Theory of Income
Fall 2018

(Fernando Alvarez)

Notes by Takuma Habu
Modifications by Simon Sangmin Oh

University of Chicago

Contents

Preface

Original Author's Note: This note is based on the Theory of Income I lectures given by Prof. Fernando Alvarez in the Fall quarter, 2017. The note contains most of the missing steps from the class notes, solutions to the exercises posed in the slides (to the best of my abilities), as well as some additional materials which should be useful to be studied in conjunction with the materials on the slides. I have reordered some topics from the class notes in a way that I think flows better.

Editor's Note: The modifications are based on the Theory of Income I lectures given by the same professor in the Fall quarter of 2018.

Part I

Welfare Theorems and Aggregation

Overview

Structure

In this part, we first review the First and the Second Welfare Theorems. In essence, the Welfare Theorems establish links between competitive equilibria and Pareto optimal allocations:

- ▷ The First Welfare Theorem states that competitive equilibrium allocations are Pareto optimal.
- ▷ The Second Welfare Theorem states that (under convexity assumptions) Pareto optimal allocations can be supported under some prices under competitive equilibria.

Second, we show that, under some regularity conditions, any Pareto optimal allocations can be obtained by maximising an object that can be interpreted as the utility function of a representative agent. We then focus on a pure exchange economy (i.e. economy without production) with differentiable utility functions, and analyse Pareto optimal, as well as competitive equilibrium allocations, and prove the welfare theorems in a different way. Finally, we discuss an aggregation result—i.e. in general, in any competitive equilibrium, the equilibrium prices and allocation depend on the distribution of wealth. However, we show that under some assumption on preferences and endowments (e.g. if preferences are homothetic and endowments are proportional), prices will be independent of the distribution of wealth.

Takeaways

Basis for solving the Planner's problem The Welfare Theorem theorems establish the basis for the approach often taken in macroeconomic models which is to solve the planner's problem—i.e. solving for the optimal allocation by maximising a represent agent's utility with respect to feasibility/market clearing condition(s)—rather than solving the agents' problem which is to maximise individual agent's utility subject to budget constraints. The planner's problem is often easier to solve as it does not involve prices. Having found a Pareto optimal allocation, we can then appeal to the Second Welfare Theorem and "decentralise" the allocation (i.e. find prices) so that it is a result of some competitive equilibria.

Basis for relying on a representative agent The aggregation results gives us the circumstances in which prices will be independent of the distribution of wealth. In other words, conditions under which we can "abstract" from heterogeneity among agents and to solve the representative agent's problem.

1 The First and Second Welfare Theorems

1.1 First Welfare Theorem

We need to place some restrictions on the form of u.

Definition 1.1. *(Local Non-Satiation)* Utility function $u^i : \mathbf{X}^i \to \mathbb{R}$ satisfies the *local non-satiation* (LNS) property if, for any $\mathbf{x} \in \mathbf{X}^i$ and any neighbourhood of $\mathbf{x}$, denote $B_\varepsilon(\mathbf{x})$, there exists $\hat{\mathbf{x}} \in B_\varepsilon(\mathbf{x})$ such that $u^i(\hat{\mathbf{x}}) > u^i(\mathbf{x})$.[1]

If an agent's utility function satisfies LNS, then there always exists another bundle that is "close" to $\mathbf{x}$ that is strictly preferred than $\mathbf{x}$. This rules out indifferences curves that are thick, and implies that agents always exhaust their income. LNS is a weaker requirement than strict increasingness as $\hat{\mathbf{x}}$ need not be greater than $\mathbf{x}$ given the definition of $B_\varepsilon(\mathbf{x})$.

Theorem 1.1. *(First Welfare Theorem) Suppose that u^i satisfies LNS for all $i \in \mathbf{I}$. Let $\{\mathbf{p}, \bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ be a competitive equilibrium. Then, $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ is a Pareto optimal allocation.*

Proof. Skipped. ■

1.2 Second Welfare Theorem

The Second Welfare Theorem is essentially the converse of the First Welfare Theorem. It says that, under certain convexity assumptions, we can find prices such that any Pareto optimal allocations is part of some competitive equilibria. Since it is often difficult to study competitive equilibrium directly (as it requires finding prices such that all markets clear), we often solve for the Pareto optimal allocation and use the Second Welfare Theorem to argue that there exists a corresponding competitive equilibrium.

For the Second Welfare Theorem to hold, we need to impose further restrictions to ensure convexity.

Definition 1.2. *(Convexity)* A function $u^i : \mathbf{X}^i \to \mathbb{R}$ is *strictly convex* if, for any $\mathbf{x}, \mathbf{x}' \in \mathbf{X}^i$ such that $\mathbf{x}' \neq \mathbf{x}$, we have

$$u^i\left(\theta\mathbf{x} + (1-\theta)\mathbf{x}'\right) > \theta u^i(\mathbf{x}) + (1-\theta)\, u^i(\mathbf{x}'),\ \forall\theta \in (0,1).$$

If u^i is strictly convex, then it is also convex (which only requires the inequalities above to hold weakly).

Definition 1.3. *(Quasiconcavity)* A function $u^i : \mathbf{X}^i \to \mathbb{R}$ is *(strictly) quasiconcave* if its upper contour set $\left\{x \in \mathbf{X}^i : u^i(\mathbf{x}) \geq u^i(\bar{\mathbf{x}})\right\}$ is (strictly) convex for all $i \in \mathbf{I}$ and all $\bar{\mathbf{x}} \in \mathbf{X}^i$. Equivalently, for any $\mathbf{x} \neq \bar{\mathbf{x}}$, we must have

$$u^i\left(\alpha\mathbf{x} + (1-\alpha)\,\bar{\mathbf{x}}\right) > \min\left\{u^i(\mathbf{x}), u^i(\bar{\mathbf{x}})\right\},\ \forall\alpha \in (0,1).$$

Definition 1.4. Suppose $\mathbf{A}, \mathbf{B} \in \mathbf{L}$. Then, the *sumset* of $\mathbf{A}$ and $\mathbf{B}$, denoted $\mathbf{A} + \mathbf{B}$, is given by the set of all sums of an element from $\mathbf{A}$ with an element from $\mathbf{B}$; i.e.

$$\mathbf{A} + \mathbf{B} \coloneqq \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in \mathbf{A}, \mathbf{b} \in \mathbf{B}\}.$$

Given two convex disjoint sets $\mathbf{A}$ and $\mathbf{B}$ in $\mathbb{R}^m$, we can always find a hyperplane (i.e. a line in $\mathbb{R}^2$, a plane in $\mathbb{R}^3$, etc.) such that A is fully contained in one side of the hyperplane, while $\mathbf{B}$ is fully contained in the other side of the hyperplane. We will use this in proving the Second Welfare Theorem.

[1] Neighbourhood is defined in the usual way: for any $\varepsilon > 0$, $B_\varepsilon(\mathbf{x}) \coloneqq \left\{\mathbf{x}' \in \mathbf{X}^i : \|\mathbf{x} - \mathbf{x}'\| < \varepsilon\right\}$, where $\|\mathbf{x}\| = \sqrt{\sum_{k=1}^m x_k^2}$ is the Euclidean norm.

Theorem 1.2. *(Separating Hyperplane Theorem) Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be subsets of* $\mathbb{R}^m$ *that are convex and disjoint (i.e.* $\mathbf{A} \cap \mathbf{B} = \emptyset$*). Then, there exists a vector* $\mathbf{p} \in \mathbb{R}^m$, $\mathbf{p} \neq \mathbf{0}$, *such that*

$$\mathbf{p} \cdot \mathbf{x} \geq \mathbf{p} \cdot \mathbf{y}, \ \forall \mathbf{x} \in A, \ \forall \mathbf{b} \in \mathbf{B}.$$

To establish the Second Welfare Theorem, we must make two additional assumptions that ensures that: (i) the marginal utility for consumers are decreasing, which ensures that utility maximisation bundle is interior; and (ii) the production function does not exhibit increasing returns to scale, which implies that marginal costs are not decreasing, so that profit maximising production level exists.

Assumption 1. *(Assumption HH) Assume that* $\mathbf{X}^i$ *are convex for all* $i \in \mathbf{I}$ *and that* $u^i : \mathbf{X}^i \to \mathbb{R}$ *are continuous and strictly quasiconcave.*

Assumption 2. *(Assumption FF) Assume that the aggregate production sumset of the economy is convex; i.e.*

$$\mathbf{Y} := \left\{ \mathbf{y} \in \mathbf{L} : \mathbf{y} = \sum_{j=1}^{J} \mathbf{y}^j, \ \mathbf{y}^j \in \mathbf{Y}^j, \ \forall j \in \mathbf{J} \right\}.$$

Theorem 1.3. *(Second Welfare Theorem) Let* $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ *be a Pareto optimal allocation. Then, there exists a price vector* $\mathbf{p}$ *such that:*

(i) all firms maximise profits such that, for all $j \in \mathbf{J}$,

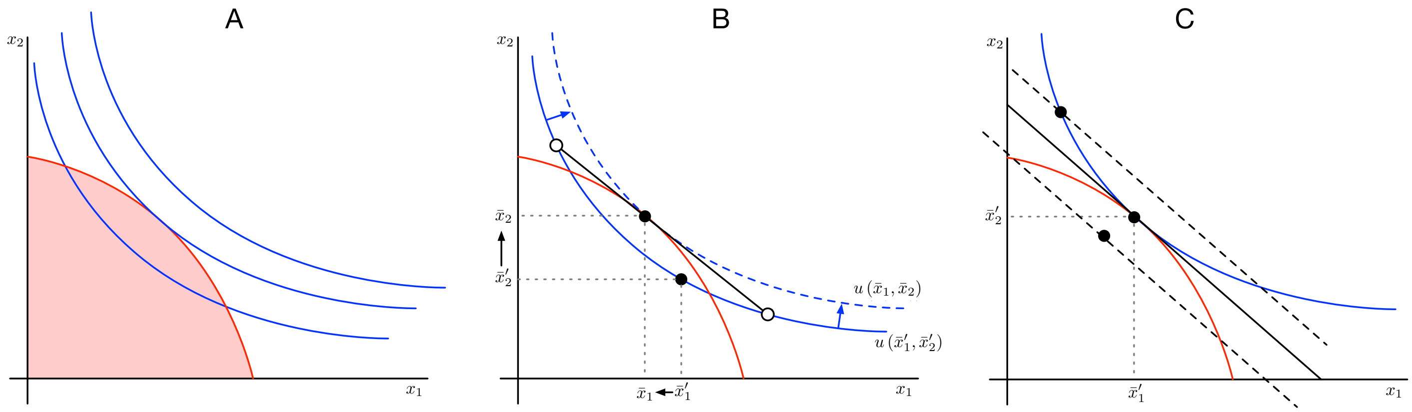
$$\mathbf{p} \cdot \bar{\mathbf{y}}^j \geq \mathbf{p} \cdot \mathbf{y}, \ \forall \mathbf{y} \in \mathbf{Y}^j;$$

(ii) given allocation $\{\bar{\mathbf{x}}^i\}$, *consumers minimise expenditure subject to attaining at least the same utility obtained by consuming* $\bar{\mathbf{x}}^i$*; i.e.*

$$\bar{\mathbf{x}}^i \in \arg\min_{\mathbf{x} \in \mathbf{X}^i} \ \mathbf{p} \cdot \mathbf{x} \quad s.t. \quad u^i(\mathbf{x}) \geq u^i\left(\bar{\mathbf{x}}^i\right).$$

Example 1.1. *(Single consumer, 2-good economy).* Before formally proving the theorem, consider the simple case of a single consumer, 2-good economy ($I = 1$ and $m = 2$) and assume that the endowment is zero.

First, under the assumptions given, the economy can be expressed as Figure A below. The red shaded region represents the aggregate production sumset of the economy, and the blue lines indicate the the consumer's indifference curves. Notice, in particular, that the production set is convex, and that the upper contour set of the indifference curves are strictly convex (i.e. utility function is strictly quasiconcave).



Consider first the consumer's problem, and the bundle $(\bar{x}_1', \bar{x}_2')$ as shown in Figure B. This bundle cannot be Pareto optimal—since u is strictly quasiconcave, we can find two points on the curve (e.g. the white circles), and a convex combination of such points so that the consumer can be on a higher difference curve. The consumer would be able to find such points unless he is at the bundle $(\bar{x}_1, \bar{x}_2)$. Thus, at any Pareto optimal allocation, the consumer's indifference curve must be tangential to the production frontier.

Is the firm's profits maximised at $(\bar{x}_1, \bar{x}_2)$? The relative prices between goods x_1 and x_2 is given by the (negative of the) slope of the indifference curve at $(\bar{x}_1, \bar{x}_2)$.[2] Given this relative price, the firm's profit is then given by the tangent line (see the solid line in Figure C). If the firm was to produce less, then the tangent line would shift down and to the left (see the lower dotted line in Figure C)—i.e. the firm would earn lower profits. Thus, the firm's profits are maximised at $(\bar{x}_1, \bar{x}_2)$.

Finally, consider whether, $(\bar{x}_1, \bar{x}_2)$ minimises the consumer's expenditures subject to attaining utility level $u(\bar{x}_1, \bar{x}_2)$. By the same argument that the tangent line represented the firm's profits, the same line also represents the consumer's expenditures. As can be seen in Figure C, if the consumer was to choose another bundle on the same indifference curve, the cost of such a bundle, represented by the dotted line above the tangent line, would be greater.

In this pedagogical example, the price vector that maximises firms profits and minimises expenditures for the consumer is given by the tangent line in the figure, which is exactly the relative price. Notice that the tangent line is simply a hyperplane that separates two convex sets: the production set and the upper contour set of the consumer's utility.

Remark 1.1. The figure would be the "same" if we were to suppose that there are two agents in the economy. The upper contour set of the consumer shown in the figure above would simply represent the sumset of the 2 consumer's upper contour sets, and the variables $\bar{x}_1$ and $\bar{x}_2$ would represent the sum of the two consumers' consumption.

Exercise. Is the hyperplane or, equivalently, the price vector always unique?

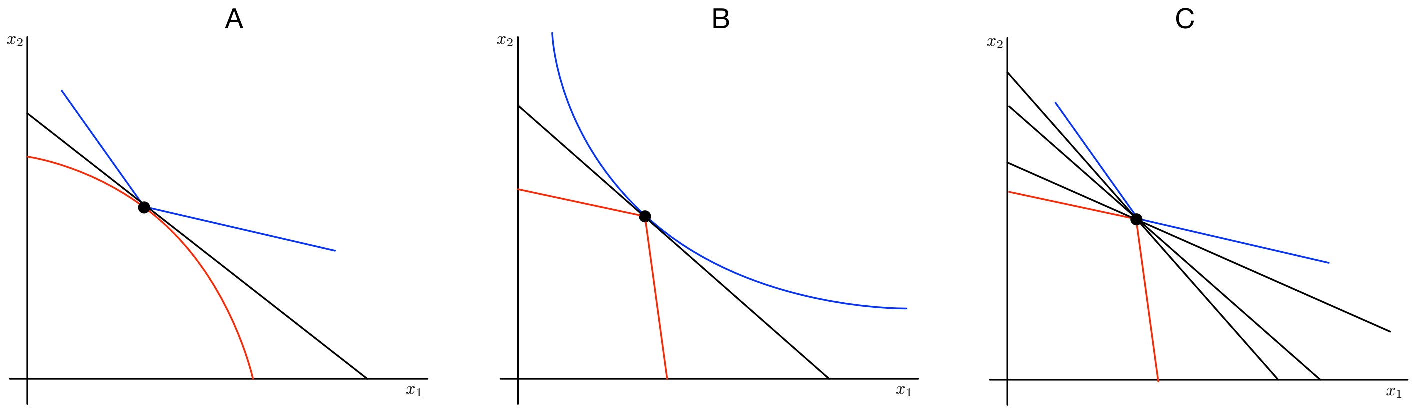
No. If, at the Pareto optimal allocation $\{\bar{x}^i, \bar{y}^j\}$, the production set *and* the upper contour sets are not differentiable (e.g. there are kinks), then there would be a continuum of price vectors that separates the two sets. See the figures below.

[2] The slope of the indifference curve is given by dx_2/dx_1, which, by the implicit function theorem, can be written as:

$$\frac{dx_2}{dx_1} = \frac{\partial u/\partial x_1}{\partial u/\partial x_2} \equiv MRS.$$

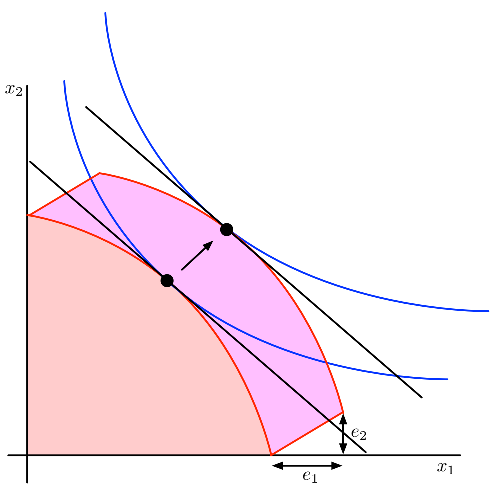
Recall that the first-order conditions for the agent's problem implies:

$$\frac{\partial u/\partial x_1}{\partial u/\partial x_2} = \frac{p_1}{p_2}.$$



Exercise. How would the analysis in the Example 1.1 change if the agent was given endowment of $e_1, e_2 > 0$?

Giving an endowment to the agent has the effect of expanding the production set, since the amount of goods available to consume in the economy is now given by the amount that firms produce plus the endowments. In the figure below, the endowment allows the consumer to reach a higher utility level. However, notice that the slope of the tangent (i.e. the relative price) is unaffected by the addition of endowments in the economy.



Proof. (Second Welfare Theorem) As was suggested by Example 1.1, the key to the proof will be the Separating Hyperplane Theorem.

Let $\mathbf{A}$ be the sumset of all the upper contour set of the households:

$$\mathbf{A} \coloneqq \left\{ \mathbf{x} \in \mathbf{L} : \mathbf{x} = \sum_{i \in \mathbf{I}} \mathbf{x}^i,\ \mathbf{x}^i \in \mathbf{X}^i,\ u^i\left(\mathbf{x}^i\right) \geq u^i\left(\bar{\mathbf{x}}^i\right),\ \forall i \in \mathbf{I} \right\}.$$

Since $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ is a Pareto optimal allocation,

$$\operatorname{int}(\mathbf{A}) \cap \mathbf{B} = \emptyset, \tag{1.1}$$

where

$$\mathbf{B} \coloneqq \mathbf{Y} + \sum_{i \in \mathbf{I}} \mathbf{e}^i.$$

That is, the interior of the sumset of upper contour sets are disjoint from the sumset of the aggregate feasible production set and the endowments. If this was not the case, then there would exist a feasible allocation (i.e. an allocation in $\mathbf{B}$) such that it is also in $\operatorname{int}(\mathbf{A})$. But if the allocation is in $\operatorname{int}(\mathbf{A})$, then it would Pareto dominate $\{\bar{\mathbf{x}}^i\}$ since the upper contour sets are convex.

Given Assumptions HH and FF, the set $\mathbf{A}$ and $\mathbf{B}$ are convex and we also have equation (1.1).[3] Therefore, by the Separating Hyperplane Theorem, there exists a vector $\mathbf{p} \neq \mathbf{0}$ such that, for all $\mathbf{x} \in \mathbf{A}$ and $\mathbf{z} \in \mathbf{B}$,

$$\mathbf{p} \cdot \mathbf{x} \geq \mathbf{p} \cdot \mathbf{z}.$$

That is, for any $\mathbf{x} \in \mathbf{A}$ and $\mathbf{y} \in \mathbf{Y}$,

$$\mathbf{p} \cdot \mathbf{x} \geq \mathbf{p} \cdot \mathbf{y} + \mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \mathbf{e}^i \right) \tag{1.2}$$

and, in particular, this must hold for $\bar{\mathbf{x}}^i$ so that

$$\mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \bar{\mathbf{x}}^i \right) \geq \mathbf{p} \cdot \mathbf{y} + \mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \mathbf{e}^i \right), \ \forall \mathbf{y} \in \mathbf{Y}. \tag{1.3}$$

Since $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ is feasible, we also have that

$$\mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \bar{\mathbf{x}}^i \right) = \mathbf{p} \cdot \left(\sum_{i \in \mathbf{J}} \bar{\mathbf{y}}^j \right) + \mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \mathbf{e}^i \right). \tag{1.4}$$

Thus, for all $\mathbf{y} \in \mathbf{Y}$,

$$\mathbf{p} \cdot \left(\sum_{i \in \mathbf{J}} \bar{\mathbf{y}}^j \right) \geq \mathbf{p} \cdot \mathbf{y}.$$

That is, firms maximise profits. If not, there would exist $\hat{\mathbf{y}} \in \mathbf{Y}^j$ such that

$$\Rightarrow \mathbf{p} \cdot \hat{\mathbf{y}} > \mathbf{p} \cdot \left(\sum_{j \in \mathbf{J}} \bar{\mathbf{y}}^j \right)$$

This implies that

$$\mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \bar{\mathbf{x}}^i \right) < \mathbf{p} \cdot \hat{\mathbf{y}} + \mathbf{p} \cdot \left(\sum_{i \in \mathbf{I}} \mathbf{e}^i \right),$$

which contradicts equation (1.3).

Similarly,

$$\bar{\mathbf{x}}^i \in \operatorname*{arg\,min}_{\mathbf{x} \in \mathbf{X}^i} \ \mathbf{p} \cdot \mathbf{x} \quad s.t. \quad u^i(\mathbf{x}) \geq u^i\left(\bar{\mathbf{x}}^i\right).$$

That is, consumers minimise their expenditures given utility level $u^i(\bar{\mathbf{x}}^i)$. If not, then there would exist $\hat{\mathbf{x}} \in \mathbf{X}^i$ such that

$$\mathbf{p} \cdot \hat{\mathbf{x}} < \mathbf{p} \cdot \bar{\mathbf{x}}^i \text{ and } u^i(\hat{\mathbf{x}}) \geq u^i\left(\bar{\mathbf{x}}^i\right).$$

[3] Take $\hat{\mathbf{x}}^i, \tilde{\mathbf{x}}^i \in \mathbf{X}^i$ such that $u^i(\hat{\mathbf{x}}^i), u^i(\tilde{\mathbf{x}}^i) \geq u^i(\bar{\mathbf{x}}^i)$. By the definition of $\mathbf{A}$, $\hat{\mathbf{x}} \coloneqq \sum_{i \in \mathbf{I}} \hat{\mathbf{x}}^i \in \mathbf{A}$ and $\tilde{\mathbf{x}} \coloneqq \sum_{i \in \mathbf{I}} \tilde{\mathbf{x}}^i \in \mathbf{A}$. Since $\mathbf{X}^i$ is convex for each $i \in \mathbf{I}$, then $\alpha \hat{\mathbf{x}}^i + (1-\alpha)\tilde{\mathbf{x}}^i \in \mathbf{X}^i$ for all $\alpha \in (0,1)$ and for all $i \in \mathbf{I}$. Moreover, since u^i is strictly quasiconcave, the set

$$u^i\left(\alpha \hat{\mathbf{x}}^i + (1-\alpha)\tilde{\mathbf{x}}^i\right) > \min\left\{u^i\left(\hat{\mathbf{x}}^i\right), u^i\left(\tilde{\mathbf{x}}^i\right)\right\} \geq u^i\left(\bar{\mathbf{x}}^i\right).$$

Hence, $\sum_{i \in \mathbf{I}} \alpha \hat{\mathbf{x}}^i + (1-\alpha)\tilde{\mathbf{x}}^i = \alpha \hat{\mathbf{x}} + (1-\alpha)\tilde{\mathbf{x}} \in \mathbf{A}$. That is, $\mathbf{A}$ is convex. That $\mathbf{B}$ is convex follows from the fact that $\mathbf{Y}$ is convex (by assumption FF) and because convexity is preserved by "addition" of a constant ($\sum_{i \in \mathbf{I}} \mathbf{e}^i$).

That is,

$$\mathbf{p}\cdot\hat{\mathbf{x}}<\mathbf{p}\cdot\left(\sum_{j\in\mathbf{J}}\bar{\mathbf{y}}^{j}\right)+\mathbf{p}\cdot\left(\sum_{i\in\mathbf{I}}\mathbf{e}^{i}\right),$$

which contradicts equation (1.3). ■

We now strengthen the notion of quasiconcavity.

Theorem 1.4. *(Arrow's Remark) Assume that u^i is strictly quasiconcave and continuous. Assume that $\bar{\mathbf{x}}^i$ is not the "cheapest" point in the budget set of household i; i.e. for all i, there exists $\tilde{\mathbf{x}}^i\in\mathbf{X}^i$ such that*

$$\mathbf{p}\cdot\tilde{\mathbf{x}}<\mathbf{p}\cdot\bar{\mathbf{x}}^{i}.$$

Then, the Pareto optimal allocation $\{\bar{\mathbf{x}}^i,\bar{\mathbf{y}}^j\}$ and the price vector $\mathbf{p}$ is a competitive equilibrium (i.e. agents not only minimise expenditures subject to attaining at $u^i(\bar{\mathbf{x}}^i)$, but they also maximise utility subject to their budget constraints).

Proof. We want to show that $\bar{\mathbf{x}}^i$, which minimises expenditure subject to attaining utility at least $u^i(\bar{\mathbf{x}}^i)$ also maximises utility subject to a budget constraint with expenditure not higher than $\mathbf{p}\cdot\bar{\mathbf{x}}^i$.

By way of contradiction, suppose that there is a $\hat{\mathbf{x}}\in\mathbf{X}^i$ such that $\mathbf{p}\cdot\hat{\mathbf{x}}\le\mathbf{p}\cdot\bar{\mathbf{x}}^i$ and $u^i(\hat{\mathbf{x}})>u(\bar{\mathbf{x}}^i)$. Then, let

$$\mathbf{x}^{\theta}\coloneqq\theta\tilde{\mathbf{x}}+(1-\theta)\,\hat{\mathbf{x}},\ \forall\theta\in(0,1)\,.$$

Note that $\mathbf{p}\cdot\tilde{\mathbf{x}}<\mathbf{p}\cdot\mathbf{x}^{\theta}<\mathbf{p}\cdot\hat{\mathbf{x}}\le\mathbf{p}\cdot\bar{\mathbf{x}}^i$. Since, $u^i(\hat{\mathbf{x}})>u^i(\bar{\mathbf{x}}^i)$, then using continuity of u^i, we have that, for all θ sufficiently close to zero,

$$u^{i}\left(\mathbf{x}^{\theta}\right)\ge u^{i}\left(\bar{\mathbf{x}}\right).$$

Moreover,

$$\begin{aligned}\mathbf{p}\cdot\mathbf{x}^{\theta}&=\theta\mathbf{p}\cdot\tilde{\mathbf{x}}+(1-\theta)\,\mathbf{p}\cdot\hat{\mathbf{x}}\\&<\theta\mathbf{p}\cdot\bar{\mathbf{x}}^{i}+(1-\theta)\,\mathbf{p}\cdot\bar{\mathbf{x}}^{i}=\mathbf{p}\cdot\bar{\mathbf{x}}^{i},\end{aligned}$$

where the inequality is strict since $\tilde{\mathbf{x}}$ is the cheapest point. Thus, $\mathbf{x}^{\theta}$ has an expenditure strictly smaller than $\bar{\mathbf{x}}^i$—a contradiction that $\bar{\mathbf{x}}^i$ minimises expenditures subject to attaining $u^i(\bar{\mathbf{x}})$. Therefore, $\bar{\mathbf{x}}^i$ must also maximise utility subject to a budget constraint. ■

When can we not the cheapest bundle? Suppose that $m=2$ and that $\mathbf{p}=(p_1,p_2)$, where $p_2=0$.

Exercise. (*Second Welfare Theorem*) To do

- ▷ Suppose that $\{\bar{\mathbf{x}}^i,\bar{\mathbf{y}}^j\}$ is a Pareto optimal allocation, and $\mathbf{p}$ is the price that (by virtue of the Second Welfare Theorem) decentralises this allocation. Describe an ownership structure—i.e. e's and θ's for which $\{\bar{\mathbf{x}}^i,\bar{\mathbf{y}}^j\}$ and $\mathbf{p}$ is a competitive equilibrium. Is there a unique one?
- ▷ Construct an example with $m=2$ where $\bar{\mathbf{x}}$ minimises expenditure subject to a given level of utility. Make sure that the utility function is quasiconcave. Display your example graphically, plotting the corresponding indifference curves and budget constraint. (Hint: Obviously, in your example the cost minimising choice must be the cheapest point. Try with one good with zero price and u increasing, but without Inada conditions.)

- ▷ What assumptions on u^i and $\mathbf{X}^i$ will be sufficient to ensure that, in any Pareto optimal allocation $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$, each agent will have a cheapest point?
- ▷ Sketch an Edgeworth box (two consumers, two goods), where a Pareto optimal allocation cannot be decentralised as a competitive equilibrium. Draw the indifference curves for one agent whose utility is not quasiconcave (i.e. the Second Welfare Theorem does not hold).
- ▷ Draw a graph of a production possibility set with one firm and two goods, consumption and labour. Assume that there is only one consumer with utility function between consumption and leisure that is quasiconcave. Draw the production possibility set that will correspond to a non-concave production function; in particular, that it will correspond to a production function that has a fixed cost and an increasing marginal cost. Make sure that your diagram is such that the Pareto optimal allocation can not be decentralised (i.e. the Second Welfare Theorem does not hold).

2 Aggregation

Definition 2.1. The *utility possibility set* $\mathbf{U}$ is defined as the set of utilities that are achievable for a feasible allocation; i.e.

$$\mathbf{U} \coloneqq \left\{ \mathbf{u} \in \mathbb{R}^I : u_i \leq u^i\left(\mathbf{x}^i\right),\ \forall i \in \mathbf{I} \text{ for some feasible allocation } \left\{\mathbf{x}^i, \mathbf{y}^j\right\} \right\}.$$

Exercise 2.1. Show that, if the aggregate possibility set $\mathbf{Y}$ and all the consumption possibility sets $\mathbf{X}^i$ are convex, then the set of feasible allocations is convex.

Solution. Recall that

$$\mathbf{Y} \coloneqq \left\{ \mathbf{y} \in \mathbf{L} : \mathbf{y} = \sum_{j \in \mathbf{J}} \mathbf{y}^j,\ \mathbf{y}^j \in \mathbf{Y}^j,\ \forall j \in \mathbf{J} \right\}.$$

The set of feasible allocation is

$$\mathbf{F} \coloneqq \left\{ \left\{ \left\{\mathbf{x}^i \in \mathbf{X}^i\right\}_{i \in \mathbf{I}}, \left\{\mathbf{y}^j \in \mathbf{Y}^j\right\}_{j \in \mathbf{J}} \right\} : \sum_{i \in \mathbf{I}} \mathbf{x}^i = \sum_{j \in \mathbf{J}} \mathbf{y}^j + \sum_{i \in \mathbf{I}} \mathbf{e}_i \right\}.$$

Take any $\hat{\mathbf{f}}, \tilde{\mathbf{f}} \in \mathbf{F}$ and construct

$$\begin{aligned} \mathbf{f}^\alpha &= \alpha \hat{\mathbf{f}} + (1-\alpha)\tilde{\mathbf{f}} \\ &= \left\{ \left\{\alpha \hat{\mathbf{x}}^i + (1-\alpha)\tilde{\mathbf{x}}^i\right\}_{i \in \mathbf{I}}, \left\{\alpha \hat{\mathbf{y}}^j + (1-\alpha)\tilde{\mathbf{y}}^j\right\}_{j \in \mathbf{J}} \right\} \end{aligned}$$

Since $\mathbf{X}^i$ is convex, $\alpha \hat{\mathbf{x}}^i + (1-\alpha)\tilde{\mathbf{x}}^i \in \mathbf{X}^i$. Similarly, since $\mathbf{Y}$ is convex $\sum_{j \in \mathbf{J}} \alpha \hat{\mathbf{y}}^j + (1-\alpha)\tilde{\mathbf{y}}^j \in \mathbf{Y}$. Finally,

$$\begin{aligned} \sum_{i \in \mathbf{I}} \left(\alpha \hat{\mathbf{x}}^i + (1-\alpha)\tilde{\mathbf{x}}^i\right) &= \alpha \sum_{i \in \mathbf{I}} \hat{\mathbf{x}}^i + (1-\alpha) \sum_{i \in \mathbf{I}} \tilde{\mathbf{x}}^i \\ &= \sum_{i \in \mathbf{I}} \mathbf{e}_i + \alpha \sum_{j \in \mathbf{J}} \hat{\mathbf{y}}^j + (1-\alpha) \sum_{j \in \mathbf{J}} \tilde{\mathbf{y}}^j \\ &= \sum_{i \in \mathbf{I}} \mathbf{e}_i + \sum_{j \in \mathbf{J}} \left(\alpha \hat{\mathbf{y}}^j + (1-\alpha)\tilde{\mathbf{y}}^j\right) \end{aligned}$$

so that feasibility is satisfied.

Assumption 3. *(Assumption CC) u^i are concave for all $i \in \mathbf{I}$.*

Exercise 2.2. Concavity and convexity.

▷ Is concavity of u^i a cardinal or an ordinal property of u^i? (recall that an ordinal property is not altered if u^i is replaced by a strictly increasing transformation, and hence only ordinal properties of u^i determine the characteristics of the demand functions).

Concavity is a cardinal property. To see why note that the function $f(x) = x^{0.5}$ is a concave function. However, suppose we transform this function using $g(x) = x^3$, which is a strictly increasing transformation (in the positive domain). Then $g(f(x)) = x^{1.5}$ but the transformed function is now convex.

▷ Show that if the convexity assumptions HH and FF are satisfied and if, furthermore, all u^i are concave (assumption CC), then U is a convex set. To do

Theorem 2.1. *Assume that u^i are strictly increasing, and that assumptions HH, CC and FF are satisfied. Then, $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ is a Pareto optimal allocation if and only if there exists a vector $\boldsymbol{\lambda} \in \mathbb{R}^l_+$ such that $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ solves the problem W:*

$$W := \max_{\{\mathbf{x}^i, \mathbf{y}^j\}} \quad \sum_{i \in \mathbf{I}} \lambda_i u^i \left(\mathbf{x}^i\right)$$
$$s.t. \qquad \left\{\mathbf{x}^i, \mathbf{y}^j\right\} \text{ is a feasible allocation.}$$

Proof. ($\Leftarrow$) We want to show that, if $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ solves the problem W, then it must be a Pareto optimal allocation. We prove this by contradiction. Suppose that $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ is not Pareto optimal, then there must be a feasible allocation $\{\hat{\mathbf{x}}^i, \hat{\mathbf{y}}^j\}$ that Pareto dominates $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ (i.e. there exists some agents who are strictly better off with other agents being at least as better off). Suppose that the weights for those who are strictly better off is positive, then

$$\sum_{i \in \mathbf{I}} \lambda_i u^i \left(\bar{\mathbf{x}}^i\right) < \sum_{i \in \mathbf{I}} \lambda_i u^i \left(\hat{\mathbf{x}}^i\right).$$

which this contradicts that $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^j\}$ solves the problem W. Thus, it must be the case that: (i) those who are strictly better off have zero weights, and (ii) those with positive weights must be equally better off under the two allocations; i.e.

$$u^i \left(\bar{\mathbf{x}}^i\right) < u^i \left(\hat{\mathbf{x}}^i\right), \ \forall i = \mathbf{I}^0 := \{i \in \mathbf{I} : \lambda_i = 0\},$$
$$u^i \left(\bar{\mathbf{x}}^i\right) = u^i \left(\hat{\mathbf{x}}^i\right), \ \forall i = \mathbf{I}^+ := \{i \in \mathbf{I} : \lambda_i > 0\}.$$

Consider a feasible allocation $\tilde{\mathbf{x}}^i$ in which we transfer some goods from those with zero weights to those with positive weights; i.e. $\tilde{\mathbf{x}}^i = \hat{\mathbf{x}}^i + \boldsymbol{\varepsilon}$ for those with $\lambda_i > 0$, and $\tilde{\mathbf{x}}^i = \hat{\mathbf{x}}^i - \boldsymbol{\varepsilon}$ for those with $\lambda_i = 0$, where $\boldsymbol{\varepsilon}$ is a positive vector. Since u^i's are strictly increasing, then

$$\begin{aligned} \sum_{i \in \mathbf{I}} \lambda_i u^i \left(\hat{\mathbf{x}}^i\right) &= \sum_{i \in \mathbf{I}^0} 0 u^i \left(\hat{\mathbf{x}}^i\right) + \sum_{i \in \mathbf{I}^+} \lambda_i u^i \left(\hat{\mathbf{x}}^i\right) \\ &< \sum_{i \in \mathbf{I}^0} 0 u^i \left(\hat{\mathbf{x}}^i - \boldsymbol{\varepsilon}\right) + \sum_{i \in \mathbf{I}^+} \lambda_i u^i \left(\hat{\mathbf{x}}^i + \boldsymbol{\varepsilon}\right) \\ \Rightarrow \sum_{i \in \mathbf{I}} \lambda_i u^i \left(\bar{\mathbf{x}}^i\right) &< \sum_{i \in \mathbf{I}} \lambda_i u^i \left(\tilde{\mathbf{x}}^i\right), \end{aligned}$$

which contradicts the assumption that $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ solves the problem W.

($\Rightarrow$) We want to show that, if $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ is a Pareto optimal allocation, then it must solve the problem W for some weights $\boldsymbol{\lambda}$. To see this, define the sets

$$\begin{aligned} \mathbf{A} &:= \left\{\bar{\mathbf{u}} \in \mathbb{R}^I : \ \bar{u}_i = u^i \left(\bar{\mathbf{x}}^i\right), \ \forall i \in \mathbf{I}\right\}, \\ \mathbf{B} &:= \operatorname{int}\left(\mathbf{U}\right) \\ &= \left\{\mathbf{u} \in \mathbb{R}^I : u_i < u^i \left(\mathbf{x}^i\right), \ \forall i \in \mathbf{I} \text{ for some feasible allocation } \left\{\mathbf{x}^i, \mathbf{y}^j\right\}\right\}. \end{aligned}$$

Then, $\mathbf{A}$ is a singleton (since u^i is strictly increasing) with the utility values of the Pareto optimal allocation, and $\mathbf{B}$ contains a set of vectors of utility levels that are strictly lower than some feasible allocation. Since $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ is a Pareto optimal allocation, then it must be that

$$\mathbf{A} \cap \mathbf{B} = \emptyset.$$

Under assumptions HH, CC and FF, both **A** and **B** are convex. Thus, by the Separating Hyperplane Theorem, there exists a vector $\boldsymbol{\lambda}^* \in \mathbb{R}^l$, $\boldsymbol{\lambda}^* \neq 0$ such that

$$\boldsymbol{\lambda}^* \cdot \bar{\mathbf{u}} \geq \boldsymbol{\lambda}^* \cdot \mathbf{u}, \ \forall \mathbf{u} \in \mathbf{B}.$$

Thus, by taking limits to all the elements in the closure of **B**, which is **U**, we have that:

$$\boldsymbol{\lambda}^* \cdot \bar{\mathbf{u}} \geq \boldsymbol{\lambda}^* \cdot \mathbf{u}, \ \forall \mathbf{u} \in \mathbf{U} \tag{2.1}$$

so that $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ solves the problem W. It remains to show that we can choose the weights to be nonnegative.

Denote the set of agents for which $\lambda_i^* < 0$ by $\mathbf{I}^-$. Thus, given the inequality (2.1), the utility of agents in the set $\mathbf{I}^-$ is being minimised. Now, consider the set of weights, $\lambda_i = \max\{0, \lambda_i^*\}$ for all $i \in I$. Since, by assumption, u^i's are strictly increasing, then

$$\boldsymbol{\lambda} \cdot \bar{\mathbf{u}} \geq \boldsymbol{\lambda} \cdot \mathbf{u}, \ \forall \mathbf{u} \in \mathbf{U}.$$

Thus, $\{\bar{\mathbf{x}}^i, \bar{\mathbf{y}}^i\}$ solves the problem W for nonnegative weights $\boldsymbol{\lambda}$. ■

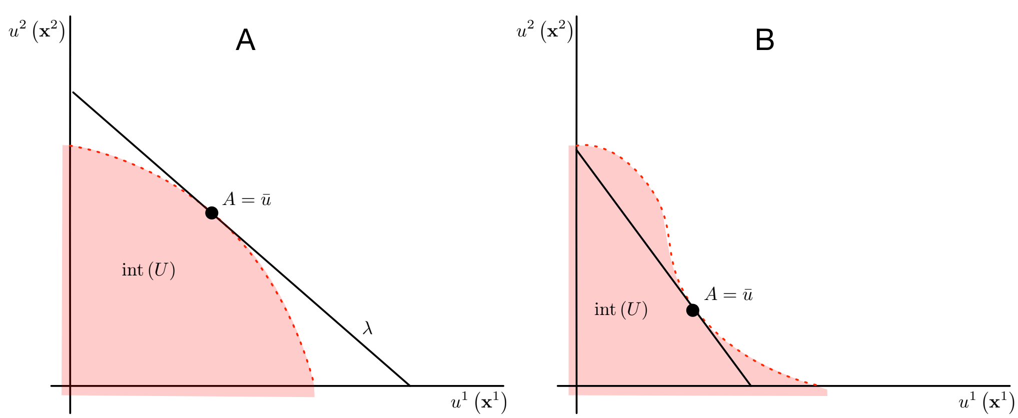
Exercise 2.3. Exercises:

- ▷ Draw a picture of a convex utility possibility set U for $I = 2$ with strictly increasing preferences. Pick a particular point in its frontier, say $\bar{u}$. Show, graphically, how to locate a vector λ such that the solution of the problem W for such λ is $\bar{u}$.

  See panel A.

- ▷ Draw a picture of a non-convex utility possibility set for U for $I = 2$. Pick a particular point in its frontier, say $\bar{u}$. Draw U and pick $\bar{u}$ in such a way that $\bar{u}$ does not solve the problem W, regardless of what λ is chosen. Can this happen if all u^i are not convex, but they are strictly quasiconcave?

  See panel B.



By the virtue of the previous theorem, as well as the Welfare Theorems, we often analyse an economy with only one agent—the representative agent. The representative agent has utility u defined as

$$u(\mathbf{x}) := \max_{\mathbf{x}^i \in \mathbf{X}^i} \sum_{i \in \mathbf{I}} \lambda_i u^i(\mathbf{x}^i)$$
$$s.t. \quad \sum_{i \in \mathbf{I}} \mathbf{x}^i = \mathbf{x}.$$

Clearly, this utility function depends on the weights $\boldsymbol{\lambda}$.

By the First Welfare Theorem, we know that competitive equilibrium allocations are Pareto optimal. Hence, we can find weights $\boldsymbol{\lambda}$ for which an economy with one representative agent with utility function u corresponds to the one with many agents. Analogously, a Pareto optimal allocation in an economy with many agents has a corresponding competitive equilibrium for some ownership structure. This competitive equilibrium corresponds to one for an economy with one representative agent for some weight $\boldsymbol{\lambda}$.

Since the definition of the representative agent does not involve the production side of the economy (i.e. it does not involve $\mathbf{Y}^j$), we will now study pure exchange economies; i.e. economies without production.

2.1 Pure Exchange Economy

We now consider the case of a pure exchange economy with smooth (i.e. differentiable) concave utility functions. This will give a precise interpretation to the weights $\boldsymbol{\lambda}$ and connect the Welfare Theorems with the utility of the representative agent.

Definition 2.2. An *exchange economy* is one where the aggregate production possibility set is given by the negative orthant, so that the agent can dispose of goods, but no goods are produced.

Feasibility is then given by, for any $\mathbf{x}^i \in \mathbf{X}^i$ for all i,

$$\sum_{i \in \mathbf{I}} \mathbf{x}^i \leq \sum_{i \in \mathbf{I}} \mathbf{e}^i := \bar{\mathbf{e}},$$

where $\bar{\mathbf{e}} \in \mathbb{R}^m_+ \subseteq \mathbf{L}$ denotes the *aggregate endowment*.

2.1.1 Pareto optimal allocations (Planner's problem)

The planner's problem is the problem W:

$$\begin{aligned} W := \max_{\mathbf{x}^i \in \mathbf{X}^i} \quad & \sum_{i \in \mathbf{I}} \lambda_i u^i\left(\mathbf{x}^i\right) \\ s.t. \quad & \sum_{i \in \mathbf{I}} \mathbf{x}^i \leq \bar{\mathbf{e}}. \end{aligned}$$

Notice, in particular, that the planner's problem does not involve prices of goods. The Lagrangian is given by

$$\begin{aligned} \mathcal{L} &= \sum_{i \in \mathbf{I}} \lambda_i u^i\left(\mathbf{x}^i\right) + \boldsymbol{\gamma} \cdot \left(\bar{\mathbf{e}} - \sum_{i \in \mathbf{I}} \mathbf{x}^i\right) \\ &= \sum_{i \in \mathbf{I}} \lambda_i u^i\left(\mathbf{x}^i\right) + \sum_{\ell \in L} \gamma_\ell \left(e_l - \sum_{i \in I} x^i_\ell\right) \end{aligned}$$

Under assumptions HH and CC, the necessary and sufficient first-order conditions for the problem W can be characterised as

$$\lambda_i \frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)}{\partial x^i_\ell} = \gamma_\ell,\ \forall \ell = 1, 2, \ldots m,\ \forall i \in \mathbf{I} \tag{2.2}$$

where $\gamma_\ell \geq 0$ are the Lagrange multipliers of the feasibility constraint of the problem W. This implies that the marginal rate of substitutions are the same for all agents for each pair of goods:

$$\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)/\partial x_\ell^i}{\partial u^i\left(\bar{\mathbf{x}}^i\right)/\partial x_k^i} = \frac{\partial u^j\left(\bar{\mathbf{x}}^j\right)/\partial x_\ell^j}{\partial u^j\left(\bar{\mathbf{x}}^j\right)/\partial x_k^j},\ \forall \ell, k \in \{1,2,\ldots,m\},\ \forall i,j \in \mathbf{I}.$$

Notice, in particular, that this expression does not contain the weights $\boldsymbol{\lambda}$.

In fact, $\boldsymbol{\lambda}$ determines the consumption across agents for each product

$$\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)/\partial x_\ell^i}{\partial u^j\left(\bar{\mathbf{x}}^j\right)/\partial x_\ell^j} = \frac{\lambda_j}{\lambda_i},\ \forall \ell \in \{1,2,\ldots,m\},\ \forall i,j \in \mathbf{I}.$$

Since the first-order conditions are necessary and sufficient, they characterise (all of) the set of Pareto optimal allocations. Furthermore, the vector $\boldsymbol{\gamma}$ has the interpretation of the marginal value of an extra unit of the aggregate endowment. As can be seen, both the solution to the problem W, as well as the vector $\boldsymbol{\gamma}$, are a function of $\boldsymbol{\lambda}$. Thus, the marginal value of an extra unit of the aggregate endowment, $\boldsymbol{\gamma}$, depends on the weights $\boldsymbol{\lambda}$.

2.1.2 Competitive equilibrium

The maximisation of the household's problem is the key part of the competitive equilibrium for the pure endowment case. The household i's problem is:

$$\max_{\mathbf{x}^i \in \mathbf{X}^i} \quad u^i\left(\mathbf{x}^i\right) \quad s.t. \quad \mathbf{p}\cdot\mathbf{x}^i \leq \mathbf{p}\cdot\mathbf{e}^i.$$

The Lagrangian is then

$$\mathcal{L} = u^i\left(\mathbf{x}^i\right) + \mu_i\left(\mathbf{p}\cdot\mathbf{e}^i - \mathbf{p}\cdot\mathbf{x}^i\right).$$

When assumptions HH and CC hold, this maximisation problem is characterised by the following necessary and sufficient conditions:

$$\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)}{\partial x_\ell} = \mu_i p_\ell,\ \forall \ell \in \{1,2,\ldots,m\}, \tag{2.3}$$

From the Envelope Theorem, we know that μ_i measures the change in maximised agent's utility if the agent's income is increased by one unit of the numeraire good, so μ_i is the marginal utility of income.[4]

These first-order conditions imply that the marginal rate of substitution is equated across agents; i.e.

$$\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)/\partial x_\ell}{\partial u^i\left(\bar{\mathbf{x}}^i\right)/\partial x_k} = \frac{\partial u^j\left(\bar{\mathbf{x}}^j\right)/\partial x_\ell}{\partial u^j\left(\bar{\mathbf{x}}^j\right)/\partial x_k},\ \forall \ell, k \in \{1,2,\ldots,m\},\ \forall i,j \in \mathbf{I}.$$

The multipliers μ_i do not enter in this expression. The relative consumption of agents i and j depend, however, on the multipliers

$$\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)/\partial x_\ell}{\partial u^j\left(\bar{\mathbf{x}}^j\right)/\partial x_\ell} = \frac{\mu_i}{\mu_j},\ \forall \ell \in \{1,2,\ldots,m\},\ \forall i,j \in \mathbf{I}.$$

As the Welfare Theorems show in the general case, there is a close connection between competitive equilibrium and Pareto optimal allocations.

[4]Denote $\mathcal{L}^*$ as the Lagrangian evaluated at the optimal $\mathbf{x}^i$. Notice that, by the Envelope Theorem, μ_i is the partial derivative of $\mathcal{L}^*$ with respect to $\mathbf{p}\cdot\mathbf{e}^i$. Adding one unit of the numeraire good is equivalent to $\mathbf{p}\cdot\mathbf{e}^i + 1$.

2.1.3 Analogy to the First Welfare Theorem

We first analyse the analog to the First Welfare Theorem. Given a competitive equilibrium $\{\bar{\mathbf{x}}^i, \mathbf{p}\}$, the household i's problem satisfies the necessary first-order conditions stated using the multiplier μ_i and prices $\mathbf{p}$. We will use them to find the $\boldsymbol{\lambda}$-weights and the Lagrange multipliers $\boldsymbol{\gamma}$ of the corresponding planning problem with solution $\{\bar{\mathbf{x}}^i\}$.

Multiplying the first-order condition with respect to good ℓ from the household's problem, (2.3), by $1/\mu_i$ yields

$$\frac{1}{\mu_i}\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)}{\partial x_\ell} = p_\ell, \;\forall \ell \in \{1, 2, \ldots, m\}\,.$$

Similar conditions holds for all households $i \in \mathbf{I}$.

This is identical to the first-order condition, (2.2), from the planner's problem with weights given by

$$\lambda_i = \frac{1}{\mu_i}$$

and with the Lagrange multiplier for good ℓ equal to the price

$$\gamma_\ell = p_\ell.$$

Since the first-order conditions for the planner's problem are sufficient, then the competitive equilibrium $\{\bar{\mathbf{x}}^i\}$ is indeed a Pareto optimal allocation.

We see that agents with low marginal utility of their income (μ_i)—i.e. those consuming a lot given the concavity of u^i—are those that are assigned high $\boldsymbol{\lambda}$-weights. Likewise, goods with high prices p_ℓ are those whose marginal social value – or in other words, the increase in utility from an extra unit of the endowment of good ℓ – (γ_ℓ) is high.

2.1.4 Analogy to the Second Welfare Theorem

Let us start with a Pareto optimal allocation $\{\bar{\mathbf{x}}^i\}$ or, equivalently, with an allocation and Lagrange multipliers $\boldsymbol{\gamma}$ satisfying the necessary first-order conditions, (2.2), of the problem W for some weights $\boldsymbol{\gamma}$. We must now show that the first-order conditions for the household problem hold for some choice of μ_i and $\mathbf{p}$.

Set $\mathbf{p} = \boldsymbol{\gamma}$ and the marginal utility of income so that $\mu_i = 1/\lambda_i$. Substituting these into the (necessary) first-order conditions for the planner's problem, (2.2), yields

$$\frac{1}{\mu_i}\frac{\partial u^i\left(\bar{\mathbf{x}}^i\right)}{\partial x_\ell} = p_\ell, \;\forall \ell, \{1, 2, \ldots, m\}\,, \;\forall i \in \mathbf{I}.$$

We set the endowments $\mathbf{e}^i = \bar{\mathbf{x}}^i$ for all $i \in \mathbf{I}$ so that $\bar{\mathbf{x}}^i$ is budget feasible for each agent $i \in I$. Since these first-order conditions are sufficient for the household's problem, and the allocation $\bar{\mathbf{x}}^i$ is budget feasible, then the agent maximises utility by choosing $\bar{\mathbf{x}}^i$.

Since the allocation is feasible, we have shown that, indeed, these prices and endowments constitute a competitive equilibrium.

Notice that those agents with high λ_i-weight are the agents with low marginal utility of income μ_i; i.e. those with high consumption.

2.1.5 Aggregation

In general, in a competitive equilibrium, the equilibrium price $\mathbf{p}$ and the equilibrium allocation $\{\bar{\mathbf{x}}^i\}$ depend on the income distribution measured by the vector of individual endowments $\{\mathbf{e}^i\}$. Similarly, in general, the shadow value of social endowment, $\boldsymbol{\gamma}$, and the Pareto optimal allocation $\{\bar{\mathbf{x}}^i\}$ depend on the $\boldsymbol{\lambda}$-weights.

However, for particular class of utility functions and endowments (equivalently, $\boldsymbol{\lambda}$-weights) the equilibrium prices $\mathbf{p}$ (equivalently, the shadow value of social endowment $\boldsymbol{\gamma}$) do not depend on the distribution of the endowment $\left\{\mathbf{e}^i\right\}$ (equivalently, on the distribution of the Pareto weights $\boldsymbol{\lambda}$).

Definition 2.3. *(Aggregation)* We say that there is *aggregation* if the equilibrium price $\mathbf{p}$ (equivalently, the shadow value of social endowment $\boldsymbol{\gamma}$) do not depend on the *distribution* of endowment/income $\left\{\mathbf{e}^i\right\}$ (equivalently on the *distribution* of the Pareto weights $\boldsymbol{\lambda}$).

Fact 2.1. *When we have aggregation, we can treat each agent as a representative agent.*

We provide an example in which there is aggregation.

Definition 2.4. *(Homotheticity)* A homothetic function is a monotonic transformation of a function which is homogenous; i.e. u^i is homothetic if

$$u^i\left(\mathbf{x}\right) = g^i\left(h\left(\mathbf{x}\right)\right),$$

where g^i is a strictly increasing concave function and where h is homogenous of degree one, i.e.

$$h\left(\eta\mathbf{x}\right) = \eta h\left(\mathbf{x}\right),\ \forall\eta > 0,\ \forall\mathbf{x},$$

and independent of i.

Proposition 2.1. *Suppose utility functions are homothetic and endowments are proportional so that there are some positive fractions $\delta^i > 0$ such that $\mathbf{e}^i = \delta^i\bar{\mathbf{e}}$ for all $i \in I$. Then, we have aggregation. Specifically, the competitive equilibrium allocation (and the Pareto optimal allocation) are equal to their endowment*

$$\bar{\mathbf{x}}^i = \delta^i\bar{\mathbf{e}},\ \forall i \in \mathbf{I},$$

and the equilibrium prices are given by

$$p_\ell = \kappa\frac{\partial h\left(\bar{\mathbf{e}}\right)}{\partial x_\ell},\ \forall\ell \in \{1, 2, \ldots, m\} \tag{2.4}$$

and where $\kappa > 0$ is an arbitrary positive constant. Observe that neither $\mathbf{p}$ nor $\{\bar{\mathbf{x}}^i\}$ depend on the distribution of $\boldsymbol{\lambda}$-weights..

Proof. The Lagrangian for the consumer's problem is:

$$\mathcal{L} = g^i\left(h\left(\mathbf{x}^i\right)\right) + \mu_i\left(\mathbf{p}\cdot\mathbf{e}^i - \mathbf{p}\cdot\mathbf{x}^i\right).$$

The first-order conditions for the agent (2.3) is given by

$$\frac{\partial g^i\left(h\left(\mathbf{x}^i\right)\right)}{\partial h}\frac{\partial h\left(\mathbf{x}^i\right)}{\partial x_\ell} = \mu_i p_\ell,\ \forall\ell \in \{1, 2, \ldots, m\},\ \forall i \in \mathbf{I}.$$

We wish to verify that the above first-order condition is satisfied when $\bar{\mathbf{x}}^i = \delta^i\bar{\mathbf{e}}$ and prices are given by (2.4). Let us evaluate first the left-hand side setting $\mathbf{x}^i = \delta^i\bar{\mathbf{e}}$:

$$\frac{\partial g^i\left(h\left(\delta^i\bar{\mathbf{e}}\right)\right)}{\partial h}\frac{\partial h\left(\delta^i\bar{\mathbf{e}}\right)}{\partial x_\ell} = \frac{\partial g^i\left(h\left(\delta^i\bar{\mathbf{e}}\right)\right)}{\partial h}\frac{\partial h\left(\bar{\mathbf{e}}\right)}{\partial x_\ell},$$

where we used the fact that h is homogenous of degree zero (partial derivatives of function that are homogenous of degree one are homogenous of degree zero). Given (2.4), the first-order condition is satisfied if

$$\begin{aligned}
\mu_i &= \frac{1}{p_\ell}\frac{\partial g^i\left(h\left(\delta^i\bar{\mathbf{e}}\right)\right)}{\partial h}\frac{\partial h\left(\bar{\mathbf{e}}\right)}{\partial x_\ell} \\
&= \frac{1}{\kappa}\frac{1}{\partial h\left(\bar{\mathbf{e}}\right)/\partial x_\ell}\frac{\partial g^i\left(h\left(\delta^i\bar{\mathbf{e}}\right)\right)}{\partial h}\frac{\partial h\left(\bar{\mathbf{e}}\right)}{\partial x_\ell} \\
&= \frac{1}{\kappa}\frac{\partial g^i\left(h\left(\delta^i\bar{\mathbf{e}}\right)\right)}{\partial h},\ \forall i \in \mathbf{I}.
\end{aligned}$$

The social shadow value of the aggregate endowment in this case is

$$\gamma_\ell = p_\ell = \kappa\frac{\partial h\left(\bar{\mathbf{e}}\right)}{\partial x_\ell},\ \forall i \in \mathbf{I}$$

for some arbitrary constant $\kappa > 0$. The corresponding $\boldsymbol{\lambda}$-weights are

$$\lambda_i = \frac{1}{\mu_i} = \frac{\kappa}{\partial g^i\left(h\left(\delta^i\bar{\mathbf{e}}\right)\right)/\partial h},\ \forall i \in \mathbf{I}. \qquad \blacksquare$$

We therefore find that the equilibrium prices $\mathbf{p}$ are independent of the distribution of the endowments parameterised by δ^i and the shadow value of the endowment $\boldsymbol{\gamma}$ is independent of the $\boldsymbol{\lambda}$-weights. In fact, the $\boldsymbol{\lambda}$-weights of the Pareto optimal allocation that corresponds to the competitive equilibrium depend upon the income distribution given by $\boldsymbol{\delta}$. Notice that larger δ^i implies larger λ_i (since $\partial g/\partial h$ is decreasing and h is increasing).

Example 2.1. Suppose, for each $i \in \mathbf{I}$,

$$u^i\left(\mathbf{x}^i\right) = \sum_{\ell=1}^m \beta_\ell^i \log\left(x_i^i - \theta_\ell^i\right),\ \sum_{\ell=1}^m \beta_l^i = 1,\ \beta_\ell^i \geq 0,\ \forall \ell \in \{1, 2, \ldots, m\}.$$

The Lagrangian for the planner's problem is

$$\mathcal{L} = \sum_{i\in\mathbf{I}} \lambda_i \sum_{\ell=1}^m \beta_\ell^i \log\left(x_\ell^i - \theta_\ell^i\right) + \boldsymbol{\gamma}\cdot\left(\bar{\mathbf{e}} - \sum_{i\in\mathbf{I}}\mathbf{x}^i\right).$$

The first-order condition gives

$$\lambda_i\frac{\beta_\ell^i}{x_\ell^i - \theta_\ell^i} = \gamma_\ell \Leftrightarrow \lambda_i\beta_\ell^i = \gamma_\ell\left(x_\ell^i - \theta_\ell^i\right)$$

Summing across $i \in \mathbf{I}$,

$$\begin{aligned}
\sum_{i\in\mathbf{I}}\lambda_i\beta_\ell^i &= \sum_{i\in\mathbf{I}}\gamma_\ell\left(x_\ell^i - \theta_\ell^i\right) \\
\Leftrightarrow \gamma_\ell &= \frac{\sum_{i\in\mathbf{I}}\lambda_i\beta_\ell^i}{\sum_{i\in\mathbf{I}}\left(x_\ell^i - \theta_\ell^i\right)} = \frac{\sum_{i\in\mathbf{I}}\lambda_i\beta_\ell^i}{\bar{e}_\ell - \sum_{i\in\mathbf{I}}\theta_\ell^i}.
\end{aligned}$$

Suppose that $\lambda_1 = 1$ and $\lambda_i = 0$ for all $i \neq 1$. Then,

$$\gamma_\ell = \frac{\lambda_1\beta_\ell^1}{\bar{e}_\ell - \sum_{i\in\mathbf{I}}\theta_\ell^i}.$$

If, instead, $\lambda_2 = 1$ and $\lambda_i = 0$ for all $i \neq 1$, then

$$\gamma'_\ell = \frac{\lambda_2 \beta_\ell^2}{\bar{e}_\ell - \sum_{i\in\mathbf{I}} \theta_\ell^i}$$

Since

$$\gamma_\ell \neq \gamma'_\ell,$$

we realise that the shadow value of aggregate endowment depends on the distribution of $\boldsymbol{\lambda}$. Since $\gamma_\ell = p_\ell$, we do not have aggregation here.

Suppose now that $\beta_\ell^i = \beta_\ell$ for all $i \in \mathbf{I}$. Then,

$$\gamma_\ell = \frac{\beta_\ell}{\bar{e}_\ell - \sum_{i\in\mathbf{I}} \theta_\ell^i}\left(\sum_{i\in\mathbf{I}} \lambda_i\right).$$

Hence, only the sum of elements of $\boldsymbol{\lambda}$ affects γ_ℓ and not the distribution of $\boldsymbol{\lambda}$. Thus, in this case, we have aggregation.

We know that only relative prices are determined in equilibrium. Hence, we can consider the ratio

$$\frac{\gamma_\ell}{\gamma_k} = \frac{p_\ell}{p_k}$$

and check whether the ratio depends on λ_i. Taking the case when $\beta_\ell^i = \beta_\ell$ for all $i \in \mathbf{I}$, observe that

$$\frac{\gamma_\ell}{\gamma_k} = \frac{\beta_\ell}{\beta_k}\frac{\bar{e}_\ell - \sum_{i\in\mathbf{I}} \theta_k^i}{\bar{e}_\ell - \sum_{i\in\mathbf{I}} \theta_\ell^i}$$

and we see that the ratio does not involve λ_i's.

Remark 2.1. Steps for checking for aggregation:

- ▷ Write down the planner's problem (maximise weighted utility subject to feasibility) and derive the first-order conditions.
- ▷ Sum across agents the first-order condition to get an expression for γ_l in terms of aggregate endowment, $\bar{e}_\ell$, and $\boldsymbol{\lambda}$-weights.
- ▷ Take the ratio of γ_l's and see if it depends on $\boldsymbol{\lambda}$-weights.

2.1.6 Examples from Problem Set 2

- ▷ Examples of no aggregation
  - ▷ Q1.1 Homothetic and non-identical preferences:

$$u^i\left(\mathbf{x}^i\right) = \sum_{\ell=1}^m \beta_\ell^i \log\left(x_\ell^i\right), \ \sum_{\ell=1}^m \beta_\ell^i = 1, \ \beta_\ell^i \geq 0, \ \forall \ell \in \{1, 2, \ldots, m\}.$$

    If $\beta_\ell^i = \beta_\ell$ for all ℓ, then we would have aggregation.

  - ▷ Q1.4 Identical and non-homothetic preferences, $\rho \neq \beta$,

$$u^i\left(\mathbf{x}^i\right) = \frac{x_1^{1-\rho}}{1-\rho} + \frac{x_2^{1-\beta}}{1-\beta}.$$

If $\beta = \rho$ then we would have aggregation.

- Examples of aggregation:
    - Q1.2 Non-identical and non-homothetic preferences (Stone-Geary):

$$u^i\left(\mathbf{x}^i\right) = \sum_{\ell=1}^{m} \beta_\ell \log\left(x_\ell^i - \theta_\ell^i\right).$$

    - Q1.3 Non-identical and non-homothetic preferences (quadratic Stone-Geary):

$$u^i\left(\mathbf{x}^i\right) = -\frac{1}{2}\sum_{l=1}^{m} \beta_\ell \left(x_\ell^i - \theta_\ell^i\right)^2.$$

    - Q1.5 Non-identical and homothetic preferences

$$u^i\left(\mathbf{x}^i\right) = \sum_{\ell=1}^{m} \beta_\ell \frac{\left(x_\ell^i - \theta_\ell^i\right)^{1-\sigma}}{1-\sigma}.$$

Part II

Overlapping Generations and Perpetual Youth Models

Overview

In this part, we first review the Overlapping Generations (OLG) model. The model assumes that each agent (or generation/cohort) lives for finite (two) periods of time, but there are infinitely many generations. We will establish that, in this setup, competitive equilibrium need not be efficient so that there is scope for government intervention to improve welfare in the form of a pay-as-you-go social security system.

In the OLG model, it is implicit that a single period lasts for decades. We therefore go on to study the Perpetual Youth model in which each period can be thought of as being around a quarter/year. The model introduces individual uncertainty as to when agents die, although, in aggregate, there is no uncertainty. This model allows a practical way to study the impact of debt and government spending on the economy (they both crowd out capital).

Takeaway

Competitive equilibria can be Pareto inefficient In the OLG model, that there are infinitely many goods and generations leads to the failure of the First Welfare Theorem when interest rates are negative (or "low" if we build in population/productivity growth). In the case of the Perpetual Youth model, inefficiency (called *dynamic inefficiency*) can also arise when income is falling over time for each generation.

Phase diagrams to study dynamics In studying the dynamics of the steady state (e.g. whether the steady state is stable/unstable etc.), it is generally easier to draw phase diagrams. We do so in the context of the Perpetual Youth model.

3 Overlapping Generations Economy

In this section, we study a version of the Overlapping Generations (OLG) economy without production. It is divided into four parts:

(i) The first section analyses an endowment OLG economy. It describes the model in detail and argues that, by way of an example, the First Welfare Theorem (as stated so far) does not always hold. Importantly, it is shown that the First Welfare Theorem does not hold if the equilibrium interest rate is negative.

(ii) Since the competitive equilibrium need not be efficient, it is feasible for the government to improve upon the competitive equilibrium allocation. In the second section, we analyse the benefit of introducing a pay-as-you-go social security system. It is shown that in cases where the competitive equilibrium fails to be efficient, an appropriate social security system could improve welfare.

(iii) The third section discusses why the First Welfare Theorem fails to hold in general. It is shown that a key step in the proof of the First Welfare Theorem breaks down when we try to compute the value of the aggregate consumption plan in the OLG model.

(iv) The last section introduces population and productivity growth. As in the standard model, it is shown that competitive equilibrium may fail to be efficient. With population growth, the condition for inefficiency remains that the interest rate must be negative. In a growing economy, however, the condition for inefficiency arises even if the interest rate is positive but sufficiently low.

3.1 Simple Overlapping Generations (OLG) Economy

3.1.1 The set up

Consider an economy with a unique consumption good in each time period $t = 1, 2, \ldots$ so that $\mathbf{L} = \mathbb{R}^\infty$. This economy is populated by agents who live and consume for only two periods. Each generation is indexed by the time of their birth so that the set of agent is $\mathbf{I} = \{0, 1, 2, \ldots\}$. Consumption by generation i in period t is written as x_t^i. There is a unit mass of agents in each generation. Agents born at dates $t \geq 1$ (i.e. generation $i \geq 1$) have the following utility function:

$$u^i\left(\mathbf{x}^i\right) = v^i\left(x_i^i, x_{i+1}^i\right),$$

where $\mathbf{x}^i = \left(x_1^i, x_2^i, \ldots\right)$. Therefore, each generation i care only about the consumption in periods in which they are alive (i.e. $t = i, i+1$). This implies that, in any competitive equilibrium or efficient allocation for all agents born at $t = 0, 1, 2, \ldots$, generation i's consumption in periods for which they are not alive should be zero: i.e.

$$x_t^i = 0, \ \forall t \neq i, i+1.$$

Agents born at date 0 (generation 0) are assumed to be different in that they are already "old" at time $t = 1$, so that the agents can only consume goods at date $t = 1$; i.e.

$$u^0\left(x_1^0, x_2^0, \ldots\right) = x_1^0.$$

Given this, in any competitive equilibrium or efficient allocation, we must have $x_t^0 = 0$ for all $t = 2, 3, \ldots$.

Generation $i \geq 1$ have positive endowments while they are alive; i.e.

$$\mathbf{e}^i = \left(e_1^i, e_2^i, \ldots, e_i^i, e_{i+1}^i, \ldots\right) = \left(0, \ldots, e_i^i, e_{i+1}^i, \ldots 0\right),$$

where $e_i^i, e_{i+1}^i > 0$ so that $e_t^i = 0$ for all $t \neq i, i+1$. Generation 0 is again different in that the agent only has endowment in $t = 1$; i.e. $e_1^0 > 0$ and e_t^0 for all $t = 2, 3, \ldots$.

To simplify, we consider a pure exchange economy (i.e. no production) so that there are no firms (see Problem Set 1, Q2 and 3). In this case, feasible allocations must satisfy:

$$\sum_{i \in \mathbf{I}} \mathbf{x}^i = \sum_{i \in \mathbf{I}} \mathbf{e}^i,$$

which is also the market clearing condition. Since in each time period t, only agents born at t and $t-1$ are alive, we can write the feasibility constraint (or period-t marketing clearing condition) as

$$x_t^{t-1} + x_t^t = e_t^{t-1} + e_t^t := \bar{e}_t, \ \forall t = 1, 2, \ldots.$$

We refer to $\bar{e}_t$ as the time-t aggregate endowment.

3.1.2 Competitive equilibrium

Let $\mathbf{p} \in \mathbb{R}^\infty$ be the price vector for the consumption good in each period.

Definition 3.1. *(Competitive equilibrium)* A competitive equilibrium is a price vector $\mathbf{p}$ and a feasible allocation $\{\mathbf{x}^i\}_{i=1}^\infty$ such that each agent maximises utility subject to its budget constraint:[5]

- ▷ Each generation $i \geq 1$'s utility maximisation problem is given by

$$\max_{x_i^i, x_{i+1}^i \in \mathbb{R}} \quad u^i\left(x_i^i, x_{i+1}^i\right)$$
$$s.t. \quad p_i x_i^i + p_{i+1} x_{i+1}^i \leq p_i e_i^i + p_{i+1} e_{i+1}^i$$

- ▷ Generation 0's problem is given by

$$\max_{x_1^0 \in \mathbb{R}} \quad x_1^0$$
$$s.t. \quad p_1 x_1^0 \leq p_1 e_1^0.$$

As stated in the proposition below, in any competitive equilibrium, agents, in each period, consume the entirety of their endowment. Consequently, there is no trade (between the old and the young in time t or between each generation in period t) in the competitive equilibrium—i.e. the equilibrium is *autarky.*

There is no intergeneration trade because the old, although he would like to receive more by trading with the young, are not able to give anything in return to the young in the next period (since the old will not be alive).

There is no intrageneration trade because we assume that agents are homogenous within each generation. Trade within each generation could arise if there were heterogeneity within each generation (e.g. in terms of endowments, see Problem Set 1, Q4); however, there will not be trade across generations as the old would not want to trade with the young.

[5]Since we are concerned with a pure exchange economy, we need not worry about whether firms maximise profits.

However, if, for example, each agent lives for 3 periods (young, middle-aged, old), then there may be trade between the young and the middle-aged (see Problem Set 1, Q5).

Proposition 3.1. *Any competitive equilibrium is autarky; i.e.*

$$\mathbf{x}^i = \mathbf{e}^i, \ \forall i \in \mathbf{I}.$$

Proof. The proof proceeds by considering the optimal consumption of the initial old first, then substituting the result into the period market clearing condition, and continuing the same procedure for the subsequent generations.

Since the initial old only care about consumption in $t = 0$ and has only endowment at time $t = 0$, for any price vector $\mathbf{p}$,

$$x_1^0 = e_1^0.$$

Using the market clearing condition for $t = 1$ ($x_1^0 + x_1^1 = e_1^0 + e_1^1$),

$$x_1^1 = e_1^0 + e_1^1 - x_1^0 = e_1^1.$$

Now suppose that

$$x_{i+1}^i = e_{i+1}^i. \tag{3.1}$$

Mearket clearing condition for $t = i + 1$ is

$$x_{i+1}^i + x_{i+1}^{i+1} = e_{i+1}^i + e_{i+1}^{i+1}.$$

Hence,

$$\begin{aligned} x_{i+1}^{i+1} &= e_{i+1}^i + e_{i+1}^{i+1} - x_{i+1}^i \\ &= e_{i+1}^{i+1}. \end{aligned}$$

That is, we showed that

$$x_{i+1}^i = e_{i+1}^i \Rightarrow x_{i+1}^{i+1} = e_{i+1}^{i+1}.$$

Since we already verified that (3.1) holds for $i = 0$, by induction, it follows that the implication above must hold for all $i \geq 0$. ■

Price ratios as interest rate We now wish to describe the equilibrium price vector for this economy. Suppose we normalise the price of period-1 consumption good to be one; i.e. $p_1 = 1$. The first-order conditions imply that

$$\frac{u_2^i\left(x_i^i, x_{i+1}^i\right)}{u_1^i\left(x_i^i, x_{i+1}^i\right)} = \frac{p_{i+1}}{p_i}.$$

That is, as usual, the marginal rate of substitution between consumption when old and young is equated with the relative price of consumptions goods across the two periods.

By the proposition above, we have that, in equilibrium, $x_i^i = e_i^i$ and $x_i^{i-1} = e_i^{i-1}$. For this allocation to solve generation $i \geq 1$'s maximisation problem, prices must be such that

$$\frac{p_{i+1}}{p_i} = \frac{u_2^i\left(e_i^i, e_{i+1}^i\right)}{u_1^i\left(e_i^i, e_{i+1}^i\right)}, \ \forall i \geq 1. \tag{3.2}$$

We can interpret the relative price as interest rates. To see this, let r_t be the the time-t net interest rate:[6]

$$\frac{1}{1+r_t} := \frac{p_{t+1}}{p_t},\ \forall t. \tag{3.3}$$

Then, we can rewrite equation (3.2) as

$$r_t = \frac{u_1^t\left(e_t^t, e_{t+1}^t\right)}{u_2^t\left(e_i^t, e_{t+1}^t\right)} - 1,\ \forall t \geq 1. \tag{3.4}$$

Notice also that

$$\begin{aligned}
p_t &= \frac{p_{t-1}}{1+r_{t-1}} \\
&= \frac{1}{1+r_{t-1}} \frac{p_{t-2}}{1+r_{t-2}} \\
&= \frac{1}{1+r_{t-1}} \frac{1}{1+r_{t-2}} \cdots \frac{p_1}{1+r_1} \\
&= \prod_{s=1}^{t-1} \frac{1}{1+r_s},
\end{aligned}$$

where we used the fact that $p_1 = 1$.

Competitive equilibrium with log, time-separable utility We analyse whether competitive allocations are Pareto optimal. To simplify the model, suppose we let, for all $i \geq 1$,

$$u^i\left(x_i^i, x_{i+1}^i\right) \equiv v\left(c_y^i, c_o^i\right) := (1-\beta)\ln c_y^i + \beta \ln c_o^i$$

for some $\beta \in (0,1)$. For the endowments, we assume that, for all $i \geq 1$,

$$e_i^i = 1-\alpha,\ e_{i+1}^i = \alpha$$

for some $\alpha \in (0,1)$. That is, we set aggregate endowment at any point in time as 1; i.e. $\bar{e}_t = 1$ for all $t \geq 1$.

Given Proposition 3.1, the competitive equilibrium consumption is given by

$$\begin{aligned}
\bar{c}_y^{i*} &= 1-\alpha, \quad \forall i \geq 1, \\
\bar{c}_o^{i*} &= \alpha, \qquad\ \ \forall i \geq 0,
\end{aligned}$$

We start by describing the equilibrium price vector for this particular economy. Using (3.2) and (3.3), we have that

$$\frac{1}{1+r_i} = \frac{p_{i+1}}{p_i} = \frac{\beta}{1-\beta}\frac{c_y^{i*}}{c_o^{i*}} = \frac{\beta}{1-\beta}\frac{1-\alpha}{\alpha}.$$

[6] Notice that the interest rate is in terms of units of goods:

$$\frac{1}{1+r_t} = \frac{1}{p_t} \Big/ \frac{1}{p_{t+1}}.$$

Using equation (3.4) and replacing i with t, we can write

$$\begin{aligned} r_t &= \frac{1-\beta}{\beta}\frac{\alpha}{1-\alpha} - 1 = \frac{\alpha(1-\beta)-\beta(1-\alpha)}{\beta(1-\alpha)} \\ &= \frac{\alpha-\beta}{\beta(1-\alpha)} =: \bar{r} \end{aligned} \tag{3.5}$$

We can then write

$$\begin{aligned} p_t &= \prod_{i=1}^{t-1} \frac{1}{(1+\bar{r})} \\ &= \left(\frac{1}{1+\bar{r}}\right)^{t-1} = \left(\frac{\beta}{1-\beta}\frac{1-\alpha}{\alpha}\right)^{t-1}, \ \forall t \geq 1. \end{aligned}$$

This gives us the expression for the equilibrium prices.

Optimal savings The quantity $s_t^t \coloneqq e_t^t - x_t^t$ can be interpreted as the saving by generation i while they are young. Note that

$$e_t^t - s_t^t = (1-\alpha) - s_t^t.$$

In this case, the budget constraint is given by

$$\begin{aligned} p_t x_t^t + p_{t+1} x_{t+1}^t &= p_t e_t^t + p_{t+1} e_{t+1}^t \\ \Leftrightarrow \frac{p_{t+1}}{p_t} x_{t+1}^t &= \underbrace{\left(e_t^t - x_t^t\right)}_{\coloneqq s_t} + \frac{p_{t+1}}{p_t} e_{t+1}^t \\ \Leftrightarrow x_{t+1}^t &= \left(\frac{p_{t+1}}{p_t^t}\right)^{-1} \left(e_t^t - x_t^t\right) + e_{t+1}^t \\ &= (1+\bar{r})\, s_t^t + \alpha. \end{aligned}$$

Hence, we can rewrite the agent's utility function in terms of s_t^t:

$$\max_{s_t^t} \ (1-\beta)\ln\left((1-\alpha) - s_t^t\right) + \beta \ln\left[(1+\bar{r})\, s_t^t + \alpha\right].$$

The first-order condition is given by

$$\begin{aligned}
\beta\left(1+\bar{r}\right)\frac{1}{\left(1+\bar{r}\right)s_t^{t*}+\alpha} &= \left(1-\beta\right)\frac{1}{1-\alpha-s_t^{t*}} \\
\Leftrightarrow \beta\left(1+\bar{r}\right)\left(1-\alpha-s_t^{t*}\right) &= \left(1-\beta\right)\left(\left(1+\bar{r}\right)s_t^{t*}+\alpha\right) \\
\Leftrightarrow \beta\left(1+\bar{r}\right)\left(1-\alpha\right)-\alpha\left(1-\beta\right) &= \beta\left(1+\bar{r}\right)s_t^{t*}+\left(1-\beta\right)\left(1+\bar{r}\right)s_t^{t*} \\
\Leftrightarrow s_t^{t*}\left(1+\bar{r}\right) &= \beta\left(1+\bar{r}\right)\left(1-\alpha\right)-\alpha\left(1-\beta\right) \\
\Leftrightarrow s_t^{t*} &= \left(1-\alpha\right)\beta-\frac{\alpha\left(1-\beta\right)}{\left(1+\bar{r}\right)} \qquad (3.6) \\
&= \left(1-\alpha\right)\beta+\left[\left(1-\alpha\right)-\left(1-\alpha\right)\right]-\frac{\alpha\left(1-\beta\right)}{\left(1+\bar{r}\right)} \\
&= \left(1-\alpha\right)-\left(1-\alpha\right)\left(1-\beta\right)-\frac{\alpha\left(1-\beta\right)}{\left(1+\bar{r}\right)} \\
&= \left(1-\alpha\right)-\left(1-\beta\right)\left[\left(1-\alpha\right)+\frac{\alpha}{\left(1+\bar{r}\right)}\right]. \qquad (3.7)
\end{aligned}$$

Hence, optimal saving s_t^{t*} is:

(i) increasing in $\bar{r}$ (see (3.7));

(ii) increasing in β (see (3.7));

(iii) decreasing in α (see (3.6)).

In words, respectively, agents save more when: (i) interest rate is higher (since returns from savings is higher), (ii) they value future consumption more, and (iii) when they have less endowment when they are old.

Notice that the term inside the square brackets, $(1-\alpha)+\alpha/(1+\bar{r})$, is the present value of the total endowment that each agent receives. The expression above means that the optimal saving for the young is his endowment, $1-\alpha$, less a share of the present value of total endowments, where the share depends on his preference for consumption when young, $(1-\beta)$. This is a feature of using the log utility function, which means that the substitution effect and the income effect from changes in the interest rate exactly offset each other.

Equilibrium interest rate Recall that, by Proposition (3.1), in equilibrium, $x_i^i = e_i^i$. This implies that, in equilibrium, savings must be zero; i.e. in equilibrium, prices/real interest rate must be such that the savings of the young are zero.

To see this, substituting (3.5) into (3.7) gives that

$$\begin{aligned}
s_t^{t*} &= \left(1-\alpha\right)-\left(1-\beta\right)\left[\left(1-\alpha\right)+\frac{\alpha}{\frac{1-\beta}{\beta}\frac{\alpha}{1-\alpha}}\right] \\
&= \left(1-\alpha\right)-\left(1-\beta\right)\left[\left(1-\alpha\right)+\frac{\beta\left(1-\alpha\right)}{1-\beta}\right] \\
&= \left(1-\alpha\right)-\left[\left(1-\beta\right)\left(1-\alpha\right)+\beta\left(1-\alpha\right)\right] \\
&= \left(1-\alpha\right)-\left(1-\alpha\right)=0.
\end{aligned}$$

We can use this to give intuition as to why the equilibrium interest rate is increasing in α and decreasing in β:

- ▷ for higher preference parameter β, young agents prefer to consume more when old and less when young—i.e. the agents want to save more at a given interest rate so that the equilibrium interest rate $\bar{r}$ has to be lower;

▷ for higher preference parameter α, the endowment when old is higher and the endowment when young is smaller—i.e. the agents prefer to save less at a given interest rate so that $\bar{r}$ has to be higher.

If $\alpha > \beta$ (i.e. endowment when old is greater than the preference to consume when old), then the equilibrium interest rate is positive, $\bar{r} > 0$, since the agent must be discouraged from saving (see (3.5)). If $\alpha < \beta$, on the other hand, then the equilibrium interest rate is negative, $\bar{r} < 0$.

3.1.3 Best symmetric allocation with log, time separable utility

For convenience, we consider the set of symmetric allocations in which all agents consume the same amount when they are young and when they are old. That is,

$$\begin{aligned} x_i^i &= c_y, \quad \forall i \geq 1, \\ x_{i+1}^i &= c_o, \quad \forall i \geq 0. \end{aligned}$$

In this case, the feasibility/market clearing condition becomes

$$x_i^i + x_i^{i-1} = 1 \Rightarrow c_y + c_o = 1.$$

We now wish to solve the planner's problem. Since generation $i \geq 1$ are homogenous, we can maximise the representative agent's utility subject to the feasibility constraint; i.e.

$$\begin{aligned} \max_{c_y, c_o} \quad & (1-\beta)\ln c_y + \beta \ln c_o \\ s.t. \quad & c_y + c_o = 1. \end{aligned}$$

The Lagrangian is given by

$$\mathcal{L} = (1-\beta)\ln c_y + \beta \ln c_o + \lambda\,(1 - c_y - c_o)$$

The first-order conditions are given by:[7]

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial c_y} &= 0 \Rightarrow \frac{1-\beta}{c_y^*} = 1, \\ \frac{\partial \mathcal{L}}{\partial c_o} &= 0 \Rightarrow \frac{\beta}{c_o^*} = 1, \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 0 \Rightarrow c_y^* + c_o^* = 1. \end{aligned}$$

Dividing the first by the second condition gives

$$\frac{1-\beta}{\beta}\frac{c_o^*}{c_y^*} = 1,$$

Substituting above into the feasibility condition yields that

$$\frac{1-\beta}{\beta} c_o^* + c_o^* = 1;$$

[7] First-order conditions here are sufficient since the objective function is strictly concave.

i.e.

$$c_y^* = 1 - \beta.$$
$$c_o^* = \beta.$$

In the best symmetric allocation, since a higher β implies that agents prefer to consume when they are older, c_0^* is increasing in β while the opposite is true for consumption when the agent is young.

3.1.4 Comparison of competitive equilibrium allocation with the best symmetric allocation

In the unique competitive equilibrium, we have

$$\begin{aligned} \bar{c}_y^{i*} &= 1 - \alpha, & \forall i \geq 1, \\ \bar{c}_o^{i*} &= \alpha, & \forall i \geq 0, \end{aligned}$$

where we use $\bar{\cdot}$ to denote competitive equilibrium allocation. In contrast, the best symmetric allocation is given by

$$\begin{aligned} \hat{c}_y^{i*} &= 1 - \beta, & \forall i \geq 1, \\ \hat{c}_o^{i*} &= \beta, & \forall i \geq 0. \end{aligned}$$

Notice that the competitive equilibrium allocation is a feasible symmetric allocation (since consumption of young and old sum to one). This means that *the best symmetric allocation is strictly preferred by the agents $i \geq 1$* unless $\alpha = \beta$ (in which case generations $i \geq 1$ are indifferent between the two allocations).

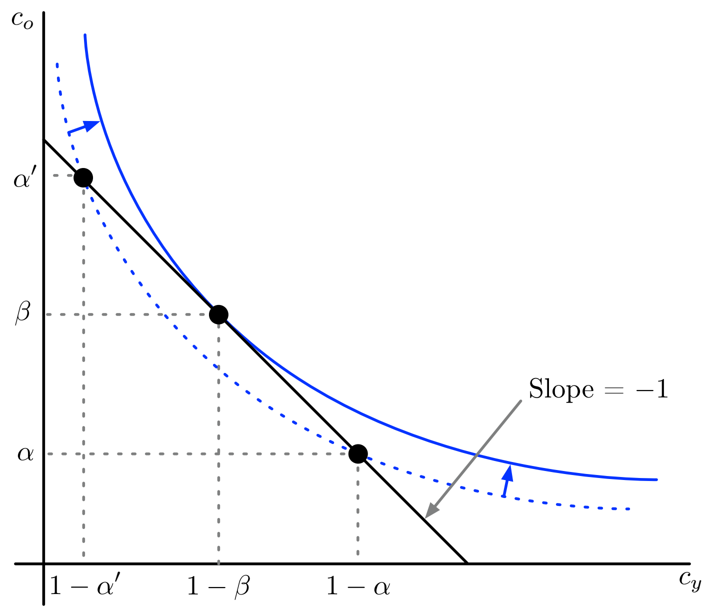
To see if the best symmetric allocation Pareto dominates the best symmetric allocation, it remains to compare the utility of the initial old, generation 0 (since the first-order condition used to derive the best symmetric allocation does not apply to generation 0).

- ▷ Case $\beta > \alpha$ ($\bar{r} < 0$): $\bar{c}_o^{0*} = \alpha < \beta = \hat{c}_o^{0*}$. In this case, the initial old strictly prefers the best symmetric allocation. Hence, all agents are strictly better off in the best symmetric allocation over the competitive equilibrium. That is, the best symmetric feasible allocation Pareto dominates the competitive equilibrium allocation. Notice that, in this case, the interest rate is negative (see equation (3.5)).

- ▷ Case $\beta = \alpha$ ($\bar{r} = 0$): $\bar{c}_o^{0*} = \alpha = \beta = \hat{c}_o^{0*}$. In this case, the two allocations exactly coincide and the interest rate is exactly zero.

- ▷ Case $\beta < \alpha$ ($\bar{r} > 0$): $\bar{c}_o^{0*} = \alpha > \beta = \hat{c}_o^{0*}$. In this case, the initial old strictly prefers the competitive equilibrium allocation to the best symmetric allocation. Thus, neither allocations Pareto dominate each other. In this case, the interest rate is strictly positive.

Consider the figure below:



The best symmetric allocation is given by where the indifference curve for generation $i \geq 1$ is tangential to the budget constraint.

We first focus on the point $(1-\alpha, \alpha)$, where $\beta > \alpha$. This point lies on an indifference curve lower than of $(1-\beta, \beta)$ so that $(1-\beta, \beta)$ is preferred by generation $i \geq 1$ over $(1-\alpha, \alpha)$. For the initial old, whose consumption equals α in the competitive equilibrium, the best symmetric allocation represents an improvement over the unique competitive equilibrium since $\beta > \alpha$. Hence, in this case, a shift to the best symmetric allocation represents a Pareto improvement.

Consider now the point $(1-\alpha', \alpha')$, where $\beta < \alpha'$. Again, this point lies on an indifference curve below that of $(1-\beta, \beta)$ so that the latter is still preferred by generation $i \geq 1$ over $(1-\alpha', \alpha')$. However, for the initial old whose consumption equals α' under the competitive equilibrium, the best symmetric allocation represents a worse allocation since $\beta < \alpha'$. Hence, in this case, the initial old would not want to move to the best symmetric allocation, despite the fact that all other generations would be better off in the best symmetric allocation.

3.1.5 Social security

Consider a tax policy indexed by a single parameter τ whereby the young are taxed τ and this amount is given to the old. For positive τ, this tax policy resembles a pay-as-you-go social security system. By construction, this policy is feasible since tax equals subsidies in each period.

With this policy, the (post-tax) endowments are now:

$$\begin{aligned} e_i^i &= (1-\alpha) - \tau, \quad && \forall i \geq 1, \\ e_{i+1}^i &= \alpha + \tau, && \forall i \geq 0. \end{aligned}$$

Since the consumers consume all their endowments in each period, consumption equal to the endowments above represents the competitive equilibrium with the social security system.

Exercise 3.1. Suppose you propose a pay-as-you-go social security system for the OLG economy as described above.

▷ Under what parameter values will this policy produce a competitive equilibrium that Pareto dominates the competitive equilibrium without the policy?

▷ What interest rate does the competitive equilibrium without the policy need to be so that the introduction of social security produces a competitive equilibrium that Pareto dominates the one without policy?

Solution. Recall that when $\beta > \alpha$, the best symmetric allocation Pareto dominates the competitive equilibrium allocation. Thus, in such a situation, we can set $\tau = \beta - \alpha > 0$ to achieve the best symmetric allocation under competitive equilibrium. Note that interest rate is negative when $\beta > \alpha$; i.e. the policy intervention can lead to a Pareto improvement when interest rates are negative.

Suppose we introduce a social security system when interest rate is positive. Who gains/loses? Clearly, since $\tau > 0$ under a social security system, the initial old is better off with than without the system since he receives something (τ) for nothing. Note that it is still the case that the equilibrium is autarky; i.e. equilibrium interest rate is such that optimal savings is zero.

The quantity $s_t^t \coloneqq e_t^t - \tau - x_t^t$ can be interpreted as the saving by generation i while they are young. Note that

$$e_t^t - \tau - s_t^t = (1-\alpha) - \tau - s_t^t.$$

In this case, the budget constraint is given by

$$\begin{aligned} p_t x_t^t + p_{t+1} x_{t+1}^t &= p_t \left(e_t^t - \tau\right) + p_{t+1}\left(e_{t+1}^t + \tau\right) \\ \Rightarrow \frac{p_{t+1}}{p_t} x_{t+1}^t &= \underbrace{\left(e_t^t - \tau - x_t^t\right)}_{\coloneqq s_t} + \frac{p_{t+1}}{p_t}\left(e_{t+1}^t + \tau\right) \\ \Rightarrow x_{t+1}^t &= \left(\frac{p_{t+1}}{p_t^t}\right)^{-1}\left(e_t^t - \tau - x_t^t\right) + \left(e_{t+1}^t + \tau\right) \\ &= \left(1 + \bar{r}\right) s_t^t + \alpha + \tau. \end{aligned}$$

Hence, we can rewrite the agent's utility function in terms of s_t^t:

$$\max_{s_t^t} \ (1-\beta)\ln\left((1-\alpha) - \tau - s_t^t\right) + \beta \ln\left[(1+\bar{r})\, s_t^t + \alpha + \tau\right].$$

First-order condition is then

$$\begin{aligned}
\frac{\beta\left(1+\bar{r}\right)}{\left(1+\bar{r}\right)s_t^{t*}+\alpha+\tau} &= \frac{1-\beta}{\left(1-\alpha\right)-\tau-s_t^{t*}} \\
\Rightarrow \beta\left(1+\bar{r}\right)\left(\left(1-\alpha\right)-\tau-s_t^{t*}\right) &= \left(1-\beta\right)\left(\left(1+\bar{r}\right)s_t^{t*}+\alpha+\tau\right) \\
\Leftrightarrow s_t^{t*}\left(1+\bar{r}\right)\left(\beta+\left(1-\beta\right)\right) &= \beta\left(1+\bar{r}\right)\left(1-\left(\alpha+\tau\right)\right)-\left(1-\beta\right)\left(\alpha+\tau\right) \\
&= s_t^{t*}\left(1+\bar{r}\right) \\
\Rightarrow s_t^{t*} &= \beta\left(1-\left(\alpha+\tau\right)\right)-\frac{\left(1-\beta\right)\left(\alpha+\tau\right)}{1+\bar{r}} \\
&= \beta\left(1-\left(\alpha+\tau\right)\right)+\left[\left(1-\left(\alpha+\tau\right)\right)-\left(1-\left(\alpha+\tau\right)\right)\right] \\
&\quad -\frac{\left(1-\beta\right)\left(\alpha+\tau\right)}{1+\bar{r}} \\
&= \left(1-\alpha-\tau\right)-\left(1-\beta\right)\left(1-\alpha-\tau\right) \\
&\quad -\frac{\left(1-\beta\right)\left(\alpha+\tau\right)}{1+\bar{r}} \\
&= \left(1-\alpha-\tau\right)-\left(1-\beta\right)\left[\left(1-\alpha-\tau\right)+\frac{\alpha+\tau}{1+\bar{r}}\right].
\end{aligned}$$

This is the same as (3.7) when $\tau=0$. The equilibrium interest rate that ensures $s_t^{t*}=0$ is given by

$$\begin{aligned}
\frac{\beta}{1-\beta}\frac{1-\alpha-\tau}{\alpha+\tau} &= \frac{1}{1+\bar{r}} \\
\Rightarrow \bar{r} &= \frac{1-\beta}{\beta}\frac{\alpha+\tau}{1-\alpha-\tau}-1 \\
&= \frac{\left(1-\beta\right)\left(\alpha+\tau\right)-\beta\left(1-\alpha-\tau\right)}{\beta\left(1-\alpha-\tau\right)} \\
&= \frac{\alpha-\beta+\tau}{\beta\left(1-\alpha-\tau\right)},
\end{aligned}$$

which is the same as (3.4) if $\tau=0$.

3.1.6 Why did the First Welfare Theorem “fail”?

We showed that, when $\beta>\alpha$, the competitive equilibrium is not Pareto optimal. However, this appears to be a counter-example to the First Welfare Theorem, which says that competitive equilibria are Pareto optimal. Since we assumed log-utility, which is strictly increasing, u^i clearly satisfies local non-satiation so that the First Welfare Theorem *should* hold. So why did it fail?

The failure is due to the fact that the OLG model has *infinitely* many agents and *infinitely* many goods. Under these circumstances, a key condition required for the First Welfare Theorem—that the value of the aggregate consumption is finite—fails; i.e. the following condition fails:

$$\mathbf{p}\cdot\sum_{i\in\mathbf{I}}\bar{\mathbf{x}}^i<\infty. \tag{3.8}$$

To see why this condition is important, recall that the proof of the First Welfare Theorem relies on the argument that any feasible allocation that Pareto dominates the competitive equilibrium costs more. Specifically, we first assume, by way of contradiction, that there is a feasible allocation $\left\{\mathbf{x}^i\right\}$ that Pareto dominates the competitive equilibrium allocation

$\{\bar{\mathbf{x}}^i\}$. We then show that this implies:

$$\begin{aligned} \mathbf{p}\cdot\mathbf{x}^i &\geq \mathbf{p}\cdot\bar{\mathbf{x}}^i, && \forall i\in\mathbf{I}, \\ \mathbf{p}\cdot\mathbf{x}^{i'} &> \mathbf{p}\cdot\bar{\mathbf{x}}^{i'}, && \text{for some } i'\in\mathbf{I}. \end{aligned}$$

We sum across the agents to obtain that

$$\mathbf{p}\cdot\sum_{i\in\mathbf{I}}\mathbf{x}^i=\sum_{i\in\mathbf{I}}\mathbf{p}\cdot\mathbf{x}^i>\sum_{i\in\mathbf{I}}\mathbf{p}\cdot\bar{\mathbf{x}}^i=\mathbf{p}\cdot\sum_{i\in\mathbf{I}}\bar{\mathbf{x}}^i.$$

Finally, we found a contradiction by comparing this against the feasibility constraint. However, notice that the above comparison is only valid if (3.8) holds since, otherwise, we will be comparing infinity to infinity; we would be trying to say "$\infty>\infty$".

In the OLG model that we set up, recall that

$$\sum_{i=1}^{\infty}\bar{\mathbf{x}}^i=(1,1,\dots)'$$

and

$$\mathbf{p}=\left(1,\frac{1}{1+\bar{r}},\left(\frac{1}{1+\bar{r}}\right)^2,\dots\right)',$$

where $\bar{r}={}^{\alpha-\beta}/_{\beta(1-\alpha)}$. Thus,

$$\begin{aligned} \mathbf{p}\cdot\sum_{i\in I}\bar{\mathbf{x}}^i &= \lim_{T\to\infty}\left(1+\frac{1}{1+\bar{r}}+\left(\frac{1}{1+\bar{r}}\right)^2+\cdots\right) \\ &= \lim_{T\to\infty}\left(\frac{1-\left(\frac{1}{1+\bar{r}}\right)^{T+1}}{1-\frac{1}{1+\bar{r}}}\right)=\lim_{T\to\infty}\left(\frac{1+\bar{r}-\left(\frac{1}{1+\bar{r}}\right)^{T}}{\bar{r}}\right). \end{aligned}$$

The sum is divergent if

$$\left|\frac{1}{1+\bar{r}}\right|>1;$$

i.e. whenever the interest rate is negative; i.e. when $\beta>\alpha$. The sum converges to a finite limit, however, when $\beta\leq\alpha$; i.e. when the interest rate is nonnegative.

The following propositions establish that the First Welfare Theorem holds when either m (the number of goods) and I (the number of agents) is finite (to simplify, we assume no production).

Proposition 3.2. *Assume that there are finitely many goods ($m < \infty$) but that the number of agents is infinite ($I = \infty$). Assume that the aggregate endowment is bounded above in each component:*

$$\bar{e}_\ell := \sum_{i \in I} e_\ell^i < \infty, \ \forall \ell \in \{1, 2, \dots, m\}.$$

Then, under these assumptions, in any competitive equilibrium $\mathbf{p}$, $\{\bar{\mathbf{x}}^i\}$, *we must have*

$$\mathbf{p} \cdot \sum_{i \in I} \bar{\mathbf{x}}^i < \infty.$$

Proposition 3.3. *Assume that there are finitely many agents ($I < \infty$) but that there are infinitely many goods ($m = \infty$). Assume that all u^i satisfy local non-satiation. Then, under these assumptions, in any competitive equilibrium* $\mathbf{p}$, $\{\bar{\mathbf{x}}^i\}$, *we must have*

$$\mathbf{p} \cdot \sum_{i \in I} \bar{\mathbf{x}}^i < \infty.$$

Proof. (Proposition 3.2) In any competitive equilibrium feasibility condition must hold:

$$\sum_{i \in \mathbf{I}} \mathbf{e}^i = \sum_{i \in \mathbf{I}} \bar{\mathbf{x}}^i \Leftrightarrow \sum_{i \in \mathbf{I}} e_\ell^i = \sum_{i \in \mathbf{I}} \bar{x}_\ell^i < \infty, \ \forall \ell \in \{1, 2, \dots, m\}.$$

It follows then that (since prices are finite):

$$\mathbf{p} \cdot \sum_{i \in \mathbf{I}} \mathbf{e}^i = \mathbf{p} \cdot \sum_{i \in \mathbf{I}} \bar{\mathbf{x}}^i < \infty. \qquad \blacksquare$$

Proof. (Proposition 3.3) Since the consumers maximise their utility and that u^i satisfies local non-satiation, then budget constraint must bind at the optimal. That is, $\mathbf{p} \cdot \bar{\mathbf{x}}^i = \mathbf{p} \cdot \mathbf{e}^i$ for all $i \in I$. Since finite sum of finite numbers are finite, it follows that

$$\mathbf{p} \cdot \sum_{i \in I} \bar{\mathbf{x}}^i = \mathbf{p} \cdot \sum_{i \in I} \mathbf{e}^i < \infty. \qquad \blacksquare$$

3.2 Growing OLG economy

We now introduce population and productivity growth into the previous setup.

Let N_t denote the number of young at time t and n denote the (constant) growth rate of population; i.e.

$$N_{t+1} = (1 + n) N_t, \ \forall t \geq 1.$$

We assume that $N_0 = 1$. Let g denote the growth rate of "productivity" of endowments for each generation so that

$$e_{t+1}^{t+1} = (1 + g) e_t^t,$$
$$e_{t+2}^{t+1} = (1 + g) e_{t+1}^t.$$

This means that the endowments of young (old) of generation $t + 1$ are greater than the endowment of young (old) of generation t by g%. Let $e_1^0 = \alpha$ and $e_0^0 = 1 - \alpha$ (note that the latter generation does not exist in the model, we're simply saying that the young in period 1 has higher productivity than than that of the old in period $t = 1$):

$$\begin{aligned} e_t^t &= (1+g)\, e_{t-1}^{t-1} = (1+g)^2\, e_{t-2}^{t-2} = \cdots = (1+g)^t\, e_0^0 = (1+g)^t\, (1-\alpha)\,, \\ e_{t+1}^t &= (1+g)\, e_t^{t-1} = (1+g)^2\, e_{t-1}^{t-2} = \cdots = (1+g)^t\, e_1^0 = (1+g)^t\, \alpha \end{aligned}$$

for all $t \geq 1$. In any period t, the aggregate endowment is given by

$$\begin{aligned} N_{t-1} e_t^{t-1} + N_t e_t^t &= N_{t-1} N_t\, (1+g)^{t-1}\, \alpha + N_t\, (1+g)^t\, (1-\alpha) \\ e_t^{t-1} + (1+n)\, e_t^t &= (1+g)^{t-1}\, \alpha + (1+n)\,(1+g)^t\,(1-\alpha) \\ &= (1+g)^{t-1}\left[(1+g)\,(1+n)\,(1-\alpha) + \alpha\right]. \end{aligned}$$

3.2.1 Optimal saving

We can solve for the optimal savings $s_t^t := e_t^t - x_t^t$. First, using the new definition of e_t^t and e_{t+1}^t, we can write consumption in terms of savings:

$$\begin{aligned} x_t^t &= (1+g)^t\,(1-\alpha) - s_t^t, \\ x_{t+1}^t &= (1+\bar{r})\, s_t^t + (1+g)^t\, \alpha. \end{aligned}$$

Using these, we can rewrite the agent's utility function in terms of s_t^t:

$$\max_{s_t^t}\ (1-\beta)\ln\left((1+g)^t\,(1-\alpha) - s_t^t\right) + \beta \ln\left[(1+\bar{r})\, s_t^t + (1+g)^t\,\alpha\right].$$

The first-order condition is given by

$$\begin{aligned} \beta\,(1+\bar{r})\,\frac{1}{(1+\bar{r})\, s_t^{t*} + (1+g)^t\,\alpha} &= (1-\beta)\,\frac{1}{(1+g)^t\,(1-\alpha) - s_t^t} \qquad (3.9)\\ \Leftrightarrow s_t^{t*} &= (1-g)^t\left((1-\alpha) - (1-\beta)\left[(1-\alpha) + \frac{\alpha}{(1+\bar{r})}\right]\right), \end{aligned}$$

which is simply the first-order condition without growth (see (3.7)) multiplied by $(1-g)^t$. That is, the tradeoff involved in deciding the optimal savings is exactly the same as before—agents still decide on the savings by deciding on the proportion of the present value of the total endowments.

3.2.2 Competitive equilibrium

The exact same arguments used in the proof of Proposition (3.1) hold so that, in any competitive equilibrium of this growing economy, $\mathbf{x}^i = \mathbf{e}^i$ for all $i \in \mathbf{I}$. Hence, for each agent of generation $t \geq 1$,

$$\begin{aligned} c_t^{t*} &= (1+g)^t\,(1-\alpha)\,, \quad \forall t \geq 1, \\ c_{t+1}^{t*} &= (1+g)^t\,\alpha. \qquad\quad \forall t \geq 0. \end{aligned}$$

Since in a competitive equilibrium, there still cannot be any saving, $\bar{r}$ should be such that aggregate savings of generation t is zero. That is, $N_t s_t^{t*} = 0$. Using (3.9) and the fact that $N_t = (1+n)^t$,

$$N_t s_t^{t*} = (1+n)^t\,(1-g)^t\left((1-\alpha) - (1-\beta)\left[(1-\alpha) + \frac{\alpha}{(1+\bar{r})}\right]\right) = 0,$$

which holds if

$$\bar{r} = \frac{\alpha - \beta}{\beta (1 - \alpha)}. \tag{3.10}$$

This is the same equilibrium interest rate as before.

We could have also obtained the equilibrium interest rate using the first-order conditions for the generation $t \geq 1$:

$$\frac{1}{1 + r_t} = \frac{p_{t+1}}{p_t} = \frac{\beta}{1 - \beta} \frac{c_t^{t*}}{c_{t+1}^{t*}} = \frac{\beta}{1 - \beta} \frac{(1 + g)^t (1 - \alpha)}{(1 + g)^t \alpha},$$

which is also solved by (3.10).

3.2.3 Best symmetric allocation

Since each generation has different endowments (due to the growth in productivity), the best symmetric allocation cannot be defined in the same way as before. However, we can define feasible symmetric allocations as those that satisfy

$$\begin{aligned} N_t c_y^t + N_{t-1} c_o^t &= N_t e_y^t + N_{t-1} e_o^t \\ &= N_t (1 - \alpha) (1 + g)^t + N_{t-1} \alpha (1 + g)^{t-1}, \end{aligned}$$

where each agent born at time t and young at t, and each agent born at time $t - 1$ and old at t consumes, respectively,

$$\begin{aligned} c_y^t &= \hat{c}_y (1 + g)^t, \\ c_o^t &= \hat{c}_o (1 + g)^{t-1}. \end{aligned}$$

Then, since $N_t = (1 + n)^t$, we can rewrite the feasible symmetric allocation as

$$\begin{aligned} \frac{N_t}{N_{t-1}} \hat{c}_y (1 + g)^t + \hat{c}_o (1 + g)^{t-1} &= \frac{N_t}{N_{t-1}} (1 - \alpha) (1 + g)^t + \alpha (1 + g)^{t-1}, \\ \Leftrightarrow (1 + n) (1 + g) \hat{c}_y + \hat{c}_o &= (1 - \alpha) (1 + n) (1 + g) + \alpha. \end{aligned} \tag{3.11}$$

We can now solve for the best feasible symmetric allocation as $\hat{c}_y$ and $\hat{c}_o$ that solves

$$\max_{\hat{c}_y, \hat{c}_o} \quad (1 - \beta) \ln \left[\hat{c}_y (1 + g)^t \right] + \beta \ln \left[\hat{c}_o (1 + g)^t \right]$$

subject to (3.11). The Lagrangian is given by

$$\begin{aligned} \mathcal{L} = (1 - \beta) \ln \left[\hat{c}_y (1 + g)^t \right] + \beta \ln \left[\hat{c}_o (1 + g)^t \right] \\ + \lambda \left[\alpha + (1 - \alpha) (1 + n) (1 + g) - ((1 + n) (1 + g) \hat{c}_y + \hat{c}_o) \right] \end{aligned}$$

so that the first-order conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \hat{c}_y} = 0 &\Rightarrow \frac{1 - \beta}{\hat{c}_y^*} (1 + g)^t = \lambda (1 + n) (1 + g), \\ \frac{\partial \mathcal{L}}{\partial \hat{c}_o} = 0 &\Rightarrow \frac{\beta}{\hat{c}_o^*} (1 + g)^t = \lambda \end{aligned}$$

and the constraint itself. Substituting λ into the first expression:

$$\begin{aligned} \frac{1-\beta}{\hat{c}_y^*} &= \frac{\beta}{\hat{c}_o^*}\left(1+n\right)\left(1+g\right) \\ \Leftrightarrow \hat{c}_o^* &= \frac{\beta}{1-\beta}\left(1+n\right)\left(1+g\right)\hat{c}_y^*. \end{aligned}$$

Substituting this into the budget constraint,

$$\begin{aligned} \left(1+n\right)\left(1+g\right)\hat{c}_y^* + \frac{\beta}{1-\beta}\left(1+n\right)\left(1+g\right)\hat{c}_y^* &= \left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha \\ \Leftrightarrow \hat{c}_y^*\left(1+n\right)\left(1+g\right)\left(1+\frac{\beta}{1-\beta}\right) &= \left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha \\ \hat{c}_y^* &= \left(1-\beta\right)\frac{\left[\left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha\right]}{\left(1+n\right)\left(1+g\right)}. \end{aligned}$$

This, in turn, implies that

$$\hat{c}_o^* = \beta\left[\left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha\right].$$

3.2.4 Comparison of competitive equilibrium allocation with the best symmetric allocation

In the competitive equilibrium, consumptions are given by

$$\begin{aligned} c_t^{t*} &= \left(1+g\right)^t\left(1-\alpha\right), \quad && \forall t \geq 1, \\ c_{t+1}^{t*} &= \left(1+g\right)^t\alpha. && \forall t \geq 0. \end{aligned}$$

Under the best symmetric allocation, we have

$$\begin{aligned} \hat{c}_y^{t*} &= \left(1+g\right)^t\left(1-\beta\right)\frac{\left[\left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha\right]}{\left(1+n\right)\left(1+g\right)} \quad && \forall t \geq 1, \\ \hat{c}_o^{t*} &= \left(1+g\right)^t\beta\left[\left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha\right] && \forall t \geq 0. \end{aligned}$$

The competitive equilibrium allocation is a feasible symmetric allocation. Hence, by the same argument as before (i.e. revealed preference), for all $t \geq 1$, the best symmetric allocation must be preferred over the competitive equilibrium allocation. Thus, for the best symmetric allocation to Pareto dominate the competitive equilibrium allocation, it remains to show that the initial old prefers the best symmetric allocation over the competitive equilibrium allocation:

$$\begin{aligned} \hat{c}_o^{0*} &\geq c_0^{0*} \\ \Leftrightarrow \beta\left[\left(1-\alpha\right)\left(1+n\right)\left(1+g\right)+\alpha\right] &\geq \alpha \\ \Leftrightarrow \left(1-\alpha\right)\left(1+n\right)\left(1+g\right) &\geq \frac{\alpha}{\beta}-\alpha = \alpha\left(\frac{1}{\beta}-1\right) \\ \Leftrightarrow \left(1+n\right)\left(1+g\right) &\geq \frac{\alpha}{1-\alpha}\frac{1-\beta}{\beta} \\ &= 1+\bar{r} \end{aligned}$$

since

$$
\begin{aligned}
\bar{r} &= \frac{\alpha - \beta}{\beta (1 - \alpha)} \\
\Leftrightarrow 1 + \bar{r} &= \frac{\beta (1 - \alpha) + \alpha - \beta}{\beta (1 - \alpha)} = \frac{\alpha (1 - \beta)}{\beta (1 - \alpha)}.
\end{aligned}
$$

Therefore, the best symmetric allocation Pareto dominates the competitive equilibrium when

$$
(1 + n)(1 + g) \geq (1 + \bar{r}).
$$

Recall that, without population and productivity growth, the best symmetric allocation Pareto dominated the competitive equilibrium (i.e. the First Welfare Theorem failed) whenever the interest rate was negative. With population and productivity growth, we now find that the First Welfare Theorem fails even if there interest rate is positive. For the Theorem to hold, it must be that the interest must be higher than the "sum" of population and productivity growth. Since $\ln(1 + x) \approx x$ for small x, we can write the inequality as

$$
n + g \gtrsim \bar{r}.
$$

3.2.5 Social security

Suppose that we implement a pay-as-you-go social security system where τ_t, which denotes the tax to an agent born at date t when young, is given by

$$
\tau_t = (1 + g)^t \hat{\tau}, \ \forall t \geq 1,
$$

and that all the revenues from the tax levied at time t are used to pay subsidies to agents born at $t - 1$ that are old at time t. At time t, the government collects from each young born in t an amount τ_t as taxes:

$$
N_t \tau_t = (1 + n)^t (1 + g)^t \hat{\tau}
$$

and distributes this to the old in $t - 1$. There are N_{t-1} number of old in time t such that each old receives in time t, an amount

$$
\frac{N_t \tau_t}{N_{t-1}} = \frac{(1 + n)^t (1 + g)^t \hat{\tau}}{(1 + n)^{t-1}} = (1 + n)(1 + g)^t \hat{\tau}.
$$

Normalising by the old's productivity $(1 + g)^{t-1}$, each old receives $(1 + n)(1 + g)\hat{\tau}$ in time t.

The post-tax endowments are now given by

$$
\begin{aligned}
e_t^t &= (1 + g)^t [(1 - \alpha) - \hat{\tau}], \\
e_{t+1}^t &= (1 + g)^t (\alpha + (1 + n)(1 + g)\hat{\tau}).
\end{aligned}
$$

This is also a competitive equilibrium allocation (since in any competitive equilibrium consumption equals endowments).

To ensure that this is equal to the best symmetric allocation, we must set

$$\begin{aligned}(1-\alpha)-\hat{\tau} &= (1-\beta)\,\frac{[(1-\alpha)(1+n)(1+g)+\alpha]}{(1+n)(1+g)} \\ \Rightarrow \hat{\tau} &= (1-\alpha)-(1-\beta)\,\frac{[(1-\alpha)(1+n)(1+g)+\alpha]}{(1+n)(1+g)} \\ &= (1-\alpha)-(1-\beta)\left[(1-\alpha)+\frac{\alpha}{(1+n)(1+g)}\right]\end{aligned}$$

Thus, we again find that social security can improve welfare when interest rates are sufficiently low.

3.3 Extension 1: Privatization of Social Security with Government (PS1 Q1)

In this problem, we stay with the endowment economy and study the effects of privatizing social security.

3.3.1 Problem 1 (Walras' Law)

Key Idea *Walras' Law* is a trick that we will use extensively. When we look at an equilibrium, there are multiple conditions for market clearing and budget constraints. These are linear equations, so we can deduce that they are not linearly independent.

This is trivial for now, but it gets more complicated later on. Once we bring capital, a version of this equation will help us get a dynamical system whose solution will be an equilibrium. In this case, the goal is to show that the market clearing condition $B_t + s_t = 0$ holds, given the budget constraints for agents and government and goods market clearing.

3.3.2 Problem 2 (Ricardian Equivalence)

Key Idea Consider two schemes:

- ▷ Pay θ when young; receive θ when old
- ▷ Pay θ' when young; receive 0 when old

When $\theta' = \theta - \theta/(1+r_t)$, the equilibrium will not change.

To show this, we *guess* that the interest rate remains the same. Then everyone has the same budget constraint. The previous allocation is feasible and it's optimal. The market clearing conditions hold by Walras' Law. We have thus constructed a new equilibrium.

Ricardian Equivalence We essentially implemented a fiscal policy that did not change anything. Private savings went up and public savings declined.

Role of Government Debt You need to issue debt for the first period in order to keep the first old generation as well off as before. The debt has interest payments, which are entirely financed by the new tax revenue θ' paid by the young generation. Specifically, the amount of debt the government needs to keep the first old generation as well off is

$$\theta - \left(\theta - \frac{\theta}{1+\bar{r}}\right) = \frac{\theta}{1+\bar{r}}$$

with interest rate payments of

$$\frac{\bar{r}\theta}{1+\bar{r}}$$

which is exactly equal to what each generation pays to the government.

3.4 Extension 2: Production Economy without Government (PS1 Q2, Q3)

Now we introduce production into the economy. For each equilibrium condition, we find an easier characterization:

- ▷ Agents maximize utility (can be simplified by expressing savings as a function of relative prices, current income, and total wealth)
- ▷ Firms maximize profits (can be simplified through first-order conditions + no arbitrage)
- ▷ Market clearing in goods
- ▷ Government budget constraint holds: $B_t + g_t = B_{t-1}\left(1 + c_{t-1}\right) + \tau_t^t + \tau_t^{t-1}$
- ▷ Labor supply is inelastic with $L_t = 1$.
- ▷ Law of motion for kapital: $I_t = K_{t+1} - (1-\delta)K_t$

Key Idea 1 Suppose you could buy a house and put it in airbnb. You return money next period $(-r_t)$; house depreciated $(-\delta)$ and rent it for one period to make (v_{t+1}). They have to be equal, which gives us

$$v_{t+1} = r_t + \delta$$

This is a necessary condition for an equilibrium. Why do we establish this? First, this only enters the agent's budget constraint and it does not enter the utility function. By establishing that this holds in the equilibrium, we can now focus on the agent's problem.

Key Idea 2 Savings are income minus consumption. Savings are a function of (1) relative prices (r_t), (2) current income, and (3) present value of income (which is the RHS of the agent's budget constraint, thereby allowing us to pin down consumption).

Key Idea 3 We establish an implicit difference equation for K_{t+1} which simplifies the analysis out. To check that this holds in equilibrium, you have to construct prices and see if markets clear and every entity is maximizing. Wage will be F_L. The rental rate of capital will be $F_K\left(K_t, 1\right)$. The interest rate would be $r_t = F_K(K_{t+1}, 1) - \delta$. Markets clear since Walras' Law holds.

Suppose you now solve for K_{t+1} as a function of K_t. Let's call that $K_{t+1} = s^*\left(K_t, \tau\right)$.

Key Idea 4 Suppose you start with social security and you get rid of it. The old would consume; the young would partly consume and partly save. If consumption is a normal good, they will consume a little bit more now and a little bit more in the future by saving more. More savings imply more capital in the next period. More capital will make the wages higher for the new generation that comes afterward. Thus the new generation is happier and happier − except for the initial old. This is the way privatization of social security is sold. You will save more and wages will increases. But the initial old will not be happy with this.

Simple Example Suppose $\delta = 1$ with $F(K, L) = AK^\gamma L^{1-\gamma}$ with log utility $(1-\beta)\log c_t^t + \beta \log c_{t+1}^t$. Log utility implies $c_t = (1-\beta)w_t$. The saving function for each generation is

$$s_t = (1-\gamma)AK^\gamma L^{-\gamma} = (1-\gamma)AK^\gamma \beta$$

which implies

$$K_{t+1} = \beta(1-\gamma)K_t^\gamma$$

In this case, the capital level converges to a steady-state. This is very well-behaved example.

What is the interest rate at K^*? If β increases, the steady state will be at a higher K^*. Since $\delta = 1$ and if v_{t+1} is not greater than 0, the interest rate has to be negative. When interest rates are negative, the allocation is not efficient. No matter where you start, you will be going to a steady state. If the steady state interest rate is negative, then the infinite sum will diverge. This has nothing to do with capital. The intuition here is that you shouldn't save too much. If we're saving so much that the net return is negative, there's an issue.

3.5 Extension 3: Heterogeneity

This extension has people with identical (log) preferences but hetereogeneous in terms of how much money they have when young vs. old. Some people when young could be borrowing and some could be saving, but total should be zero. Once again, if interest rates are negative, we will have the same dynamic inefficiency.

Because we are given homothetic and identical preferences, we can solve the problem given the averages of the α_ns, derive a solution, and check that this satifies the equilibrium. This is a preview of aggregation, which we cover many lectures later.

3.6 Extension 4: Multiple-Periods

Now suppose agents live multiple periods in the basic model. What are the challenges in solving this model? It's that at $t = 1$, there are multiple old people and the original autarky argument needs to be thrown away. This is what makes finding the equilibrium tougher than before.

Key Idea 1 Our goal is to establish that examining the sign of the interest rate is enough to determine if the equilibrium can be Pareto improved. We show this by adhering to the following steps:

(i) Derive the FOCs for a competitive equilibrium $\{\bar{c}^t\}$ with $\tau = 0$.

(ii) Construct a feasible allocation $\{c^{*t}(\tau)\}$.

(iii) We show that moving towards the feasible allocation makes the allocation more Pareto-efficient.

Why does this work? In step 3, we take the derivative of the utility U with respect to τ. Recall that the scheme suggested in step 2 implies that the person starts getting paid at year $Y + 1$. Since (1) marginal utilities are proportionally to prices and (2) $\tau = 0$ implies $c^{*t} = \bar{c}^t$, the directionality of the derivative depends on the interest rate. And once again, the punchline remains the same: *if interest rate ares strictly negative, a competitive equilibrium can be Pareto improved upon by the introduction of social security.* Note that this does not immediately give us the Pareto-optimal solution. Computing such allocation requires more work and more information.

4 Perpetual Youth Model

The model we study here is called the *Model of Perpetual Youth* as individuals have a fixed probability of death, independent of their age. This (possibly) unrealistic assumption is what gives the model its tractability. As in the case of OLG, when an individual dies, they leave no bequests, and so each new individual starts life with no assets.

This immediately raises a problem—what do individuals do about their uncertain death? The solution lies in the fact that, although the timing of death of an individual is uncertain, we can say with certainty what the average number of deaths in a generation will be (assuming there are many people in each generation). The model contains insurance type institutions who pay annuity payments to each individual in return for receiving all their assets on death. This neatly solves the problem of what to do with a persons assets once they die.

4.1 The setup

4.1.1 Probability of death

The constant probability of death assumption is equivalent to saying that the "time until death" random variable follows an exponential distribution. Let X be this variable, then its density function is given by

$$f_X(t) = p\exp(-pt), \ \forall t \geq 0,$$

where p can be interpreted as the *rate* of death. As such p need not be bounded between zero and one. What is bounded between zero and one is the probability than an individual dies within any given discrete time.

The expected lifetime of an agent, i.e. the expected value of X, is given by

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty tp\exp(-pt)\,dt \\ \text{(integration by parts)} &= [-t\exp(-pt)]_0^\infty + \int_0^\infty \exp(-pt)\,dt \\ &= (0-0) + \left[-\frac{1}{p}\exp(-pt)\right]_0^\infty = \frac{1}{p}.\end{aligned}$$

Thus, we can think of $1/p$ as an index of the effective horizon of individuals in the model. As the probability of death (p) decreases, the horizon index increases. In the limit, as $p \to 0$, the horizon becomes infinite and we are in the infinitely lived agents (Ramsey) model. Thus, this setup generalises the Ramsey model.

At every instance of time, a new generation composed of people with constant probability of death p is born. Each generation is large enough so that p is also the rate at which the cohort size decreases through time. Thus, although each person is uncertain about the time of his death, the size of a generation declines deterministically through time. We normalise the size of each new generation as p so that a cohort born at time s has a size, as of time t, of one (see section 4.7.1 for derivation):

$$N(s,t) := p\exp(-p(t-s)), \ \forall t \geq s.$$

The size of the total population at any time t is equal to 1:

$$\begin{aligned}\int_{-\infty}^t N(s,t)\,ds &= \int_{-\infty}^t p\exp(-p(t-s))\,ds \\ &= [-\exp(-p(t-s))]_{-\infty}^t = 1.\end{aligned}$$

Notice that we have implicitly assumed zero population growth rate.

4.1.2 Availability of insurance (annuities)

Although there are no uncertainty in the aggregate, since individuals are uncertain about their time of death, there is scope for insurance.

The typical individual in the model who has accumulated wealth faces the possibility of dying before he can use it. He would be better off if he could sell the claim on his death in the event he dies, in exchange for command over resources while he is alive. There will thus be a demand for insurance that takes the form of the insurance company making premium payments to the living in exchange for receipt of their "estates" in the event they die.[8] This form of insurance is called an annuity.

We assume that insurance companies operate competitively so that they just break even. Then, the insurance premium must be p per unit time—individuals will pay a rate p to receive one good contingent on death. In the absence of a bequest motive and negative bequests prohibited (i.e. individuals cannot be in debt when they die), individuals will contract to have all of their non-human (i.e. financial) wealth, denoted v_t, go to the insurance company contingent on their deaths. In exchange, the insurance company will pay them a premium of pv_t per unit time.

The insurance company just balances its book. It receives payment from those who die, at rate pv_t, and pays out premia at the rate pv_t. It faces no uncertainty because the proportion of population dying per unit of time, p, is not stochastic.

Let r be the riskless interest rate. If everyone who is alive at time t invests v, then in period $t+\Delta$, the investment would be worth $v\left(1+\Delta r\right)$. However, only $\left(1-p\Delta\right)$ of those who were alive at time t is alive at $t+\Delta$. Thus, those who are alive at $t+\Delta$ receive, $v\left(1+\Delta r\right)/\left(1-p\Delta\right)$ from their investment of v at time t in time $t+\Delta$. Thus, their return is given by:[9]

$$v\frac{1+\Delta r}{1-p\Delta} = v + v\left(r+p\right)\Delta + o\left(\Delta\right).$$

Hence, the rate of return in continuous time is given by

$$\lim_{\Delta\to 0}\frac{\frac{1+\Delta r}{1-p\Delta}-1}{\Delta} = r + p + \frac{1}{v}\frac{o\left(\Delta\right)}{\Delta} = r+p.$$

4.1.3 Differential versus integral equations

The model we deal with here is micro-founded; i.e. aggregate variables are derived from individual optimisation problems. That is, aggregate variables are integral of individual variables. Individuals, in turn, are also solving a dynamic problem in which they find the optimal consumption profile over time. Thus, one way to solve the problem is to work with integral expressions. However, this is extremely difficult as there are infinitely many consumers in infinitely many time periods. To simplify the problem, we therefore convert the problem into a system of differential equations.

[8] This is the reverse of the usual sense in which people think of insurance in which the insured makes payments to an insurance company while alive in exchange for a payment if a certain event occurs.

[9] To see this, define $f\left(\Delta\right) := v\left(1+\Delta r\right)/\left(1-p\Delta\right)$. Taylor expansion around zero yields

$$\begin{aligned} f\left(\Delta\right) &= f\left(0\right) + f'\left(0\right)\Delta + o\left(\Delta\right) \\ &= v + v\left.\frac{r\left(1-p\Delta\right)+p\left(1+\Delta r\right)}{\left(1-p\Delta\right)^2}\right|_{\Delta=0}\Delta + o\left(\Delta\right) \\ &= v + v\left(r+p\right)\Delta + o\left(\Delta\right). \end{aligned}$$

4.2 Individual consumption

4.2.1 Objective function

Since we focus here on individual consumption, we will suppress the s index which denotes the date of birth of an agent.

Let $\theta \in (0, \infty)$ be the discount rate so 1 util at time $t + \Delta$ is worth $1/(1 + \Delta\theta)$ at t. Since every individual faces uncertainty about the timing of their death (note, this is the only uncertainty in the model), individuals face a maximisation problem under uncertainty: at time t, they maximise

$$\mathbb{E}_t \left[\int_t^\infty u\left(c\left(z\right)\right) e^{-\theta(z-t)} dz \right],$$

where $c(z)$ is the agent's consumption at time z and $e^{-\theta(z-t)}$ is the discount factor.

Agents have expected utility over consumption (only) while they are alive. The probability of being alive at time z is given by $e^{-p(z-t)}$. In case of death, utility is assumed to be zero; if alive, utility is $u(c(z))$. Hence, we can write

$$\begin{aligned} \mathbb{E}_t \left[\int_t^\infty u\left(c\left(z\right)\right) e^{-\theta(z-t)} dz \right] &= \int_t^\infty u\left(c\left(z\right)\right) e^{-\theta(z-t)} e^{-p(z-t)} dz \\ &= \int_t^\infty u\left(c\left(z\right)\right) e^{-(\theta+p)(z-t)} dz. \end{aligned} \tag{4.1}$$

Notice that the probability of death adds to the discounting; i.e. p increases the individual's rate of time preference.

4.2.2 Budget constraint and human wealth

To make the problem tractable, from now now, we simplify $u(\cdot)$ to log-utility.

If an individual has non-human wealth $v(t)$ at time t, he receives $r(t)v(t)$ in interest, where $r(t)$ is the interest rate, and he also receives $pv(t)$, which is the premium from the insurance company. Thus, the individual's dynamic budget constraint (while alive) is:

$$\frac{dv(t)}{dt} = (r(t) + p)\, v(t) + y(t) - c(t), \tag{4.2}$$

where $v(t)$ is non-human (i.e. financial) wealth at time t, and $y(t)$ is the labour income at time t. We also need a no-Ponzi-game (NPG) condition (see section 4.7.3):

$$\lim_{z \to \infty} v(z)\, R(t, z) = 0, \tag{4.3}$$

where $R(t, z)$ is the price of a good in time z in terms of goods in t:

$$R(t, z) := \exp\left[-\int_t^z (r(\mu) + p)\, d\mu \right].$$

Using the NPG condition, we can integrate the dynamic budget constraint to obtain the intertemporal budget constraint (see section 4.7.3):

$$v(t) = \int_t^\infty \left[c(z) - y(z)\right] R(t, z)\, dz, \tag{4.4}$$

We see that p enters into the discounting factor $R(t, z)$ (in addition to the riskless rate $r(z)$). Intuitively, this is because the probability of death makes the consumers more impatient.

We also define human wealth as the present value of all future income: i.e.

$$h(t) = \int_t^\infty y(z) R(t,z) \, dz \tag{4.5}$$

with the boundary condition:

$$\lim_{z \to \infty} h(z) R(t,z) = 0.$$

4.2.3 Optimal consumption

The household's problem is to maximise (4.1) subject to the budget constraint (4.4) with respect to $c(z)$ and $v(z)$. The solution is given by (see section 4.7.4):

$$c(t) = (\theta + p)(v(t) + h(t)). \tag{4.6}$$

We can interpret $(\theta + p)$ as the propensity to consume out of total individual wealth (i.e. sum of non-human and human wealth). Notice that consumption is independent of the interest rate (due to assumption of log-utility) and independent of the individual's age (in contrast to the OLG model; since probability of death is constant).

The law of motion (or, Euler equation) for consumption is

$$\frac{dc(t)}{dt} = (r(t) - \theta) c(t), \tag{4.7}$$

(see 4.7.4).

Comparison with infinitely lived consumers Recall that the Perpetual Youth Model converges to the infinitely lived consumer (Ramsey) model as $p \to 0$. Thus, we can think about how consumption of Blanchard-Yaari consumers differ to that of their infinitely lived counterparts. The probability of death, p, has two effects on consumption. First, as discussed above, p reduces the present value of the agents' wealth (as people are less patient when they know they might die in the future). However, p also induces consumers to consume a higher proportion of their total wealth (see (4.6)). In fact, this latter effect will dominate because Blanchard-Yaari consumers start off with zero wealth and, in a continuing steady state, would spend their life saving and increasing their capital stock. The reason why this can be a steady state is that the saving is offset by the fact that some individuals are dying and being replaced by new born who have zero wealth.

Capital accumulation We will see later that we require $r(t) > \theta$ in the steady state. From (4.7), this implies that, in the steady state, individual consumption rises over time. Now, imagine a steady state where labour income is constant. If consumption is rising, then there must be savings going on. Saving leads to accumulation of assets and this continues throughout life; agents continue to accumulate and save, which allows personal consumption to continue to grow (as we see below, aggregate consumption of each generation does not necessarily grow because each generation shrinks in size over time).

It might seem odd that an individual will carry on accumulating more and more assets as they grow older. First, we naturally should think that lives have a clear end point so that, as we reach the end-point, we should think about running down our assets. However, in this set up, your expected length of life is the same whether you are one or a hundred years old. Second, we think about the steady state in the infinite life case where there is no attempt to build up or run down assets but that is because the rate of interest in that steady state is equal to impatience, whereas here, it is greater so there is always a motive to save. (See section 4.7.5.)

Comparison with finitely lived consumers To see how (4.6) compares with the optimal consumption from the finite period model, recall that, in the finite consumption model with log utility, consumption of good i was given by a share of the present value of wealth/income equal to the terms in front of the log. More precisely, the solution to

$$\max_{c_1,c_2,\ldots,c_n} a_1 \ln(c_1) + a_2 \ln(c_2) + \cdots + a_n \ln(c_n)$$
$$s.t. \qquad R_1 c_1 + R_2 c_2 + \cdots + R_n c_n = v + h$$

was given by

$$c_i = \underbrace{\frac{a_i}{\sum_{i=1}^n a_i}}_{\text{share}} \underbrace{\frac{(v+h)}{R_i}}_{\text{PV of wealth}} .$$

Normalising $R_1 = 1$ and a_i's so that $\sum_{i=1}^n a_i = 1$ yields that $c_1 = a_1 (v+h)$.

In the Perpetual Youth Model, we have that

$$\int_t^\infty e^{-(\theta+p)(z-t)} dz = \left[-\frac{1}{\theta+p} e^{-(\theta+p)(z-t)} \right]_{z=t}^\infty = \frac{1}{\theta+p},$$

and $R(t,t) = 1$. That is, here, the weights have been normalised to $1/(\theta+p)$ so that

$$c(t) = (\theta+p)(v(t) + h(t)).$$

Just as in the finite case, consumption in this model is also independent of the the interest rate. This is due the the assumption of log utility, which implies that the substitution and income effect from changes in the interest rate exactly offset each other.

4.2.4 Summary

We summarise the results for the solution to the individual's problem below. We also re-introduce the s index which denotes the date of birth of an agent. The dynamics are given by

$$\frac{dc(s,t)}{dt} = (r(t) - \theta) c(s,t),$$
$$\frac{dv(s,t)}{dt} = (r(t) + p) v(s,t) + y(s,t) - c(s,t)$$

together with the boundary conditions:

$$\lim_{T\to\infty} v(s,T) \exp\left[-\int_s^T (r(\mu) + p) d\mu \right] = 0,$$
$$\lim_{T\to\infty} y(s,T) \exp\left[-\int_s^T (r(\mu) + p) d\mu \right] = 0,$$

where the second condition ensures that $h(t)$ is finite.

4.3 Aggregation

We use capital letters to denote aggregates. In the model, agent are not representative consumers; however, individuals born at time s (if they are alive) are identical; hence, the heterogeneity in the model comes from the existence of (infinitely) many generations at time t. Recall that we denote the size of generation s at time t as

$$N(s,t) = p\exp[-p(t-s)]. \tag{4.8}$$

4.3.1 Distribution of labour income

To derive the dynamic behaviour of human wealth, $Y(t)$, we must specify the distribution of labour income across people at any point in time. In doing so, we want to capture the idea of retirement; i.e. we want labour income to eventually decrease with age. Denote $Y(t)$ as the aggregate labour income of agents alive at time t.

The idea is to assume that each generation alive at time t receives a share of the aggregate labour income at time t, denoted $Y(t)$, but that older generation is endowed with a smaller share than younger generations. Let $y(s,t)$ denote the labour income in time t of an agent born at time s (who is alive). Then, we can model $y(s,t)$ using the exponential distribution, similar to the way we modelled the "time until death" variable, i.e.[10]

$$y(s,t) = \frac{\alpha+p}{p} Y(t) e^{-\alpha(t-s)},\ \alpha \geq 0. \tag{4.9}$$

This means that, at any particular point in time t, the share of $Y(t)$ endowed to generation s falls at the rate α as s increases.

4.3.2 Aggregate human wealth

The aggregate human wealth is defined as

$$H(t) \coloneqq \int_{-\infty}^{t} N(s,t)\, h(s,t)\, ds. \tag{4.10}$$

Substituting the expression for $y(s,t)$ from (4.9) and differentiating with respect to t gives that (see section 4.7.6):

$$\frac{dH(t)}{dt} = (r(t) + p + \alpha) H(t) - Y(t). \tag{4.11}$$

4.3.3 Aggregate non-human wealth

The aggregate non-human wealth is defined as

$$V(t) \coloneqq \int_{-\infty}^{t} N(s,t)\, v(s,t)\, ds.$$

[10] To see why we p appears in $y(s,t)$, first write

$$y(s,t) = aY(t)e^{-\alpha(t-s)}.$$

Then,

$$\begin{aligned} Y(t) &= \int_{-\infty}^{t} N(s,t)\, y(s,t)\, ds = \int_{-\infty}^{t} pe^{-p(t-s)} aY(t) e^{-\alpha(t-s)} ds \\ &= aY(t) \int_{-\infty}^{t} pe^{-(\alpha+p)(t-s)} ds = aY(t) \left[\frac{p}{\alpha+p} e^{-(\alpha+p)(t-s)} \right]_{-\infty}^{t} = aY(t) \frac{p}{\alpha+p} \\ \Rightarrow a &= \frac{\alpha+p}{p}. \end{aligned}$$

Differentiating above with respect to t while substituting in (4.8), and then using (4.2) (see section 4.7.7) gives that:

$$\frac{dV(t)}{dt} = r(t)V(t) + Y(t) - C(t). \tag{4.12}$$

Notice that this is different from that of the individual budget constraint, (4.2), where an individual received a return p on their assets while they are alive. However, assets for the economy as a whole do not have the term p because people are dying at rate p. In other words, the amount $pV(t)$ is a transfer (through insurance companies) from those who die to those who remain alive. Thus, the insurance does not add to aggregate wealth.

4.3.4 Aggregate consumption

Thus, aggregate consumption is defined as

$$C(t) := \int_{-\infty}^{t} N(s,t)\, c(s,t)\, ds.$$

However, since the propensity to consume from total wealth is independent of the age (agents have log utility), we may aggregate consumption as

$$C(t) = (p+\theta)(H(t) + V(t)). \tag{4.13}$$

Differentiating with respect to t and using (4.11) and (4.12) gives (see section 4.7.8):

$$\frac{dC(t)}{dt} = (r(t) + \alpha - \theta)C(t) - (p+\theta)(p+\alpha)V(t). \tag{4.14}$$

Setting $p = 0$ so that agents are infinitely lived, which implies that $\alpha = 0$, we see that the consumption dynamics is given by

$$\frac{dC(t)}{dt} = (r(t) - \theta)C(t).$$

Comparing (4.14) against above, we see that in the Perpetual Youth Model, the dynamics of consumption depends on the level on the aggregate non-human wealth $V(t)$. This is despite the fact that, at the individual level, the consumption dynamics does not depend on $v(t)$ (see equation (4.47) in section 4.7.4).

Notice that α, which appears in two places in (4.14), has two effects on the dynamics.

- ▷ First, α has the effect of increasing the consumption rate by $\alpha C(t)$. This can be seen as a correction for the fact that labour income falls as individuals grow older.
- ▷ Second, α decreases the rate of consumption by $(p+\theta)\alpha V(t)$. This can be seen as a correction for the fact that the new borns do not have wealth and must accumulate assets throughout their life. As younger individuals save more for their future, consumption rate falls. Notice that this term appears only when aggregating (compare (4.14) and (4.47)) and thus can be seen as an aggregate effect.

4.3.5 A summary

We have now derived the dynamics that describe the aggregate behaviour in the Perpetual Youth Model:

$$\begin{aligned}\frac{dH(t)}{dt} &= (r(t)+p+\alpha)\,H(t)-Y(t),\\ \frac{dV(t)}{dt} &= r(t)\,V(t)+Y(t)-C(t), &(4.15)\\ \frac{dC(t)}{dt} &= (r(t)+\alpha-\theta)\,C(t)-(p+\theta)(p+\alpha)\,V(t). &(4.16)\end{aligned}$$

We also need a no-Ponzi-game condition, which ensures that agents have finite wealth; i.e.

$$\lim_{T\to\infty} Y(t)\exp\left[-\int_t^T (r(z)+\alpha+p)\,dz\right]=0.$$

If the condition is violated, then agents would have unbounded human wealth and could borrow an unbounded amount. Notice that the expression means that if $\alpha+p>0$, then the boundary condition can be satisfied even if we have negative interest rates.

In the next section, we consider three cases of general equilibrium:

(i) Pure endowment economy: $Y(t)=C(t)$ and $V(t)=0$

(ii) Closed economy with capital accumulation

- ▷ Private sector buys capital: $V(t)=K(t)$
- ▷ If depreciation is zero, the rental rate of capital is $r(t)=F'(K(t))$ assuming everyone supplies labor inelastically.
- ▷ Since there is measure of 1 people working, using the fact that the production function is CRS:

$$F_Y Y = F - F_K K \Leftrightarrow Y(t)=F(K(t))-F'(K(t))K(t)$$

(iii) Adding government debt

4.4 General equilibrium in a pure endowment economy

Let us see what the equilibrium is in a pure endowment setting; i.e. when $Y(t)$ is given exogenously. Labor income is the only source of income. To simplify, we assume that $Y(t)$ is constant over time; i.e.

$$Y(t)=Y,\ \forall t.$$

Solving for $r(t)$ By market clearing condition, this implies that

$$Y=C(t),\ \forall t$$

so that aggregate consumption is constant in equilibrium. Since there is no capital in the economy, it must be that

$$V(t)=0,\ \forall t.$$

This condition does not mean that there is no savings or borrowings; i.e. $v(s,t)$ could be different from zero for some s in any period t. However, it does require there to be a net zero supply of savings/assets.

That $C(t) = Y$ (so that it is constant) implies that $dC(t)/dt = 0$. This, together with the fact that $V(t) = 0$ implies that (4.16) becomes

$$0 = (r(t) + \alpha - \theta) Y.$$

Since $Y > 0$, the equation holds (in equilibrium) if and only if

$$r(t) = \theta - \alpha.$$

Dynamics for Consumption Substituting this into the individual's dynamics for consumption, (4.7), yields

$$\begin{aligned} \frac{dc(s,t)}{dt} &= (r(t) - \theta) c(s,t) \\ &= -\alpha \\ \Leftrightarrow \frac{1}{c(s,t)} \frac{dc(s,t)}{dt} &= -\alpha. \end{aligned}$$

This is a differential equation in t where the left-hand side is equal to the derivative of $\ln c(s,t)$ with respect to t. We can solve the differential equation:

$$\begin{aligned} \int_s^t \frac{d}{dz} \ln c(s,z)\, dz &= -\int_s^t \alpha dz \\ \Leftrightarrow \ln c(s,t) - \ln c(s,s) &= -\alpha(t-s) \\ \Leftrightarrow c(s,t) &= c(s,s) e^{-\alpha(t-s)}. \end{aligned} \tag{4.17}$$

To pin down $c(s,s)$, we use (4.4) and noting that $v(s,s) = 0$ because each generation is born with no assets (recall that in writing (4.4) we had suppressed the reference to generation s):

$$\begin{aligned} 0 = v(s,s) &= \int_s^\infty [c(s,z) - y(s,z)] R(s,z)\, dz \\ &= \int_s^\infty [c(s,z) - y(s,z)] \exp\left[-\int_s^z (r(\mu) + p)\, d\mu\right] dz \\ &= \int_s^\infty [c(s,z) - y(s,z)] \exp\left[-\int_s^z (\alpha - \theta + p)\, d\mu\right] dz \\ &= \int_s^\infty [c(s,z) - y(s,z)] e^{-(\alpha-\theta+p)(z-s)} dz. \end{aligned}$$

Substituting (4.17) and (4.9) yields

$$\begin{aligned} 0 &= \int_s^\infty \left[c(s,s) e^{-\alpha(z-s)} - \frac{\alpha+p}{p} Y e^{-\alpha(z-s)}\right] e^{-(\theta-\alpha+p)(z-s)} dz \\ &= \left(c(s,s) - \frac{\alpha+p}{p} Y\right) \int_s^\infty e^{-(\theta+p)(z-s)} dz \\ &= \left(c(s,s) - \frac{\alpha+p}{p} Y\right) \left[-\frac{1}{\theta+p} e^{-(\theta+p)(z-s)}\right]_s^\infty \\ &= \left(c(s,s) - \frac{\alpha+p}{p} Y\right) \frac{1}{\theta+p}. \end{aligned}$$

Since $\theta + p \neq 0$, this equation holds if and only if

$$c(s,s) = \frac{\alpha + p}{p} Y.$$

Substituting this into (4.17) yields

$$\begin{aligned} c(s,t) &= \frac{\alpha + p}{p} Y e^{-\alpha(t-s)} \\ &= y(s,t). \end{aligned}$$

That is

$$c(s,t) = y(s,t), \ \forall t \geq s.$$

This, in turn, implies by (4.4) that

$$v(s,t) = 0, \ \forall t \geq s;$$

i.e. the equilibrium is autarky (i.e. there is no trade).

Summary In this setup, the equilibrium here is given by

$$\begin{aligned} Y(t) &= C(t) = Y, \\ V(t) &= 0, \\ r(t) &= \theta - \alpha, \\ c(s,t) &= y(s,t), \ \forall t \geq s, \\ v(s,t) &= 0, \ \forall t \geq s. \end{aligned}$$

Exercise 4.1. Exercises:

▷ *What did we assume for the distribution of wealth among those agents alive at time $t = s$?*

Since

$$y(s,s) = \frac{\alpha + p}{p} Y e^{-\alpha(s-s)} = \frac{\alpha + p}{p} Y,$$

the share of income is given by $\frac{\alpha+p}{p}$ at time $t = s$ for each generation s.

▷ *What is the equilibrium interest rate for $\alpha = 0$? How are the individual consumptions?*

Since

$$r(t) = \theta - \alpha,$$

when $\alpha = 0$, the equilibrium interest equals the individual's discount rate, θ. Moreover, $\alpha = 0$ implies that

$$y(s,t) = Y.$$

Together with the fact that $y(s,t) = c(s,t)$ for all $t \geq s$, we realise that

$$c(s,t) = Y, \ \forall t \geq s$$

when $\alpha = 0$.

▷ *What values of α makes sense if $p = 0$?*

Recall that p was the probability of death as well as the rate a which a new generation is born. So $p = 0$ implies that every generation lives forever and that no new generation is born. To make sense of these, we assume that all generations are born at the beginning of the period ($t = -\infty$). But this means that there is really only one generation in the model (we are back in the Ramsey model) and each individual is identical. Thus, it makes sense to ensure that each individual receives the same share of income all of the time; i.e. $\alpha = 0$ so that

$$y(s,t) = Y.$$

Intuitively, the idea of 'retirement' does not apply when agents are infinitely lived.

▷ *What is the equilibrium interest rate for $p = 0$? How are the individual consumptions?*

If $p = 0$, then we argued above that $\alpha = 0$. Hence, equilibrium interest rate is again θ.

▷ *Consider the class of allocations with $c(s,t) = Y$. Are these Pareto superior to the equilibrium?*

Recall that

$$c(s,t) = y(s,t) = \frac{\alpha + p}{p} e^{-\alpha(t-s)} Y.$$

For $c(s,t) = Y$ to be Pareto superior we require that

$$\frac{\alpha + p}{p} e^{-\alpha(t-s)} \geq 1, \ \forall t \geq s$$

with the inequality holding strict for at least some t. Rearranging gives

$$\begin{aligned} -\alpha(t-s) &\geq \ln \frac{p}{\alpha + p} \\ t - s &\leq -\frac{1}{\alpha} \underbrace{\ln \frac{p}{\alpha + p}}_{\leq 0}, \end{aligned}$$

where the right-hand side is some positive constant. However, the left-hand side grows to infinity. hence, the inequality cannot hold for all $t \geq s$. That is, $c(s,t) = Y$ is not Pareto superior.

▷ *Can $r(t) = \theta$ be an equilibrium?*

This can be an equilibrium if $\alpha = 0$, which makes sense when $p = 0$ so that agents are infinitely lived.

▷ *When is the equilibrum allocation Pareto optimal?*

Only when $\sum_{t=0}^{\infty} 1/p_t = +\infty$ (will be shown later).

4.5 General equilibrium in a closed economy with capital accumulation

We now introduce capital into the model that we have developed so far.

4.5.1 The setup

We now introduce capital as the only form of non-human wealth:

$$V(t) = K(t), \ \forall t. \tag{4.18}$$

Let $F(K(t))$ be the production function for net output (i.e. after depreciation):

$$F(K(t)) = \mathbb{F}(K(t), 1) - \delta K, \tag{4.19}$$

where $\mathbb{F}(K, L)$ is the (neoclassical) production function with constant returns to scale, and we normalise labour supply to be inelastic at $L = 1$.[11]

In the discrete time case, the law of motion for capital is given by

$$K(t + \Delta) = I(t) + (1 - \Delta\delta) K(t).$$

Taking the limit as $\Delta \downarrow 0$, then

$$\frac{dK(t)}{dt} = I(t) - \delta K(t).$$

The feasibility condition is that

$$I(t) + C(t) = \mathbb{F}(K(t), L(t)).$$

Substituting out $I(t)$ using the law of motion yields

$$\frac{dK(t)}{dt} + C(t) = \mathbb{F}(K(t), L(t)) - \delta K(t).$$

Equilibrium Payment to Factors First, define $F(K) \equiv \mathbb{F}(K, 1) - \delta K$ (taking K as exogenous) and assume zero depreciation. Then supposing that labour supply is inelastic at $L(t) = 1$ for all t, and using the definition of F above, we can write the law of motion as

$$\frac{dK(t)}{dt} = F(K(t)) - C(t).$$

Firms rent capital and labor from the households. Let $\upsilon(t)$ denote the rental rate of capital, then firm's problem in any period t is

$$\max_{K(t), L(t)} \quad \mathbb{F}(K(t), L(t)) - w(t) L(t) - \upsilon(t) K(t).$$

The first-order conditions imply that

$$\mathbb{F}_K(K(t), L(t)) = \upsilon(t), \ \mathbb{F}_L(K(t), L(t)) = w(t).$$

Since $\mathbb{F}$ is constant returns to scale, its partial derivatives are homogenous of degree zero. So, we may write

$$\mathbb{F}_K\left(\frac{K(t)}{L(t)}, 1\right) = \upsilon(t), \ \mathbb{F}_L\left(\frac{K(t)}{L(t)}, 1\right) = w(t).$$

[11] A neoclassical production function, $\mathbb{F}(K, L)$, satisfies the Inada conditions:

- $\mathbb{F}(0, 0) = 0$;
- $\mathbb{F}(K, L)$ is concave (i.e. negative semidefinite Hessian matrix)
- $\lim_{K \to 0} \partial\mathbb{F}/\partial K = \infty$ and $\lim_{L \to 0} \partial\mathbb{F}/\partial L = \infty$;
- $\lim_{K \to, \infty} \partial\mathbb{F}/\partial K = 0$ and $\lim_{L \to \infty} \partial\mathbb{F}/\partial L = 0$.

Note that

$$\begin{aligned}
\mathbb{F}_L\left(K\left(t\right),L\left(t\right)\right) &= \frac{\partial}{\partial L\left(t\right)}\left(L\left(t\right)\mathbb{F}\left(\frac{K\left(t\right)}{L\left(t\right)},1\right)\right)\\
&= \mathbb{F}\left(\frac{K\left(t\right)}{L\left(t\right)},1\right) - L\left(t\right)\mathbb{F}_K\left(\frac{K\left(t\right)}{L\left(t\right)},1\right)\frac{K\left(t\right)}{L\left(t\right)^2}\\
&= \mathbb{F}\left(\frac{K\left(t\right)}{L\left(t\right)},1\right) - \mathbb{F}_K\left(\frac{K\left(t\right)}{L\left(t\right)},1\right)\frac{K\left(t\right)}{L\left(t\right)}.
\end{aligned}$$

Letting $L\left(t\right)=1$, we have

$$\mathbb{F}_L\left(K\left(t\right),1\right) = \mathbb{F}\left(K\left(t\right),1\right) - \mathbb{F}_K\left(K\left(t\right),1\right)K\left(t\right).$$

Differentiating (4.19) with respect to $K\left(t\right)$ yields

$$F'\left(K_t\right) = \mathbb{F}_K\left(K\left(t\right),1\right) - \delta \Leftrightarrow \mathbb{F}_K\left(K\left(t\right),1\right) = F'\left(K_t\right) + \delta.$$

Hence,

$$\begin{aligned}
\mathbb{F}_L\left(K\left(t\right),1\right) &= \mathbb{F}\left(K\left(t\right),1\right) - \left(F'\left(K_t\right)+\delta\right)K\left(t\right)\\
&= \mathbb{F}\left(K\left(t\right),1\right) - \delta K\left(t\right) - F'\left(K_t\right)K\left(t\right)\\
&= F\left(K\left(t\right)\right) - F'\left(K_t\right)K\left(t\right)\\
&= w\left(t\right).
\end{aligned}$$

Since aggregate income $Y\left(t\right)$ represents labour income and there total labour supply equals the unit measure of total population

$$Y\left(t\right) = w\left(t\right) = F\left(K\left(t\right)\right) - F'\left(K\left(t\right)\right)K\left(t\right). \tag{4.20}$$

Moreover, since $F_K\left(K\left(t\right),1\right) = \upsilon\left(t\right)$, we have

$$F'\left(K_t\right) = \upsilon\left(t\right) - \delta.$$

Recall that in discrete time, in any equilibrium, we must have that $\upsilon_{t+\Delta} = r_t + \delta$. Suppose not. Consider the following strategy that an agent could follow: borrow money to fund purchase of (one unit) capital in period t, and rents the capital out. In period $t+\Delta$, the agent:

- ▷ receives the rental rate $\upsilon_{t+\Delta}$;
- ▷ pays r_t in interest;
- ▷ loses δ in depreciation.

If $\upsilon_{t+\Delta} > r_t + \delta$, the agent can borrow more to purchase more capital. If, on the other hand, $\upsilon_{t+\Delta} < r_t + \delta$, then those who would owns capital would wish to sell. Therefore, for there to be no arbitrage, then it must be that $\upsilon_{t+\Delta} = r_t + \delta$. Taking limits, as $\Delta \downarrow 0$,

$$\upsilon\left(t\right) = r\left(t\right) + \delta,$$

which allows us to write

$$F'\left(K\left(t\right)\right) = \upsilon\left(t\right) - \delta = r\left(t\right). \tag{4.21}$$

Notice that interest payment (i.e. return on non-human wealth) plus labour income equals output:

$$\begin{aligned} r(t) K(t) + Y(t) &= F'(K(t)) K(t) + F(K(t)) - F'(K(t)) K(t) \\ &= F(K(t)). \end{aligned} \tag{4.22}$$

Substituting (4.18), (4.21) and (4.22) into (4.15) and (4.16) yields the following dynamic equations for the production economy:

$$\dot{C} \coloneqq \frac{dC(t)}{dt} = [F'(K(t)) + \alpha - \theta] C(t) - (p+\alpha)(p+\theta) K(t), \tag{4.23}$$

$$\dot{K} \coloneqq \frac{dK(t)}{dt} = F(K(t)) - C(t). \tag{4.24}$$

4.5.2 Steady state and the phase diagram

Steady state is a point where $\dot{C} = 0$ and $\dot{K} = 0$:

$$\begin{aligned} &\dot{C} = 0 \Leftrightarrow C = \frac{(p+\alpha)(p+\theta)}{F'(K) + \alpha - \theta} K, \\ &\dot{K} = 0 \Leftrightarrow C = F(K). \end{aligned}$$

To study the dynamics and to find the steady state, we draw the phase diagram in (K, C) space. We first consider each expression above in turn:

- $\dot{C} = 0$: Since F is increasing and concave (i.e. $F'(\cdot) > 0$ and $F'(\cdot)$ is decreasing), thus C is an increasing function of K. It has an asymptote as K converges to $\hat{K}$ such that $F'\left(\hat{K}\right) = \theta - \alpha$. Since

  $$\lim_{K \to 0} F'(K) = \lim_{K \to 0} \mathbb{F}_K(K, 1) - \delta = \infty,$$

  C converges to zero as $K \to 0$.

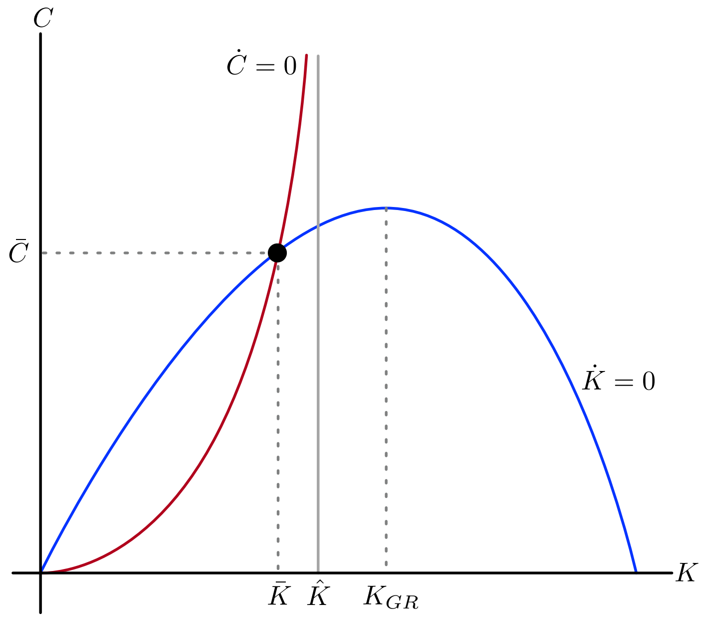
- $\dot{K} = 0$: Consumption traces out the (net) production function, which is concave and initially increasing. There is a maximum at $F'(K_{GR}) = 0$ and for $K \geq K_{GR}$,[12] and it converges to zero (with Inada conditions) as depreciation rate becomes greater than the marginal product of capital:

  $$\lim_{K \to \infty} F'(K) = \lim_{K \to \infty} \mathbb{F}_K(K, 1) = -\delta.$$

  So the slope becomes constant and equal to $-\delta$ as K becomes large.

The figure below plots the $\dot{C} = 0$ and $\dot{K} = 0$ locus in the (K, C) space. We assume that $\theta - \alpha > 0$ so that $\hat{K} < K_{GR}$.

[12] $\bar{K}$ is often referred to as the *golden rule* level.



We now think about dynamics of this system by considering movements above/below the constant consumption locus, and above/below the constant capital locus.

- ▷ Suppose we are above the constant consumption locus; i.e. $C > \frac{(p+\alpha)(p+\theta)}{F'(K)+\alpha-\theta}K$, multiplying both sides by the denominator on the right-hand side yields

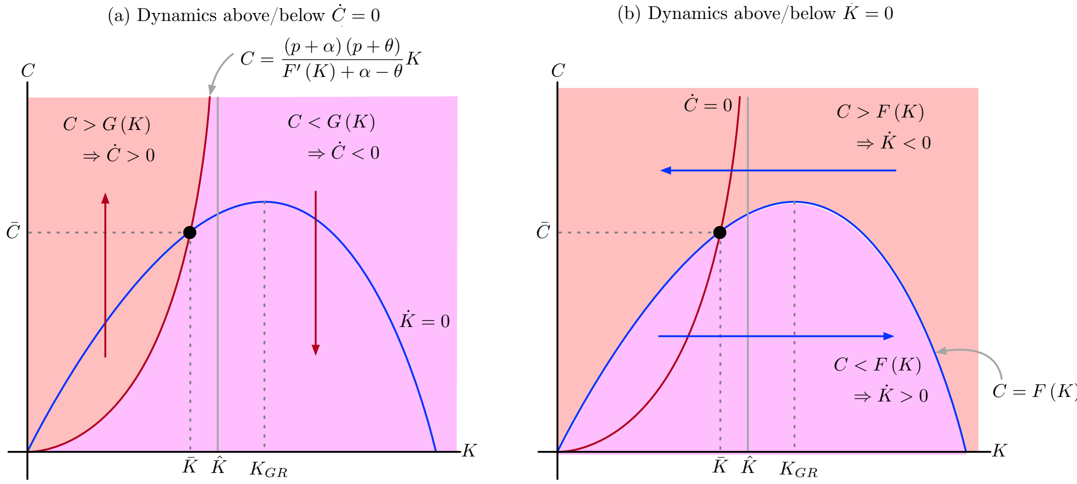
$$[F'(K(t)) + \alpha - \theta] C > (p+\alpha)(p+\theta) K.$$

  From (4.23), we can see that $dC(t)/dt > 0$ so that consumption must be increasing. Similar reasoning gives that consumption must be falling when we are below the constant consumption locus.

- ▷ Suppose that we are above the constant capital locus; i.e. $C > F(K)$. Then, from (4.24), it follows that $dK(t)/dt < 0$ so that capital must be falling (holding consumption constant). Similarly, if we are below the constant capital locus, then capital must be increasing.

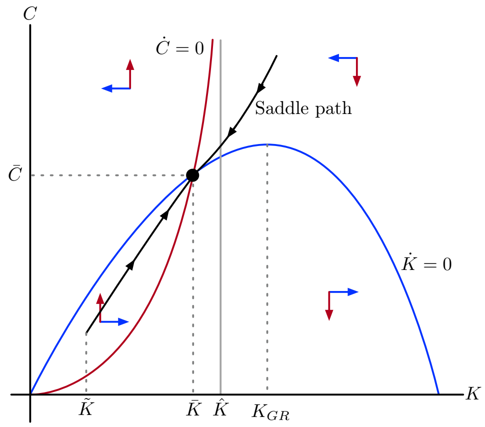
The figure below overlays the dynamics over the previous phase diagram.

(a) Dynamics above/below $\dot{C} = 0$ (b) Dynamics above/below $\dot{K} = 0$



The figure below combines the dynamics described above into a single figure. It shows that there is a *saddle path* that converges to the unique interior steady state, $\left(\bar{K}, \bar{C}\right)$, which is the intersection of $\dot{C} = 0$ and $\dot{K} = 0$, and where the steady state level of capital, $\bar{K}$, satisfies:

$$F\left(\bar{K}\right) = \frac{(p+\alpha)(p+\theta)}{F'\left(\bar{K}\right) + \alpha - \theta}\bar{K}.$$



In (a perfect foresight) equilibrium, the economy must be on the saddle path since any other path would converge (in finite time) to negative C or K (violating feasibility conditions).

Exercise. Questions:

▷ *What is the relationship between $C = S(K)$ and $C = (\theta + p)(K + H)$?*

$S(K)$ is the saddle path. Since $C = (\theta + p)(K + H)$ must hold at all times, it must be that

$$C = S(K) = (\theta + p)(K + H).$$

▷ *What happens to the graph if $\delta = 0$?*

The curve is only concavely increasing.

▷ *Is $S(\cdot)$ necessary linear?*

Not in general.

▷ *Assume $0 < \tilde{K} < \bar{K}$. What are the characteristics of the time path for $C(t)$ and $K(t)$. Are they monotone? What about $Y(t)$, $r(t)$ and $H(t)$?*

Starting from $\tilde{K} < \bar{K}$, both aggregate consumption, $C(t)$, and capital, $K(t)$, are increasing monotonically (since the slope of the saddle path is positive). Since F is a concave function and the interest rate is given by $r = F'(K)$, interest rate is falls as $K(t)$ increases over time. Recall from (4.20) that

$$\begin{aligned} Y &= F(K) - F'(K)K \\ \Rightarrow \frac{dY}{dK} &= F'(K) - F''(K)K - F'(K) \\ &= -F''(K)K > 0. \end{aligned}$$

Hence, output/labour income is increasing over time as K increases.

Since the only form of non-human wealth is capital (i.e. $V = K$), the interest rate is given by the marginal product of capital; i.e. $r = F'(K)$. Since F is concave, this means that interest rate, $r(t)$, is falling as $K(t)$ increases over time. $Y(t)$ is an exogenous process (it is the share of $Y(t)$ that a particular generation receives that falls at rate α) so the time path for $Y(t)$ is whatever we assume it to be. Suppose that $Y(t)$ is constant over time. Since aggregate labour income grows over time,

$$y(s,t) = \frac{\alpha + p}{p} Y(t)\, e^{-\alpha(t-s)}$$

grows over time and the interest rate is falling over time, it follows that $h(s,t)$ grows over time so that $H(t)$ grows over time (see (4.5) and (4.10)).

4.5.3 Comparative statics of the steady state

We now consider how the steady-state level of capital, $\bar{K}$, change as parameters α, p and θ change. Recall that the dynamics are given by the following equations:

$$\begin{aligned}
\dot{C} &= [F'(K(t)) + \alpha - \theta]\, C(t) - (p+\alpha)(p+\theta)\, K(t), \\
\dot{K} &= F(K(t)) - C(t).
\end{aligned}$$

The steady-state level of $\bar{K}$ satisfies

$$F(\bar{K}) = \frac{(p+\alpha)(p+\theta)}{F'(\bar{K}) + \alpha - \theta} \bar{K}.$$

To conduct comparative statistics, we can see how the right-hand side changes with the parameters; i.e. how the $\dot{C} = 0$ locus shifts with changes in the parameters.

Probability of death p As p increases, for any value of K, the right-hand side is greater; i.e. the $\dot{C} = 0$ locus shifts up/left. This means that the steady-state level of capital falls. Therefore, a higher p implies that expected life is shorter so that there is less incentive to save. Hence, the steady-state level of aggregate capital is smaller.

Discount rate θ Since

$$\begin{aligned}
\frac{\partial}{\partial \theta}\left(\frac{p+\theta}{F'(K) + \alpha - \theta}\right) &= \frac{(F'(K) + \alpha - \theta) + (p+\theta)}{(F'(K) + \alpha - \theta)^2} \\
&= \frac{F'(K) + \alpha + p}{(F'(K) + \alpha - \theta)^2} > 0,
\end{aligned}$$

we realise that an increase in θ also shifts the $\dot{C} = 0$ locus up/left. So, the steady-state level of capital again falls. Therefore, a higher θ implies that individuals value the future less so that they save less, which leads to a lower level of steady-state aggregate capital.

Rate of decline in labour income α Observe that

$$\begin{aligned}
\frac{\partial}{\partial \alpha}\frac{p+\alpha}{F'(K) + \alpha - \theta} &= \frac{(F'(K) + \alpha - \theta) - (p+\alpha)}{(F'(K) + \alpha - \theta)^2} \\
&= \frac{F'(K) - (\theta + p)}{(F'(K) + \alpha - \theta)^2}.
\end{aligned} \tag{4.25}$$

Hence, the comparative static depends on the sign of $F'(K) - (\theta + p)$.

Claim 4.1. If $\alpha = 0$, then $F'\left(\bar{K}\right) < \theta + p$.

Proof. By way of contradiction, suppose that $F'\left(\bar{K}\right) \geq \theta + p$. Then, we can find $\varepsilon \geq 0$ such that

$$F'\left(\bar{K}\right) = \theta + p\left(1+\varepsilon\right),\ \varepsilon \geq 0. \tag{4.26}$$

Substituting into (4.23), $\alpha = 0$, the expression above, we obtain

$$\begin{aligned}\dot{C} &= \left[F'\left(K\right) + \alpha - \theta\right] C - \left(p+\alpha\right)\left(p+\theta\right) K \\ &= p\left(1+\varepsilon\right) C - p\left(p+\theta\right) K.\end{aligned}$$

Since $\dot{C} = 0$ in equilibrium,

$$\left(1+\varepsilon\right)\bar{C} = \left(p+\theta\right)\bar{K}.$$

By market clearing, we also have that $\bar{C} = F\left(\bar{K}\right)$ so that

$$\left(1+\varepsilon\right) F\left(\bar{K}\right) = \left(p+\theta\right)\bar{K}.$$

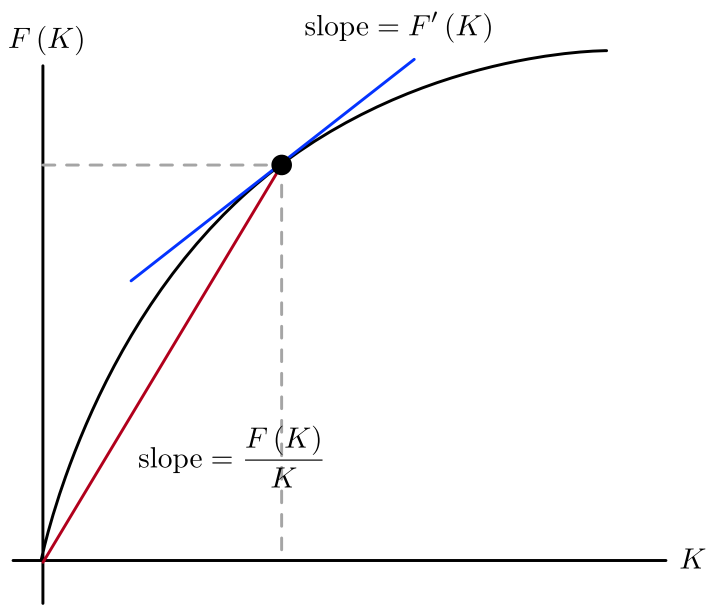
From (4.26), we know that $p + \theta = F'\left(\bar{K}_0\right) - p\varepsilon$ so that

$$\begin{aligned}\left(1+\varepsilon\right) F\left(\bar{K}\right) &= \left(F'\left(\bar{K}\right) - p\varepsilon\right)\bar{K} \\ \Leftrightarrow F\left(\bar{K}\right) - \bar{K}F'\left(\bar{K}\right) &= -\varepsilon\left(p\bar{K} + F\left(\bar{K}\right)\right).\end{aligned}$$

Notice that concavity of $F\left(\cdot\right)$ implies that

$$\frac{F\left(\bar{K}\right)}{\bar{K}} > F'\left(\bar{K}\right),$$

which says that the average productivity of capital is greater than the marginal productivity of capital. See also figure below.



Hence, we must have[13]

$$-\varepsilon\left(p\bar{K}+F\left(\bar{K}\right)\right)>0,$$

which holds if and only if $\varepsilon<0$ (since p, $\bar{K}$ and $F\left(\bar{K}\right)$ are all positive)—contradicting the fact that $\varepsilon\geq 0$. ■

The claim implies that when $\alpha=0$, the partial derivative (4.25) is strictly negative; i.e. a small increase in α from 0 shifts $\dot{C}=0$ locus to the right/down, which leads to a higher $\bar{K}$. But, for greater value of $\bar{K}$, $F'\left(\bar{K}\right)$ is even smaller by the concavity of F. Hence $\bar{K}$ is increasing for all (nonnegative) values of α.

4.5.4 Dynamic efficiency

Case: $\alpha=0$ Setting $\alpha=0$ implies that each generation at any given time t receives a constant share of the total aggregate income. In this case, the equilibrium level of capital, $\bar{K}$, solves:

$$F\left(\bar{K}\right)=\frac{p\left(p+\theta\right)}{F'\left(\bar{K}\right)-\theta}\bar{K}.$$

Recall that the golden rule of capital satisfies $F'\left(K_{GR}\right)=0$. Since output (i.e. consumption) cannot be negative—i.e. $F'\left(\bar{K}\right)>\theta$—it must be that $\bar{K}<K_{GR}$. Moreover, as $r=F'\left(K\right)$, we must also have that $r>0$. Together with what we established in Claim 4.1,

$$0<\theta<F'\left(\bar{K}\right)=r<\theta+p. \tag{4.27}$$

This means that, when relative labour income is constant (i.e. $\alpha=0$) and people have finite horizons (i.e. $\theta>0$), the interest rate, r, must be higher than the subjective discount rate θ. This ensures that individuals, who are born with no assets, has the incentive to save (recall that individually, they discount future utility by $\theta+p$) over their life time. If the interest rate is $r=\theta$, then there is no incentive to save/dissave so that there would be no aggregate capital accumulation. This cannot be an equilibrium since it implies $K\left(t\right)=0$ for all t; i.e. there will be no production (recall that $F\left(0\right)=0$ assumption), implying zero consumption.[14]

Unlike in the OLG model where equilibrium could be inefficient (when $r<0$), in this case, we obtain the result that the equilibrium is necessarily efficient (recall that population growth is zero in this model).

Case $\alpha>0$ Unlike in the case where $\alpha=0$, when $\alpha>0$, we will find that the equilibrium may be inefficient. The intuition is that, when an individual expects labour income to decline over time, he is likely to want to accumulate more capital to finance consumption later in life when income is low. This increased capital accumulation by individuals who take the path of interest rates as given is likely to drive the interest down.

The equilibrium level of capital solves

$$F\left(\bar{K}\right)=\frac{\left(p+\alpha\right)\left(p+\theta\right)}{F'\left(\bar{K}\right)+\alpha-\theta}\bar{K}.$$

Since, in the steady state, the denominator must be positive:

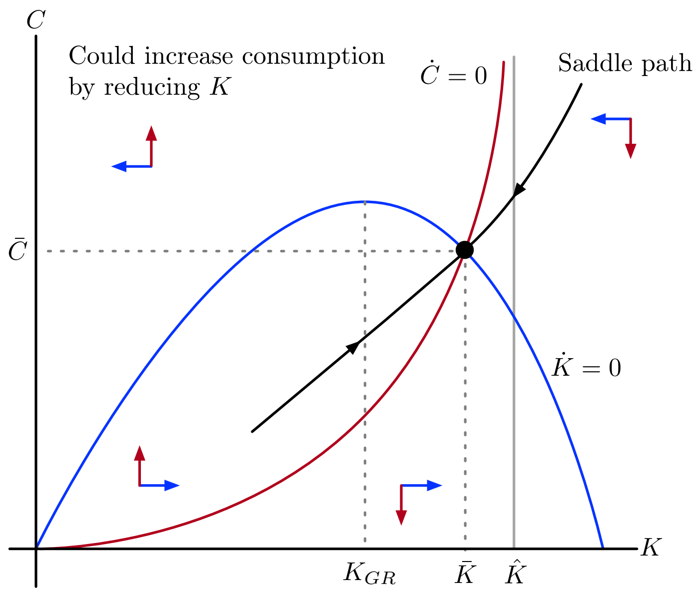
$$F'\left(\bar{K}\right)>\theta-\alpha. \tag{4.28}$$

Hence, the steady-state capital stock is such that the steady-state marginal product of capital exceeds $\theta-\alpha$. However, notice that $\theta-\alpha$ can be positive or negative.

[13]For example, if $F\left(K\right)=K^{0.5}$, then $F'\left(K\right)=0.5K^{-0.5}$ and $KF'\left(K\right)=0.5K^{0.5}<F\left(K\right)$.

[14]Consumption smoothing also implies that individuals would choose a flat consumption profile given that they receive constant labour income in each period.

If $\alpha > \theta$, $\bar{K}$ can be larger than the golden rule level of capital, K_{GR} (where $F'(K_{GR}) = 0$) so that $F'\left(\bar{K}\right)$, which equals the interest rate in equilibrium, could be negative. In such a case, the equilibrium is not efficient—intuitively, households are saving too much and everyone can be made better off (i.e. consumption increased) by reducing savings (see figure below). Importantly, this inefficiency arises even though agents are, individually, maximising their utility.



When $\alpha > 0$, we see that life cycle has two opposing effects on capital accumulation. The fact that people do not live forever leads to less capital accumulation (i.e. higher consumption); the shorter the horizon, the lower the steady-state level of capital. But, the fact that people retire (i.e. labour income declines over time), leads to more capital accumulation. The net effect is ambiguous, and if the second effect is strong enough, the economy may be dynamically inefficient.

4.5.5 Infinite horizon case: $p = 0$

Recall that if $p = 0$, then $\alpha = 0$. (You can't have any difference between people). In this case,

$$F'\left(\bar{K}\right) = r = \theta > 0 = F'(K_{GR})$$

so that $\bar{K} < K_{GR}$. In this case,

$$\begin{aligned} \dot{C} &= \left(F'(K(t)) - \theta\right) C(t), \\ \dot{K} &= F(K(t)) - C(t). \end{aligned}$$

Thus the two loci are given by

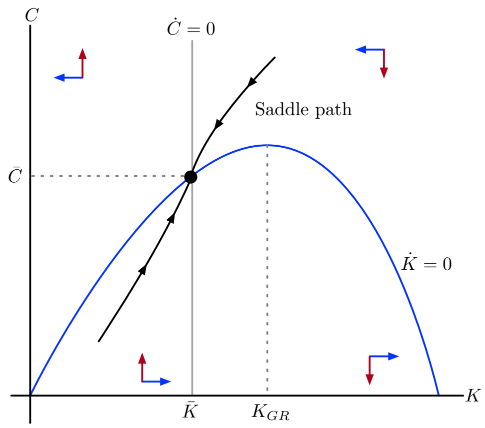
$$\begin{aligned} \dot{C} = 0 &: F'\left(\bar{K}\right) = \theta \\ \dot{K} = 0 &: C = F(K). \end{aligned}$$

The dynamics are as follows.

- ▷ $\dot{C} = 0$: If $K > \bar{K}$, then $F'(K) < F'\left(\bar{K}\right)$ so that $\dot{C} < 0$. Hence, to the right of $\dot{C} = 0$, consumption is decreasing. To the left, consumption is increasing.

▷ $\dot{K}=0$: If $C>\bar{C}$, then capital is falling; i.e. above the $\dot{K}=0$ locus, capital is falling and below, capital is increasing.

Thus, we essentially have the same dynamics as before. We will see later that this is the neoclassical growth model case (or the Ramsey model).



▷ This is equivalent to permanent income. You also get infinitely elastic long-run supply of capital (is this realistic?)

▷ On the OLG, on the other hand, the interest rates adjust to induce agents to save more.

4.6 Adding government and fiscal policy

The government budget constraint is the same whether or not people have finite lives. Let $B(t)$ denote the level of government debt at time t, $G(t)$ denote government purchases at time t and $T(t)$ denote the taxes levied on all agents alive at time t. The government's dynamic budget constraint is

$$\frac{dB(t)}{dt}=r(t)\,B(t)+G(t)-T(t) \tag{4.29}$$

with the boundary condition that

$$\lim_{T\to\infty} B(T)\,\mathcal{R}(t,z)\,dz,$$

where

$$\mathcal{R}(t,z):=\exp\left[-\int_t^z r(\mu)\,d\mu\right].$$

Notice that, in contrast to the discount factor for the agents, $R(t,z)$, the discount factor for the government, $\mathcal{R}(t,z)$, does not contain p. This is because the government can issue debt at the rate $r(t)$ and not $r(t)+p$ since, although individually agents demand a return of $r(t)+p$, p will die in the next "period".

Equivalently, in present-value terms, the government's budget constraint is given by

$$B(t)=\int_t^\infty (G(z)-T(z))\,\mathcal{R}(t,z)\,dz.$$

In this form, the budget constraint says that the current level of debt must be equal to the present value of primary surpluses. That is, if the government is currently a net debtor, it must intend to run surplus at some time in the future.[15]

[15] Since the government's discount factor does not contain p, here, we require $r>0$ to ensure that the government could not run a deficit forever without the debt exploding. If not, the government could issue debt equal to the negative interest rate it pays (i.e. the positive payment it receives from the public that is holding the debt).

This does not imply that debt is ultimately repaid or even that debt is ultimately constant—all it means is that the debt ultimately grows less rapidly than the interest rate.

4.6.1 Ricardian Equivalence (Partial Equilibrium)

We will consider the impact of a tax cut in time t funded by an increase in taxes after s periods. When agents were infinitely lived, recall that this tax policy had no effect on consumption; i.e. Ricardian Equivalence held. As we demonstrate below, the same cannot be said in the case of the Perpetual Youth Model. Intuitively, this is because Blanchard-Yaari individuals know that, with some probability, they will not be alive to pay the higher tax in the future, which leads them to consume more now.

Case $\alpha = 0$: Suppose that $\alpha = 0$; i.e. per capita labour income is independent of age. Now, imagine that the government cuts taxes at time t by an amount $dT(t)$ and funds this by an increase in taxes after s periods. The amount of the tax increase in period $t+s$, denoted $dT(t+s)$, must be

$$\begin{aligned} dT(t+s) &= -dT(t)\,\mathcal{R}(t,t+s) \\ &= -dT(t)\exp\left[\int_t^{t+s} r(\mu)\,d\mu\right] \end{aligned} \tag{4.30}$$

In words, the increase in taxes at time $t+s$ must be equal to the initial decrease at time t compounded at the rate of interest (for the government).

With government taxes, the equation for the human capital is given by (see 4.50)

$$\begin{aligned} H(t) &= \int_t^{\infty} (Y(z) - T(z))\,R(t,z)\,dz \\ &= \int_t^{\infty} (Y(z) - T(z))\exp\left[\int_t^{z} (r(\mu)+p)\,d\mu\right] dz \end{aligned}$$

since taxes are lump sum.[16]

$B(t)$ is unchanged at time t as the increase in the deficit, which is a flow, does not instantaneously change the stock of debt. Thus, this reallocation has an effect on aggregate demand at time t only to the extent that it has an effect on consumption through its effect on human wealth, given by

$$\begin{aligned} dH(t) &= -dT(t) - dT(t+s)\exp\left[-\int_t^{t+s} (r(\mu)+p)\,d\mu\right] \\ &= -dT(t) + dT(t)\exp\left[\int_t^{t+s} r(\mu)\,d\mu\right]\left[-\int_t^{t+s} (r(\mu)+p)\,d\mu\right] \\ &= -dT(t) + dT(t)\exp\left[-\int_t^{t+s} p d\mu\right] \\ &= -dT(t)\left(1 - e^{-ps}\right). \end{aligned} \tag{4.31}$$

Since we are considering a tax cut, $dT(t) < 0$, this means that tax policy change has a positive impact on human wealth for given pre-tax current and future income and given interest rates (this is why the analysis is based on partial equilibrium). This higher human wealth leads to higher consumption (see (4.13)).

The term $1 - e^{-ps}$ is the probability, for individuals currently alive, of not being around to have to pay the future taxes

[16] To see why this holds, follow the derivation in 4.7.6 using $\tilde{Y}(t) = Y(t) - T(t)$ instead of $Y(t)$.

when the government levies them. Thus, consumption increases because individuals know that with some probability, they will not be alive in the period in which they would have to pay the government taxes (mathematically, this appears as the discount rate in the government budget constraint being r while for the human wealth, the discount rate is $r+p$). Hence, it is unsurprising that the longer the taxes are deferred, the larger is the effect on the tax cut on human wealth.

When $p=0$ (i.e. agents are infinitely lived), (4.31) reduces to $dH(t)=0$ so that the timing of taxes has no impact on current human wealth, and thus no effect on aggregate demand. In this case, a tax cut, although it increases current disposable income, leads to individuals saving exactly as much as the increased disposable income so that consumption does not change at all. They do this because the individuals know that, sooner or later, they will have to pay the higher taxes in the future; therefore, they willingly absorb the debt issued by the government at an unchanged interest rate.

Case $\alpha > 0$: Suppose now that $\alpha > 0$. We first analyse the case when taxes also decline with age, at the same rate as wages do (i.e. taxes are proportional to labour income). Then, the equation for aggregate human capital is

$$H(t)=\int_t^{\infty}(Y(z)-T(z))\exp\left[\int_t^z(r(\mu)+\alpha+p)\,d\mu\right]dz.$$

Hence, if in period t, the government cuts taxes by dT_t, (4.30) remains unchanged. However, now, the effect on human wealth is different:

$$\begin{aligned} dH(t) &= -dT(t)-dT(t+s)\exp\left[-\int_t^{t+s}(r(\mu)+\alpha+p)\,d\mu\right] \\ &= -dT(t)\left(1-\exp\left[-\int_t^{t+s}(\alpha+p)\,d\mu\right]\right). \end{aligned}$$

Notice that $dH(t)$ is greater than when $\alpha=0$. This is because the probability of not being around to pay the future tax is now higher, which, in turn, is caused by the fact that when the individual is tax later, he is aware that the tax would be smaller (as he will be older).

This means that when taxes are raised equally among all generations at time t: i.e.

$$H(t)=\int_t^{\infty}\left(Y(z)\exp\left[\int_t^z(r(\mu)+\alpha+p)\,d\mu\right]-T(z)\exp\left[\int_t^z(r(\mu)+p)\,d\mu\right]\right)dz,$$

then the outcome will be the same as when $\alpha=0$.

4.6.2 Pure endowment with government

Let us assume that

$$Y(t)=Y,\ G(t)=G,\ \forall t;$$

i.e. both labour income and government purchases are constant over time. Market clearing implies

$$Y=C(t)+G,\ \forall t$$

so that aggregate consumption is constant in equilibrium. We no longer must have $V(t)=0$ since we have assets in the form of government bonds, $B(t)$. With pure endowment with government, we have that the individual's cumulated savings equal government debt: i.e.

$$V(t)=\int_{-\infty}^{t}N(s,t)\,v(s,t)\,ds=B(t).$$

The law of motion for aggregate consumption becomes

$$\frac{dC\left(t\right)}{dt}=\left(r\left(t\right)+\alpha-\theta\right)\left(Y-G\right)-\left(p+\alpha\right)\left(p+\theta\right)B\left(t\right).$$

Since aggregate consumption is constant in equilibrium, $dC\left(t\right)/dt=0$ in equilibrium so that

$$r\left(t\right)=\frac{\left(p+\alpha\right)\left(p+\theta\right)B\left(t\right)}{Y-G}+\left(\theta-\alpha\right). \tag{4.32}$$

Substituting this into the government's dynamic budget constraint, (4.29), yields

$$\frac{dB\left(t\right)}{dt}=\frac{\left(p+\alpha\right)\left(p+\theta\right)B\left(t\right)^{2}}{Y-G}+B\left(t\right)\left(\theta-\alpha\right)+G-T\left(t\right).$$

Setting $T\left(t\right)=G=0$ for all t, above simplifies to

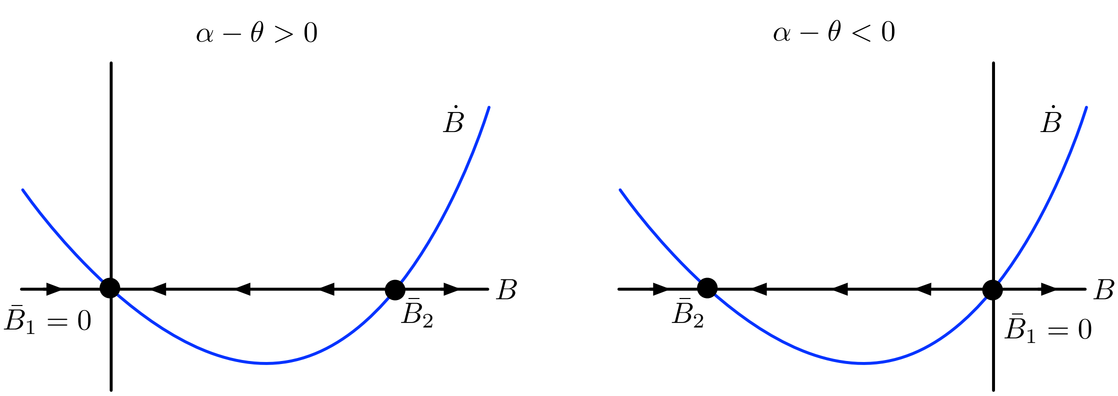
$$\begin{aligned}\frac{dB\left(t\right)}{dt}=\dot{B}&=\frac{\left(p+\alpha\right)\left(p+\theta\right)}{Y}B\left(t\right)^{2}+\left(\theta-\alpha\right)B\left(t\right)\\&=B\left(t\right)\left[\frac{\left(p+\alpha\right)\left(p+\theta\right)}{Y}B\left(t\right)+\left(\theta-\alpha\right)\right]\end{aligned} \tag{4.33}$$

Since the right-hand side is quadratic, the system has two steady states (set $dB\left(t\right)/dt=0$):

$$\bar{B}_{1}=0,\ \ \bar{B}_{2}=\frac{\left(\alpha-\theta\right)\bar{Y}}{\left(p+\alpha\right)\left(p+\theta\right)}.$$

Thus, if $\bar{r}=\theta-\alpha$, from (4.32), we see that $\bar{B}=0$, which, in turn, implies $\bar{V}=0$. Thus, $\bar{r}=\theta-\alpha$ is the autarky steady-state interest rate.

Suppose first that $\alpha-\theta>0$ so that $\bar{B}_{2}>0$. Then, we can draw $\dot{B}\left(t\right)$ as below.



Observe that the direction of arrows are determined by the sign of $\dot{B}\left(t\right)$. Hence, in this case, we see that $\bar{B}_{2}$ is unstable while $\bar{B}_{1}$ is stable. This means that, from any $B\left(0\right)<\bar{B}_{2}$, we would converge to $\bar{B}_{1}$. For any $B\left(0\right)>\bar{B}_{2}$, there is no convergence (so no equilibrium) and $B\left(0\right)=\bar{B}_{2}=B\left(t\right)$ for all t. What does $\alpha-\theta>0$ mean? Recall (4.28), which implies that

$$\bar{r}=F'\left(\bar{K}\right)>\theta-\alpha.$$

So, when $\bar{r}<0$, then

$$0>\bar{r}>\theta-\alpha\Rightarrow\alpha-\theta>0.$$

Thus, $\alpha - \theta > 0$ holds when the interest rate is negative (though the inequality may hold even if $\bar{r} > 0$).

Now suppose that $\alpha - \theta < 0$. Then, $\bar{B}_2 < 0$ and $\bar{B}_2$ becomes the stable root as can be seen from the figure above.

Exercise 4.2. Exercises:

▷ Identical to the case of 2-period OLG model with log utility. Try it.

[To do]

▷ How will the analysis change if $T - G \neq 0$?

If $T - G \neq 0$, then we would have:

$$\dot{B} = \frac{(p+\alpha)(p+\theta)}{Y-G} B(t)^2 + (\theta - \alpha) B(t) + G - T.$$

Since there is now a constant term, we no longer have one of the roots equalling zero, there will be two nonzero roots. Using the quadratic equation, the roots are given by

$$\begin{aligned}\bar{B} &= \frac{-(\theta-\alpha) \pm \sqrt{(\theta-\alpha)^2 - 4\frac{(p+\alpha)(p+\theta)}{\bar{Y}-\bar{G}}\left(\bar{G}-\bar{T}\right)}}{2\frac{(p+\alpha)(p+\theta)}{\bar{Y}-\bar{G}}} \\ &= \frac{(\alpha-\theta) \pm \sqrt{(\alpha-\theta)^2 - 4\frac{(p+\alpha)(p+\theta)}{\bar{Y}-\bar{G}}\left(\bar{G}-\bar{T}\right)}}{2\frac{(p+\alpha)(p+\theta)}{\bar{Y}-\bar{G}}}\end{aligned}$$

If $\bar{G} > \bar{T}$, then

$$\sqrt{(\alpha-\theta)^2 - 4\frac{(p+\alpha)(p+\theta)}{\bar{Y}-\bar{G}}\left(\bar{G}-\bar{T}\right)} < |\alpha - \theta|.$$

If, in addition, $\alpha - \theta > 0$, then both roots will be positive (roots could be complex if $\bar{G}$ is sufficiently greater than $\bar{T}$). If $\alpha - \theta < 0$, then both roots will be negative.

If, instead, $\bar{T} > \bar{G}$, then

$$\sqrt{(\alpha-\theta)^2 - 4\frac{(p+\alpha)(p+\theta)}{\bar{Y}-\bar{G}}\left(\bar{G}-\bar{T}\right)} > |\alpha - \theta|.$$

one root will be negative and the other will be positive.

▷ Which steady state is stable when $T - G \neq 0$? How does it depend on $\theta - \alpha$?

As was the case when $T - G = 0$, the stable root is the smaller root of the two (assuming the roots are real).

Remark 4.1. Check 2010/11 Mid-Term 1 for question on this.

4.6.3 Closed economy with capital accumulation and government

Assume $\alpha = 0$. With government debt, note that non-human wealth can now be either capital or government debt; i.e.

$$V(t) = K(t) + B(t), \tag{4.34}$$

We assume that government spending is separable from consumers' utility so that aggregate consumption function (4.13) remains unchanged. The dynamic equation for consumption, (4.23), then becomes,

$$\frac{dC(t)}{dt} = (r(t) - \theta) C(t) - p(p+\theta)(K(t) + B(t)).$$

Of course, the government must purchase goods from the economy too, so modifying (4.22) accordingly, market clearing condition becomes

$$F\left(K\left(t\right)\right)=Y\left(t\right)+r\left(t\right)K\left(t\right)+G\left(t\right), \tag{4.35}$$

and the law of motion for capital is now

$$\frac{dK\left(t\right)}{dt}=F\left(K\left(t\right)\right)-C\left(t\right)-G\left(t\right).$$

To summarise, we now have that

$$\begin{aligned}\frac{dC\left(t\right)}{dt}&=\left(r\left(t\right)-\theta\right)C\left(t\right)-p\left(p+\theta\right)\left(K\left(t\right)+B\left(t\right)\right),\\ \frac{dK\left(t\right)}{dt}&=F\left(K\left(t\right)\right)-C\left(t\right)-G\left(t\right),\\ \frac{dB\left(t\right)}{dt}&=rB\left(t\right)+G\left(t\right)-T\left(t\right).\end{aligned}$$

Constant debt Now consider a fiscal policy of keeping debt constant, by adjusting taxes. Then

$$\begin{aligned}\frac{dC\left(t\right)}{dt}&=\left(r\left(t\right)-\theta\right)C\left(t\right)-p\left(p+\theta\right)\left(K\left(t\right)+B\right),\\ \frac{dK\left(t\right)}{dt}&=F\left(K\left(t\right)\right)-C\left(t\right)-G\left(t\right),\\ T\left(t\right)&=rB+G\left(t\right),\\ r&=F'\left(K\right).\end{aligned}$$

Steady state, given B and G, is characterised by

$$\left(F'\left(K^*\right)-\theta\right)C^*=p\left(p+\theta\right)\left(K^*+B^*\right) \tag{4.36}$$

$$F\left(K^*\right)=C^*+G, \tag{4.37}$$

$$T^*=F'\left(K^*\right)B+G. \tag{4.38}$$

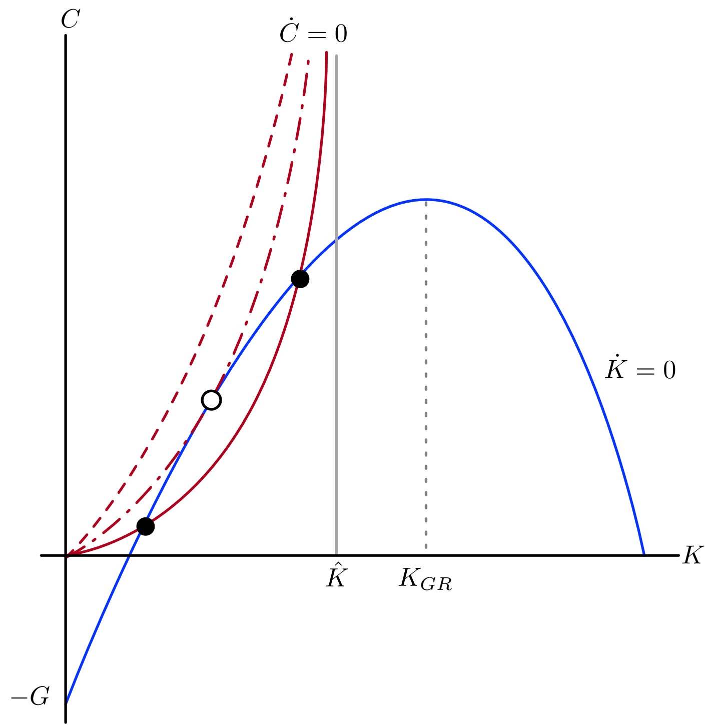
The steady-state level of capital solves

$$F\left(K^*\right)-G=\frac{p\left(p+\theta\right)\left(K^*+B\right)}{F'\left(K^*\right)-\theta}. \tag{4.39}$$

Exercise 4.3. Exercises

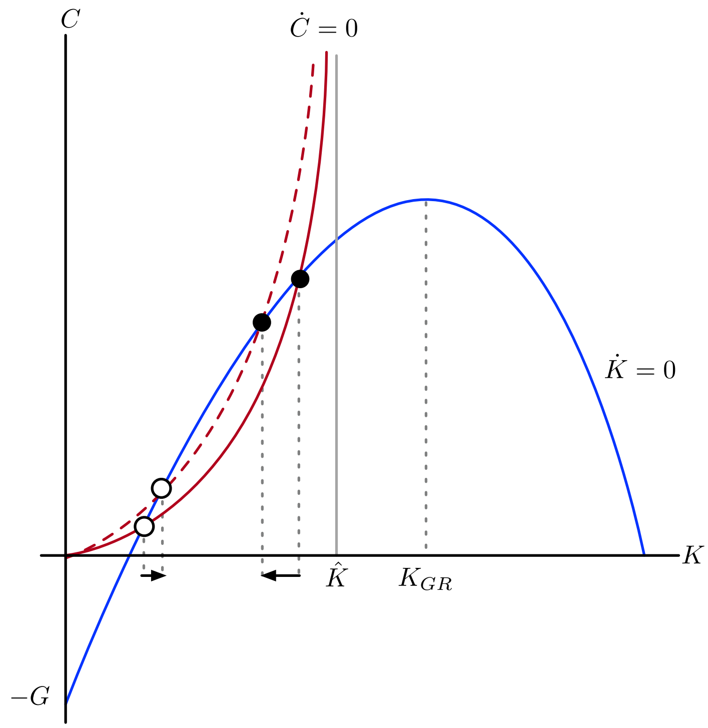
- ▷ Is there a unique steady state K^*?

  Note the the right-hand side converges to zero as $K\to 0$, and converges to infinity as $K\to\hat{K}$, where $F'\left(\hat{K}\right)=\theta$. The right-hand side equals $-G$ when $K=0$. Depending on the value of G and B, we may have none, one or two steady state. The figure below plots the case for a fixed value of $G>0$ and for different levels of B.



▷ What is the effect on K^*, r^* of an increase in the initial level B?

Consider the right-hand side of (4.39). For any fixed value of K, a higher B increases the value of the function. This means that $\dot{C} = 0$ locus shifts up/left. Consider the case in which there are two steady states. We can see from the figure below that: (i) for the steady state with a higher level of capital, an increase in B leads to a reduction in the steady-state capital; (ii) for the steady state with a lower level of capital, in contrast, an increase in B leads to an increase in the steady-state capital. Since $r^* = F'(K^*)$, in case (i) r^* rises while in case (ii), r^* falls.



- ▷ What is the effect on K^*, r^* of an increase on G?

  An increase in G works in the opposite way to an increase in the initial level of B with respect to the impact on K^* (and so r^*). We can see from the figure above that a higher G leads to a shift in the $\dot{K}=0$ locus down/right.

- ▷ How do the answers depend on p?

  A higher p leads to a greater shift in the $\dot{C}=0$ locus. Hence, the effect of changes in B and G on K^* (and so r^*) is greater with a higher value of p.

Non-constant debt Suppose we fix $G(t)=G$ and $T(t)=T$. Then, using the fact that $r(t)=F'(K(t))$, the steady states are given by:

$$\left(F'\left(K^*\right)-\theta\right)C^*=p\left(p+\theta\right)\left(K^*+B^*\right) \tag{4.40}$$

$$F\left(K^*\right)=C^*+G, \tag{4.41}$$

$$F'\left(K^*\right)B^*=T-G. \tag{4.42}$$

Note that the last equation does not imply that an increase in government spending reduces the national debt, rather that, if the government decides on a higher steady-state level of spending with the same steady-state level of taxes, it must, to balance the steady-state budget, reduce the national debt. Thus, to achieve this new steady state, it must temporarily increase taxes or decrease spending.[17]

[17] We implicitly held interest rate constant.

Using the steady-state equations, steady-state capital K^* solves

$$\begin{aligned} F(K^*) - G &= C^* \\ &= \frac{p(p+\theta)}{F'(K^*) - \theta}(K^* + B^*) \\ &= \frac{p(p+\theta)}{F'(K^*) - \theta}\left(K^* + \frac{T-G}{F'(K^*)}\right). \end{aligned} \tag{4.43}$$

Exercise 4.4. Exercises

▷ Is there a unique steady state? Signs of B and r at steady state. Compare with pure endowment.

The steady state is given by the intersection two loci:

$$\begin{aligned} \dot{C} = 0 : C &= \frac{p(p+\theta)}{F'(K^*) - \theta}\left(K^* + \frac{T-G}{F'(K^*)}\right) \\ \dot{K} = 0 : C &= F(K^*) - G. \end{aligned}$$

For $T - G > 0$, $\dot{C} = 0$ locus has the same properties as before: strictly increasing, asymptotes at $\hat{K}$ where $F'\left(\hat{K}\right) = \theta$ and converges to zero as $K \to 0$. As the figure below shows, when $T - G > 0$, there can be none, one or two steady states. From (4.42), we see that the sign of B^* depends on the sign of $T - G$ since $F'(K^*) > 0$, Moreover, since $F'(K^*) = r^*$, we realise that interest rate must always be positive. That is, unlike in the pure endowment case, r^* cannot be negative.

If, on the other hand, $T - G < 0$, then we cannot be certain that $\dot{C} = 0$ is increasing so the comparative statistic is more difficult.

▷ How do K^*, B^* and r^* vary with G and/or T? What is the difference between increases in G and decreases in T?

Observe that an increase in G affects both loci, while an increase in T affects only the $\dot{C} = 0$ locus. A decrease in T shifts the $\dot{C} = 0$ locus down/right. We see that the comparative statistic in this case is similar to the one studied above with constant debt. In contrast, an increase in G has the same effect on the $\dot{C} = 0$ locus, while also shifting the $\dot{K} = 0$ locus down. Hence, the effect on K^* is ambiguous (and so B^* and r^*).

▷ What is the effect on steady state quantities if p increases? What is the limit case as $p \downarrow 0$?

As before, a higher p exacerbates the impact of changes in T and G on the $\dot{C} = 0$ locus. As $p \downarrow 0$, we move to the infinite horizon model (i.e. neoclassical model with government). Then, as we aw earlier

$$\begin{aligned} \dot{C} = 0 &: F'\left(\bar{K}\right) = \theta \\ \dot{K} = 0 &: C = F(K) - G. \end{aligned}$$

Thus, the $\dot{C} = 0$ locus becomes a vertical line so that there will be a unique steady state.

Tax cuts Suppose now that the government temporarily cuts taxes, holding its spending constant. This raises the deficit and increases the debt. Then, later, the government raises taxes to bring the budget back to balance and thereafter keeps adjusting taxes to maintain a constant level of debt. Thus, we are considering a sequence of policy actions that ends up increasing the debt but leaving G unaffected.

For given levels of government spending, what happens if government debt is higher? This requires an initial cut in taxes but, thereafter, a permanently higher levels of taxes to finance the additional debt.

We can show that: (see section 4.7.10)

$$\frac{dK^*}{dB^*} = \frac{p\left(p+\theta\right)}{F''\left(K^*\right)C^* - \frac{Y-T}{K^*+B^*}\left(r^*-\theta\right)}.$$

We therefore find that the stock of government debt matters because it displaces capital from the portfolio of savers. A sufficient condition for an increase in the debt to decrease the capital stock is $Y > T$ and $K^* + B^* > 0$ (since $F'' < 0$). Note that $(K^* + B^*) > 0$ means that, if government debt is negative (i.e. surplus), then it is not too negative. (4.40) also implies $r^* > 0$. Thus, the equilibrium in the model will be dynamically efficient. This is the same result as we obtained when $\alpha = 0$ without government.

The analysis here suggests that the government could choose a level of government debt that produces the golden rule of capital; however, this implies the government should be holding assets, not debt. In other words, either the government should invest in capital directly, or encourage the private sector to do by so lending to it.

We could similarly examine the effects of an increase in G on the capital stock, holding constant the level of the primary deficit (or surplus) by a corresponding increase in taxes: (see section 4.7.10)

$$\frac{dK^*}{dG} = \frac{r-\theta}{F''\left(K^*\right)C^* - \frac{Y-T}{K^*+B^*}\left(r^*-\theta\right)}.$$

The derivative is negative if $Y > T$ and $K^* + B^* > 0$; i.e. an increase in G reduces the steady-state capital stock.

4.7 Appendix

4.7.1 Derivation of the size of cohort born in time s as of time t

It helps to think in terms of discrete time to see why this might be true. Let $N(s,t)$ be the size of the cohort born in time s as of time t. We want to think about the cohort size of generation s at time $t+\Delta$. During the interval Δ, $p\Delta$ of the generation dies. Hence, the size of the cohort at time $t+\Delta$ is given by

$$N\left(s,t+\Delta\right) = N\left(s,t\right)\left(1-p\Delta\right)$$
$$\Leftrightarrow \frac{N\left(s,t+\Delta\right) - N\left(s,t\right)}{\Delta} = -N\left(s,t\right)p.$$

We can then obtain the differential equation as:

$$\lim_{\Delta\downarrow 0}\frac{N\left(s,t+\Delta\right) - N\left(s,t\right)}{\Delta} = \frac{dN\left(s,t\right)}{dt} = -pN\left(s,t\right),$$

which can be rearranged as

$$\frac{1}{N\left(s,t\right)}\frac{dN\left(s,t\right)}{dt} = -p.$$

This is a differential equation where the left-hand side is the derivative of $\ln N(s,t)$ with respect to t; i.e.

$$\begin{aligned}\frac{d}{dt}\ln N(s,t) &= -p\\ \int_s^t \ln N(s,t)\,dt &= -\int_s^t pdt\\ &-p(t-s)+C\\ \Rightarrow N(s,t) &= A\exp(-p(t-s)).\end{aligned}$$

Adding the boundary condition is that $N(s,s)=p$ (since we normalised the initial cohort size to p) gives that

$$N(s,t) = p\exp(-p(t-s)).$$

4.7.2 Variable interest-rate discount factor

Recall that the relationship between intertemporal prices and the interest rates were given by

$$\frac{p_{t+1}}{p_t} = \frac{1}{1+r_t} \Leftrightarrow 1+r_t = \frac{p_t}{p_{t+1}} = \frac{1}{p_{t+1}},$$

where we normalised $p_t = 1$. Here, write $R(t,z)$ as the intertemporal price of goods in period z as at period t. Now consider the discrete time case with Δ being the length of the time period. Between time t to z, there are $(z-t)/\Delta$ periods. Then ,

$$(1+r_{t+1}\Delta)(1+r_{t+2}\Delta)\cdots(1+r_z\Delta) = \prod_{i=1}^{\frac{t-z}{\Delta}}(1+r_{t+i}\Delta) = \frac{1}{R(t,z)}.$$

Taking logs,

$$-\log R(t,z) = \sum_{i=1}^{\frac{t-z}{\Delta}}\log(1+r_{t+i}\Delta).$$

Taylor expansion of log around $\Delta = 0$ yields

$$\begin{aligned}\log(1+r_{t+i}\Delta) = & \log(1+r_{t+i}\Delta)|_{\Delta=0} + \left(\left.\frac{r_{t+i}}{1+r_{t+i}\Delta}\right|_{\Delta=0}\right)(\Delta-0)+o(\Delta)\\ &= \left(r_{t+i}+\frac{o(\Delta)}{\Delta}\right)\Delta.\end{aligned}$$

where $\lim_{\Delta\to 0} o(\Delta)/\Delta = 0$. Hence,

$$\begin{aligned}\lim_{\Delta\downarrow 0}\log R(t,z) &= -\lim_{\Delta\downarrow 0}\sum_{i=1}^{\frac{t-z}{\Delta}}\left(r_{t+i}+\frac{o(\Delta)}{\Delta}\right)\Delta\\ &= -\int_t^z r(\mu)\,d\mu,\end{aligned}$$

where the last step follows by the definition of Riemann integral. In this setup, the discount rate is $r(\mu)+p$ so, with variable interest rates, the intertemporal prices $R(t,z)$ is given by

$$R(t,z) \coloneqq \exp\left[-\int_t^z (r(\mu)+p)\,d\mu\right].$$

4.7.3 Derivation of the intertemporal budget constraint

The intertemporal budget constraint for the individual is given by:

$$\frac{dv(z)}{d(z)} = (r_z + p) v_z + y_z - c_z,$$

where we use subscript z to denote a variable to be a function of z; e.g. $v_z \equiv v(z)$. Notice that the return on assets include the annuity payment pv_z that the insurance company pays for receiving the individual's assets if they die.

Suppose first that interest rates were constant at r. Rearranging (4.2) and multiplying both sides by the discount factor, we obtain a differential equation:

$$\begin{aligned} y_z - c_z &= \frac{dv_z}{dz} - (r+p) v_z \\ \Rightarrow (y_z - c_z) e^{-(r+p)(z-t)} &= e^{-(r+p)(z-t)} \left[\frac{da_z}{dz} - (r+p) v_z\right] \\ &= \frac{d\left[v_z e^{-(r+p)(z-t)}\right]}{dz}. \end{aligned}$$

Integrating both sides

$$\begin{aligned} \int_{z=t}^{\infty} (y_z - c_z) e^{-(r+p)(z-t)} dz &= \int_{z=t}^{\infty} \frac{d\left[v_z e^{-(r+p)(z-t)}\right]}{dz} dz \\ &= \left[v_z e^{-(r+p)(z-t)}\right]_{z=t}^{\infty} \\ &= \left(\lim_{z\to\infty} rv(z)\right) - v_t, \\ &= -v_t, \end{aligned} \tag{4.44}$$

Notice that we have assumed that $\lim_{z\to\infty} rv(z) = 0$.

With variable interest rates, recall that

$$R(t,z) := \exp\left[-\int_t^z (r_\mu + p) d\mu\right].$$

Replacing the discounting term in (4.44) with $R(t,z)$ gives

$$v(t) = \int_{z=t}^{\infty} (c(z) - y(z)) R(t,z) dz. \tag{4.45}$$

The condition that $\lim_{z\to\infty} rv(z)$ can be interpreted as a no-Ponzi game (NPG) condition. It prevents the individual from going infinitely into debt and protecting themselves by buying life insurance. With variable interest rates, the NPG condition, if an individual is still alive at time z, is given by

$$\lim_{z\to\infty} v(z) R(t,z) = 0. \tag{4.46}$$

We can interpret the condition as follows:

- ▷ Suppose that (4.46) is strictly positive; i.e. the present value of non-human capital is positive in the limit. However, this would mean that the agent did not maximise utility as he could have used the capital to increase consumption. Hence, $R(t,z) v(z) > 0$ is required to ensure that a solution to the problem exists.

▷ Suppose instead that (4.46) is strictly negative. This means that the individual accumulates debt forever at a rate higher than the effective rate of interest facing him. Thus, the condition rules out the possibility of a Ponzi scheme (i.e. paying debt with new higher debt). This is why (4.46) is often called the *no Ponzi scheme condition*.

Similarly, we can show that

$$\frac{dh_z}{dz} = (r_z + p)\, h_z - y_z \Rightarrow \int_t^{\infty} y(z)\, R(t,z)\, dz$$

with the boundary condition:

$$\lim_{z\to\infty} R(t,z)\, h(z) = 0.$$

we can interpret this condition in a similar way to (4.46). If this limit does not converge, then the agent's human wealth is unbounded and there will be no solution to the problem.

4.7.4 Derivation of the optimal consumption plan

The household's problem is

$$\begin{aligned} \max_{c_z} & \int_t^{\infty} u(c_z)\, e^{-(\theta+p)(z-t)} dz \\ s.t. \quad & \frac{dv_z}{dz} = (r_z + p)\, v_z + y_z - c_z. \end{aligned}$$

The Lagrangian is given by

$$\mathcal{L} = \int_t^{\infty} \left(e^{-(\theta+p)(z-t)} \left[\ln(c_z) + \lambda_z \left((r_z + p)\, v_z + y_z - c_z - \dot{v}_z \right) \right] \right) dz.$$

Using integration by parts, we can write

$$\begin{aligned} \int_t^{\infty} e^{-(\theta+p)(z-t)} \lambda_z \dot{v}_z dz &= \left[e^{-(\theta+p)(z-t)} \lambda_z v_z \right]_t^{\infty} - \int_t^{\infty} v_z \frac{d}{dz} \left(e^{-(\theta+p)(z-t)} \lambda_z \right) dz \\ &= C - \int_t^{\infty} v_z e^{-(\theta+p)(z-t)} \left(\dot{\lambda}_z - (\theta+p)\, \lambda_z \right), \end{aligned}$$

where C is a constant. Substituting this expression into $\mathcal{L}$ yields

$$\mathcal{L} = \int_t^{\infty} \left(e^{-(\theta+p)(z-t)} \left[\underbrace{\ln(c_z) + \lambda_z \left((r_z+p)\, v_z + y_z - c_z \right)}_{:=H(\cdot)} + v_z \left(\dot{\lambda}_z - (\theta+p)\, \lambda_z \right) \right] \right) dz + C.$$

The first-order conditions are

$$\begin{aligned} \frac{\partial H(\cdot)}{\partial c_z} = 0 \Leftrightarrow \lambda_z &= \frac{1}{c_z} \\ \Rightarrow \dot{\lambda}_z &= \frac{d}{dz}\left(\frac{1}{c_z} \right) = -\frac{\dot{c}_z}{c_z^2}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial H(\cdot)}{\partial v_z} + \dot{\lambda}_z - (\theta+p)\, \lambda_z = 0 \\ \Rightarrow \lambda_z (r_z + p) + \dot{\lambda}_z - (\theta+p)\, \lambda_z = 0. \end{aligned}$$

Substituting the expression for λ_z and $\dot{\lambda}_z$ yields

$$\begin{aligned}
\frac{1}{c_z}\left(r_z + p - \theta - p\right) &= \frac{\dot{c}_z}{c_z^2} \\
\Leftrightarrow \dot{c}_z = \frac{dc_z}{dz} &= \left(r_z - \theta\right) c_z, \qquad (4.47)
\end{aligned}$$

which gives the Euler equation for this maximisation problem.

To obtain the level of consumption at time z when the consumer is at time $t < s$, we integrate from t to s:

$$\begin{aligned}
\int_t^z \left(r_z - \theta\right) dz &= \int_t^z \frac{dc_z}{dz}\frac{1}{c_z} dz \\
&= \left[\ln\left(c_z\right)\right]_t^z \\
&= \ln\left(\frac{c_z}{c_t}\right) \\
\Rightarrow c_z &= c_t \exp\left[\int_t^z \left(r_z - \theta\right) dz\right].
\end{aligned}$$

Substituting this expression into the intertemporal budget constraint (4.45) yields:

$$\begin{aligned}
v_t &= \int_{z=t}^{\infty} \left(c_z - y_z\right) R\left(t, z\right) dz \\
\Leftrightarrow v_t + h_t &= \int_{z=t}^{\infty} c_z R\left(t, z\right) dz \\
&= c_t \int_{z=t}^{\infty} \exp\left[\int_t^z \left(r_z - \theta\right) dz\right] \exp\left[-\int_t^z \left(r_z + p\right) dz\right] dz \\
&= c_t \int_{z=t}^{\infty} \exp\left[-\int_t^z \left(p + \theta\right) dz\right] dz \\
&= c_t \int_{z=t}^{\infty} \exp\left[-\left(p + \theta\right)\left(z - t\right)\right] dz \\
&= c_t \exp\left[\left(p + \theta\right) t\right] \int_{z=t}^{\infty} \exp\left[-\left(p + \theta\right) z\right] dz \\
&= c_t \exp\left[\left(p + \theta\right) t\right] \left[-\frac{1}{p + \theta} \exp\left[-\left(p + \theta\right) z\right]\right]_{z=t}^{\infty} \\
&= \frac{c_t}{p + \theta} \exp\left[\left(p + \theta\right) t\right] \exp\left[-\left(p + \theta\right) t\right] \\
&= \frac{c_t}{p + \theta} \\
\Leftrightarrow c_t &= \left(p + \theta\right)\left(v_t + h_t\right).
\end{aligned}$$

4.7.5 Illustration of continuous growth of assets

To see the continuous growth of assets, suppose that labour income and interest rate is constant over time (i.e. $y_t = y$ and $r_t = r$ for all t) . Then, the budget constraint (4.2) becomes

$$\frac{dv_t}{dt} = \left(r + p\right) v_t + y - c_t. \qquad (4.48)$$

We can also rewrite human wealth at time t, (4.5),

$$\begin{aligned}
h_t &= \int_t^\infty y_z R(t,z)\,dz \\
&= \int_t^\infty y \exp\left[-(r+p)(z-t)\right] dz \\
&= y \exp\left[(r+p)t\right] \int_t^\infty \exp\left[-(r+p)z\right] dz \\
&= y \exp\left[(r+p)t\right] \left[-\frac{1}{r+p}\exp\left[-(r+p)z\right]\right]_{z=t}^\infty \\
&= \frac{y}{r+p}\exp\left[(r+p)t\right]\exp\left[-(r+p)t\right] \\
&= \frac{y}{r+p}.
\end{aligned}$$

This implies that human capital is simply the labour income discounted by the return on an asset. We can now write consumption as

$$c_t = (\theta+p)\left(v_t + \frac{y}{r+p}\right).$$

Substituting this into (4.48) yields

$$\begin{aligned}
\frac{dv_t}{dt} &= (r+p)v_t + y - \left[(\theta+p)\left(v_t + \frac{y}{r+p}\right)\right] \\
&= (r-\theta)v_t + y\left(1 - \frac{\theta+p}{r+p}\right) \\
&= (r-\theta)v_t + y\left(\frac{r-\theta}{r+p}\right). \qquad (4.49)
\end{aligned}$$

Notice that if $r > \theta$, then $dv_t/dt > 0$ so that the individual will always be accumulating assets as long as they are alive—this implies that individuals never reaches a steady state.

We can also derive the growth rate of assets using (4.49) as

$$\frac{dv_t/dt}{v_t} = (r-\theta) + \frac{y}{v_t}\left(\frac{r-\theta}{r+p}\right).$$

Since we know that v_t is always increasing, then we can see that the growth rate of assets tends to $r-\theta$ over time.

Note that agent are not representative consumers—each generation will hold different amount of assets.

4.7.6 Derivation of law of motion for Aggregate Human Wealth

Substituting in the expression for $y(s,t)$, (4.9), in the expression for $h(t)$, (4.5), gives

$$\begin{aligned}
h(s,t) &= \int_t^\infty y_z R(t,z)\,dz \\
&= \int_t^\infty \left(\frac{\alpha+p}{p} Y_z e^{-\alpha(z-s)}\right) R(t,z)\,dz. \\
&= \frac{\alpha+p}{p}\left[\int_t^\infty Y_z e^{-\alpha(z-t)} R(t,z)\,dz\right] e^{-\alpha(t-s)},
\end{aligned}$$

where the term in $[\cdot]$ is independent of time of birth s.

We can substitute the expression for $h(s,t)$ into the definition of H_t:

$$\begin{aligned} H_t &= \int_{-\infty}^{t} N(s,t)\, h(s,t)\, ds \\ &= \int_{-\infty}^{t} \left(pe^{-p(t-s)}\right)\left(\frac{\alpha+p}{p}\left[\int_t^\infty Y_z e^{-\alpha(z-t)} R(t,z)\, dz\right] e^{-\alpha(t-s)}\right) ds \\ &= (\alpha+p)\int_{-\infty}^{t} e^{-(\alpha+p)(t-s)}\left[\int_t^\infty Y_z e^{-\alpha(z-t)} R(t,z)\, dz\right] ds \\ &= (\alpha+p)\left[\int_t^\infty Y_z e^{-\alpha(z-t)} R(t,z)\, dz\right]\int_{-\infty}^{t} e^{-(\alpha+p)(t-s)} ds \\ &= (\alpha+p)\left[\int_t^\infty Y_z e^{-\alpha(z-t)} R(t,z)\, dz\right]\left[\frac{1}{\alpha+p} e^{-(\alpha+p)(t-s)}\right]_{-\infty}^{t} \\ &= \int_t^\infty Y_z e^{-\alpha(z-t)} R(t,z)\, dz \\ &= \int_t^\infty Y_z \exp\left[-\alpha(z-t)\right] \exp\left[-\int_t^z (r_\mu + p)\, d\mu\right] dz \\ &= \int_t^\infty Y_z \exp\left[-\int_t^z (r_\mu + \alpha + p)\, d\mu\right] dz. \end{aligned} \tag{4.50}$$

Hence, aggregate human wealth is the future aggregate income, discounted at a mark-up over the rate of interest equal to $\alpha + p$, for those who are currently alive.

To proceed, recall the rule for differentiation under the integral sign (Leibniz's Rule). Suppose we need to differentiate with respect to x the function:

$$G(y) = \int_{p(y)}^{q(y)} f(x,y)\, dx,$$

where appropriate conditions hold. Then, the derivative of G is given by

$$\frac{dG}{dy} = f(q(y), y)\frac{dq}{dy} - f(p(y), y)\frac{dp}{dy} + \int_{p(y)}^{q(y)} \frac{\partial f(x,y)}{\partial y} dx.$$

Using this rule, we can differentiate (4.50) with respect to t (replace y above with t and x with z). We also define

$$f(x,y) = f(z,t) = Y_z \exp\left[-\int_t^z (r_\mu + \alpha + p)\, d\mu\right],\ q(y) = \infty,\ p(y) = t.$$

It follows that

$$\frac{dq}{dt} = 0,\ \frac{dp}{dt} = 1$$

so that we only need to concern ourselves with

$$\begin{aligned} f(p(y), y) &= Y_t \exp\left[-\int_t^t (r_\mu + \alpha + p)\, d\mu\right] \\ &= Y_t. \end{aligned}$$

We therefore have that

$$\begin{aligned}
\frac{dH_t}{dt} &= -Y_t + \int_t^{\infty} \frac{\partial}{\partial t}\left(Y_z \exp\left[-\int_t^z \left(r_\mu + \alpha + p\right) d\mu\right]\right) dz \\
&= -Y_t + \int_t^{\infty} Y_z \frac{\partial}{\partial t}\left(\exp\left[-\int_t^z \left(r_\mu + \alpha + p\right) d\mu\right]\right) dz
\end{aligned}$$

Consider

$$\begin{aligned}
\frac{\partial}{\partial t}\left(-\int_t^z \left(r_\mu + \alpha + p\right) d\mu\right) &= \left(r_t + \alpha + p\right) - \int_t^z 0 d\mu \\
&= r_t + \alpha + p.
\end{aligned}$$

Using this, we can now write

$$\begin{aligned}
\frac{dH_t}{dt} &= -Y_t + \left(r_t + \alpha + p\right) \underbrace{\int_t^{\infty} Y_z \exp\left[-\int_t^z \left(r_\mu + \alpha + p\right) d\mu\right]}_{H_t} \\
&= \left(r_t + \alpha + p\right) H_t - Y_t.
\end{aligned}$$

4.7.7 Derivation of law of motion for Aggregate Non-human Wealth

In principle, we can use the same method as what we did to derive the aggregate human wealth to drive the law of motion for aggregate non-human wealth. However, there is an easier way, which is to to use the individual's budget constraint, (4.2). We could not do this for human wealth as we did not have a similar expression for dh_t/dt.

Recall that

$$\begin{aligned}
V_t &:= \int_{-\infty}^t N\left(s,t\right) v\left(s,t\right) ds \\
&= \int_{-\infty}^t p e^{-p(t-s)} v\left(s,t\right) ds.
\end{aligned}$$

Then,

$$\begin{aligned}
\frac{dV_t}{dt} &= \frac{d}{dt}\left(\int_{-\infty}^t p e^{-p(t-s)} v\left(s,t\right) ds\right) \\
&= \left[\left(p e^{-p(t-s)}\right) v\left(s,t\right) | s = t\right] + \int_{-\infty}^t \frac{\partial}{\partial t}\left(p e^{-p(t-s)} v\left(s,t\right)\right) ds \\
&= \underbrace{p v\left(t,t\right)}_{=0} + \int_{-\infty}^t \frac{\partial}{\partial t}\left(p e^{-p(t-s)} v\left(s,t\right)\right) ds \\
&= \int_{-\infty}^t \left(p e^{-p(t-s)} \frac{dv\left(s,t\right)}{dt} - v\left(s,t\right) p e^{-p(t-s)}\right) ds \\
&= -p \underbrace{\int_{-\infty}^t v\left(s,t\right) e^{-p(t-s)} ds}_{=V_t} + p \int_{-\infty}^t e^{-p(t-s)} \frac{dv\left(s,t\right)}{dt} ds \\
&= -p V_t + p \int_{-\infty}^t e^{-p(t-s)} \frac{dv\left(s,t\right)}{dt} ds.
\end{aligned}$$

Recall that the individual's budget constraint is

$$\frac{dv\left(t\right)}{dt} = \left(r\left(t\right) + p\right)v\left(t\right) + y\left(t\right) - c\left(t\right).$$

Hence,

$$\begin{aligned}
\frac{dV_t}{dt} &= -pV_t + p\int_{-\infty}^{t} e^{-p(t-s)}\left[\left(r\left(t\right) + p\right)v\left(t\right) + y\left(t\right) - c\left(t\right)\right]ds \\
&= -pV_t + \left(r\left(t\right) + p\right)\underbrace{\int_{-\infty}^{t} v\left(s,t\right)pe^{-p(t-s)}ds}_{=V_t} \\
&\qquad \underbrace{\int_{-\infty}^{t} y\left(s,t\right)pe^{-p(t-s)}ds}_{=Y_t} - \underbrace{\int_{-\infty}^{t} c\left(s,t\right)pe^{-p(t-s)}ds}_{=C_t} \\
&= -pV_t + \left(r_t + p\right)V_t + Y_t - C_t \\
&= r_tV_t + Y_t - C_t.
\end{aligned}$$

4.7.8 Derivation of law for motion of Aggregate Consumption

Recall that (see (4.13), and sections 4.7.6 and 4.7.7):

$$\begin{aligned}
C_t &= \left(p + \theta\right)\left(H_t + V_t\right), \\
\frac{dH_t}{dt} &= \left(r_t + \alpha + p\right)H_t - Y_t, \\
\frac{dV_t}{dt} &= r_tV_t + Y_t - C_t.
\end{aligned}$$

Totally differentiating the expression for C_t and substituting the expressions for dH_t/dt and dV_t/dt yields:

$$\begin{aligned}
\frac{dC_t}{dt} &= \left(p + \theta\right)\left(\frac{dH_t}{dt} + \frac{dV_t}{dt}\right) \\
&= \left(p + \theta\right)\left(\left[\left(r_t + p + \alpha\right)H_t - Y_t\right] + \left[r_tV_t + Y_t - C_t\right]\right) \\
&= \left(p + \theta\right)\left(\left(r_t + p + \alpha\right)H_t + r_tV_t - C_t\right).
\end{aligned}$$

Substituting out H_t using the expression for C_t gives

$$\begin{aligned}
\frac{dC_t}{dt} &= \left(p + \theta\right)\left(\left(r_t + p + \alpha\right)\left(\frac{C_t}{p + \theta} - V_t\right) + r_tV_t - C_t\right) \\
&= \left(p + \theta\right)\left(C_t\left(\frac{r_t + p + \alpha}{p + \theta} - 1\right) + \left(p + \alpha\right)V_t\right) \\
&= \left(r_t + \alpha - \theta\right)C_t - \left(p + \alpha\right)\left(p + \theta\right)V_t.
\end{aligned}$$

4.7.9 Deriving the law of motion for capital with government

With government, the individual's budget constraint is given by

$$\frac{dv_z}{dz} = \left(r_z + p\right)a_z + y_z - c_z - \tau_z,$$

where τ_z is the individual lump sum tax in period z. We can follow the same derivation in section 4.7.7 to obtain that

$$\frac{dV_t}{dt} = r_t V_t + Y_t - C_t - T_t,$$

where T_z is the aggregate lump sump tax given by

$$\int_{-\infty}^{t} \tau(s,t)\, p e^{-p(t-s)} ds.$$

Recall that, without government, the total output was equal to the sum of human wealth and the return on non-human wealth (see (4.22)). With government, human wealth is reduced by an amount equal to the tax: i.e.

$$\begin{aligned} F(K_t) &= Y_t - T_t + rV_t \\ &= Y_t - T_t + r(K_t - B_t), \end{aligned}$$

where we used the fact that $V(t) = K(t) - B(t)$ (see (4.34)) in the last equality. Notice that

$$\frac{dK_t}{dt} = \frac{\partial K_t}{\partial V_t}\frac{dV_t}{dt} = \frac{dV_t}{dt} = r_t V_t + Y_t - C_t - T_t.$$

Substituting out Y_t yields that

$$\frac{dK_t}{dt} = r_t V_t + [F(K_t) + T_t - r(K_t - B_t)] - C_t - T_t$$

4.7.10 Comparative statics on B^* and G

Recall that we derived the following equations that describe the steady-state:

$$\begin{aligned} (F'(K^*) - \theta) C^* &= p(p+\theta)(K^* + B^*) \qquad (4.51) \\ F(K^*) &= C^* + G, \\ F'(K^*) B^* &= T - G. \end{aligned}$$

Replacing C^* in the first equation using the second, we obtain

$$(F'(K^*) - \theta)(F(K^*) - G) = p(p+\theta)(K^* + B^*)$$

Differentiating with respect to B^* yields

$$\begin{aligned} F''(K^*) \frac{dK^*}{dB^*}(F(K^*) - G) + (F'(K^*) - \theta) F'(K^*) \frac{dK^*}{dB^*} &= p(p+\theta)\left(\frac{dK^*}{dB^*} + 1\right) \\ \Rightarrow [F''(K^*) C^* + r^*(r^* - \theta)] \frac{dK^*}{dB^*} &= p(p+\theta)\left(\frac{dK^*}{dB^*} + 1\right), \end{aligned}$$

so that

$$\frac{dK^*}{dB^*} = \frac{p(p+\theta)}{F''(K^*) C^* + r^*(r^* - \theta) - p(p+\theta)}.$$

Substituting out $p\left(p+\theta\right)$ using (4.51) gives

$$\frac{dK^*}{dB^*}=\frac{p\left(p+\theta\right)}{F''\left(K^*\right)C^*+r^*\left(r^*-\theta\right)-\frac{\left(r-\theta\right)C^*}{K^*+B^*}}.$$

Using the fact that $C^*=r^*\left(B+K\right)+Y-T$, and multiplying this by $\left(r-\theta\right)/K+B$ yields that ???

$$\frac{\left(r-\theta\right)C^*}{K^*+B^*}=r\left(r^*-\theta\right)+\frac{Y-T}{K^*+B^*}\left(r-\theta\right)$$

so that

$$\frac{dK^*}{dB^*}=\frac{p\left(p+\theta\right)}{F''\left(K^*\right)C^*-\frac{Y-T}{K^*+B^*}\left(r^*-\theta\right)}.$$

Hence, if $Y<T$ and $K^*+B^*>0$, then $dK^*/dB^*>0$.

Similarly, we can differentiate (4.51) with respect to G to obtain that

$$\begin{aligned}F''\left(K^*\right)\frac{dK^*}{dG^*}\left(F\left(K^*\right)-G\right)+\left(F'\left(K^*\right)-\theta\right)\left(F'\left(K^*\right)\frac{dK^*}{dG^*}-1\right)&=p\left(p+\theta\right)\frac{dK^*}{dG^*}\\ \Rightarrow F''\left(K^*\right)C^*\frac{dK^*}{dG^*}+\left(r-\theta\right)\left(r\frac{dK^*}{dG^*}-1\right)&=p\left(p+\theta\right)\frac{dK^*}{dG^*},\\ \left(F''\left(K^*\right)C^*+\left(r-\theta\right)-p\left(p+\theta\right)\right)\frac{dK^*}{dG^*}&=r-\theta,\end{aligned}$$

which gives that

$$\begin{aligned}\frac{dK^*}{dG^*}&=\frac{r-\theta}{F''\left(K^*\right)C^*+\left(r-\theta\right)-p\left(p+\theta\right)}\\ &=\frac{r-\theta}{F''\left(K^*\right)C^*-\frac{Y-T}{K^*+B^*}\left(r^*-\theta\right)}.\end{aligned}$$

Hence, if $Y<T$ and $K^*+B^*>0$, then $dK^*/dB^*>0$.

Part III

Uncertainty

5 Uncertainty

By carefully choosing the commodity space, set of agents, etc. several interesting economic issues can be analysed as competitive equilibrium. In this section, we analyse how uncertainty affects the extent of risk sharing and the price of risky securities.

Assume that the state of the economy s can take m_1 different values. To simplify, we will assume that there are m_2 physically different goods, indexed by r. We will index a commodity both by its physical attributes as well as by the state, so that there are $m = m_1 \times m_2$ goods, or $L = \mathbb{R}^m$. Thus, the interpretation of vector $\mathbf{x}$ is the consumption of each of the physically different goods in each of the states. We write x_{sr} for the good in state s of physical characteristic r.

Utility function u^i is a function of a vector on $\mathbb{R}^m$. The endowment of agent i, denoted by $\mathbf{e}^i$ are also indexed by state s and physical characteristic r.

The production possibilities of the economy can also depend on the state, thus, the production possibility sets of each firm are $Y^j \subseteq \mathbb{R}^m$.

5.1 Risk aversion and Jensen's Inequality

We assume that agents have expected utility: i.e.

$$u^i(\mathbf{x}) = \sum_{s=1}^{m_1} v^i(x_{s1}, x_{s2}, \ldots, x_{sm_2})\,\pi_s^i,$$

for some function $v^i : \mathbb{R}^{m_2} \to \mathbb{R}$ and some vector of (subjective) probabilities $\boldsymbol{\pi}^i \in \mathbb{R}_+^{m_1}$, with $\sum_{s=1}^{m_1} \pi_s^i = 1$. The subutility function v^i has the interpretation of the utility that an agent will enjoy if he consumes the m_2 physically different goods in state s. In this case, it is convenient to regard the vector $x_{s1}, x_{s2}, \ldots, x_{sm_2}$ as a random variable with m_1 possible realisations.

To analyse attitudes toward risk, we simplify and consider the case of only one physically different good so that $m_1 = m$; i.e.

$$u^i(\mathbf{x}) = \sum_{s=1}^{m} v^i(x_s)\,\pi_s^i.$$

Definition 5.1. We say that u^i is *risk averse* if

$$v^i\left(\sum_{s=1}^{m} x_s \pi_s^i\right) > \sum_{s=1}^{m} v^i(x_s)\,\pi_s^i \tag{5.1}$$

for any random variable x.

Notice that, if $m = 2$, this coincides with the definition of v^i being strictly concave and the assumption that x, as a random variable, is not degenerate.

Exercise 5.1. Show that if v^i is concave, then u^i is concave.

Solution. If v^i is concave, then for any $x_s, y_s \in \mathbb{R}$ and $\alpha \in [0,1]$:

$$v^i\left(\lambda x_s + (1-\lambda) y_s\right) \geq (1-\lambda) v^i(x_s) + \lambda v^i(y_s).$$

Substituting this expression into the definition of u^i yields

$$\begin{aligned}\sum_{s=1}^{m} v^i\left(\lambda x_s + (1-\lambda) y_s\right) \pi_s^i &\geq \sum_{s=1}^{m} \left[(1-\lambda) v^i(x_s) + \lambda v^i(y_s)\right] \pi_s^i \\ u^i\left(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\right) &\geq (1-\lambda)\sum_{s=1}^{m} v^i(x_s)\pi_s^i + \lambda \sum_{s=1}^{m} v^i(y_s)\pi_s^i \\ &= (1-\lambda) u^i(\mathbf{x}) + \lambda u^i(\mathbf{y}),\end{aligned}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_m)$. Hence, u^i is concave.

To understand how concavity implies the Jensen's inequality, (5.1), we will consider the case where v^i is concave and where the random variable can take 3 values so that $m = 3$.

Claim 5.1. For any positive weights η_s,

$$v^i\left(\sum_{s=1}^{3} x_s \eta_s\right) > \sum_{s=1}^{3} v^i(x_s)\eta_s.$$

Proof. We can write the expected utility as

$$\sum_{s=1}^{3} v^i(x_s)\eta_s = \left(\sum_{s'=1}^{2} \eta_{s'}\right)\left(\sum_{s=1}^{2} v^i(x_s)\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}}\right) + v^i(x_3)\eta_3.$$

We can also write

$$\sum_{s=1}^{3} x_s\eta_s = \left(\sum_{s=2}^{2} \eta_{s'}\right)\left(\sum_{s=1}^{2} x_s\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}}\right) + x_3\eta_3.$$

The term $\eta_s / \left(\sum_{s'=1}^{2}\eta_{s'}\right)$ has the interpretation of conditional probability of s being 1 or 2. Using the definition of concavity:

$$\sum_{s=1}^{2} v^i(x_s)\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}} < v\left(\sum_{s=1}^{2} v^i(x_s)\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}}\right).$$

Multiplying this by $\sum_{s'=1}^{2}\eta_{s'}$ and adding $v\left(x_3\right)\eta_3$, we get

$$\begin{aligned}
\sum_{s=1}^{3} v^i\left(x_s\right)\eta_s &= \left(\sum_{s=1}^{2} v^i\left(x_s\right)\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}}\right)\left(\sum_{s'=1}^{2}\eta_{s'}\right) + v^i\left(x_3\right)\eta_3 \\
\text{(by concavity)} &< v^i\left(\underbrace{\sum_{s=1}^{2} v^i\left(x_s\right)\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}}}_{=\tilde{x}_1}\right)\left(\underbrace{\sum_{s'=1}^{2}\eta_{s'}}_{=\tilde{\eta}_1}\right) + v^i\left(x_3\right)\eta_3 \\
&= v^i\left(\tilde{x}_1\right)\tilde{\eta}_1 + v^i\left(x_3\right)\eta_3 \\
\text{(by concavity)} &< v^i\left(\tilde{x}_1\tilde{\eta}_1 + x_3\eta_3\right) \\
&< v^i\left(\left(\sum_{s=1}^{2} v^i\left(x_s\right)\frac{\eta_s}{\sum_{s'=1}^{2}\eta_{s'}}\right)\left(\sum_{s'=1}^{2}\eta_{s'}\right) + x_3\eta_3\right) \\
&= v^i\left(\sum_{s=1}^{3} x_s\eta_s\right).
\end{aligned}$$

■

Proof for when $n > 3$ can then be done by induction.

5.2 Equilibrium Risk Sharing

We will study the risk sharing implications for equilibrium allocations and Pareto optimal allocations in the context of a pure endowment economy with one good, m states of nature, and preferences given by expected utility displaying risk aversion and all agents using the *same* probabilities $\boldsymbol{\pi}$. Under these assumptions, we will show that the consumption of each agent depends solely on the realisation of the aggregate endowment, and that all the individual consumptions move together with aggregate endowment.

5.2.1 Risk sharing

Let

$$\bar{\mathbf{e}} = \sum_{i\in I}\mathbf{e}^i \in \mathbb{R}^m$$

be the aggregate endowment where $\mathbf{e}^i \in \mathbb{R}^m$ is the endowment of agent i, and let $\pi_s^i = \pi_s \in \mathbb{R}_+$ for all i be the common probability of states s, and $v^i : \mathbb{R} \to \mathbb{R}$ be the subutility function of agent i. We assume that v^i's are differentiable, strictly increasing and strictly concave. We can interpret the vectors $\mathbf{x}^i$, $\mathbf{e}^i$ and $\bar{\mathbf{e}}$ as random variables.

Theorem 5.1. *Fix an arbitrary vector of weights, λ. The corresponding Pareto optimal allocation can be described by a set of strictly increasing functions g^i of the aggregate endowment; i.e. the optimal allocation can be written as*

$$x_s^i = g^i\left(\bar{e}_s\right),\ \forall i \in I.$$

Proof. Using the assumption of expected utility and that probabilities are common across all agents, we can write the objective function of the Pareto optimal problem as

$$\begin{aligned}\sum_{i\in I}\lambda_i u^i\left(\mathbf{x}^i\right) &= \sum_{i\in I}\lambda_i\left(\sum_{s=1}^{m} v^i\left(x_s^i\right)\pi_s^i\right)\\ \text{(additive separability)} &= \sum_{s=1}^{m}\left(\sum_{i\in I}\lambda_i v^i\left(x_s^i\right)\pi_s^i\right)\\ \text{(common beliefs)} &= \sum_{s=1}^{m}\left(\sum_{i\in I}\lambda_i v^i\left(x_s^i\right)\right)\pi_s.\end{aligned}$$

Thus, the planner's problem is to maximise above subject to the feasibility constraint:

$$\begin{aligned}\max_{\{x_s^i\}_{s\in\{1,2,\ldots,m\}i\in I}} &\sum_{s=1}^{m}\left(\sum_{i\in I}\lambda_i v^i\left(x_s^i\right)\right)\pi_s\\ s.t. \quad &\sum_{i\in I} x_s^i = \bar{e}_s \;\forall s = 1,2,\ldots,m.\end{aligned}$$

The solution to the problem above can be obtained by solving the sub-problem for each state; i.e. for each state $s \in \{1, 2, \ldots, m\}$, solve

$$\begin{aligned}\max_{\{x_s^i\}_{i\in I}} &\sum_{i\in I}\lambda_i v^i\left(x_s^i\right)\\ s.t. \quad &\sum_{i\in I} x_s^i = \bar{e}_s.\end{aligned}$$

Notice that the probability π_s does not enter in the state-s problem. The only difference between the sub-problems for different state s is the aggregate endowment in each state, $\bar{e}_s$. This establishes that the solution of the state s problem is given by

$$x_s^i = g^i\left(\bar{e}_s\right).$$

Now we show that g^i functions are strictly increasing. Consider two states s and s' with $\bar{e}_s > \bar{e}_{s'}$. Then, it must be that for at least some i,

$$x_s^i > x_{s'}^i;$$

otherwise, the feasibility constraint will not hold with equality in state s. Clearly, g^i is increasing in $\bar{e}_s$. It remains to show that g^j for any other agent j is also increasing.

Proof. Using the first-order conditions for the sub-problem (equate the first-order conditions for agent i and j using the Lagrange multiplier on the feasibility constraint), we obtain:

$$\lambda_i\frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i} = \lambda_j\frac{\partial v^j\left(x_s^j\right)}{\partial x_s^j}, \tag{5.2}$$

$$\lambda_i\frac{\partial v^i\left(x_{s'}^i\right)}{\partial x_{s'}^i} = \lambda_j\frac{\partial v^j\left(x_{s'}^j\right)}{\partial x_{s'}^j}. \tag{5.3}$$

Since $x_s^i > x_{s'}^i$ and v^i is strictly concave (i.e. its derivative is decreasing):

$$\frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i} < \frac{\partial v^i\left(x_{s'}^i\right)}{\partial x_{s'}^i}.$$

From (5.2) and (5.2), this implies that

$$\frac{\partial v^j\left(x_s^j\right)}{\partial x_s^j} < \frac{\partial v^j\left(x_{s'}^j\right)}{\partial x_{s'}^j}.$$

Using the fact that v^j is strictly concave, above implies that

$$x_s^j > x_{s'}^j.$$

Hence, g^j is strictly increasing. ■

■

Remark 5.1. The fact that g^i does not depend on the distribution of $\boldsymbol{\lambda}$ is a feature of Gorman aggregation. In the Gorman aggregation problem sets, we saw a number of examples of functions that leads to g^i. Note also that if people have different $\boldsymbol{\pi}$'s then the theorem would not hold.

Remark 5.2. Rewrite the first-order condition from the previous problem as

$$\lambda_i v_i'(g^i(\bar{e})) = \gamma(\bar{e})$$
$$\sum_{i=1}^{I} g_i(\bar{e}) = \bar{e}$$

Differentiating each for both sides:

$$\lambda_i v_i''(g_i(\bar{e}))g_i'(\bar{e}) = \gamma'(\bar{e})$$
$$\sum_{i=1}^{I} g_i'(\bar{e}) = \bar{e}$$

Combining the two equations:

$$g_i'(\bar{e}) = \frac{\gamma'(\bar{e})}{\gamma(\bar{e})} \cdot \frac{v_i'(g_i(\bar{e}))}{v_i''(g_i(\bar{e}))}$$

Aggregating on each side:

$$1 = \frac{\gamma'(\bar{e})}{\gamma(\bar{e})} \cdot \sum_{i=1}^{I} \left(\frac{v_i'(g_i(\bar{e}))}{v_i''(g_i(\bar{e}))}\right)$$

Rearranging:

$$\frac{\gamma'(\bar{e})}{\gamma(\bar{e})} = \frac{1}{\sum_{i=1}^{I} \left(\frac{v_i'(g_i(\bar{e}))}{v_i''(g_i(\bar{e}))}\right)}$$

- ▷ The more risk-tolerant insures the less risk-tolerant, if there is a heterogeneity in risk tolerance across the population.
- ▷ A person's level of consumption not only depends on her curvature but also on others:

$$g_i'(\bar{e}) = \frac{\gamma'(\bar{e})}{\gamma(\bar{e})} \cdot \frac{v_i'(g_i(\bar{e}))}{v_i''(g_i(\bar{e}))} = \frac{1}{\sum_{i=1}^{I} \left(\frac{v_i'(g_i(\bar{e}))}{v_i''(g_i(\bar{e}))}\right)} \cdot \frac{v_i'(g_i(\bar{e}))}{v_i''(g_i(\bar{e}))}$$

- ▷ Suppose everyone has exponential utility. Then g_i' will not depend on $\bar{e}$ and will be a constant.
    - ▷ The optimal risk-sharing rule is affine.

5.2.2 Competitive equilibrium and risk sharing

Since the welfare theorems hold for this economy, the previous theorem implies that, in any competitive equilibrium, the consumption allocations depend only on the aggregate endowment $\bar{e}_s$ and *not* on the individual realisation of e_s^i. This is the sense in which there is complete risk sharing. Note that consumption of all agents in state s is higher if the realisation s of the aggregate endowment is higher.

5.2.3 Competitive equilibrium and state prices

In a competitive equilibrium, the budget constraint of agent i is

$$\sum_{s=1}^{m} p_s x_s^i = \sum_{s=1}^{m} p_s e_s^i,$$

where p_s's are referred to as *state prices* (or Arrow-Debreu price); i.e. the price of a security that pays one unit of the numeraire in state s and zero otherwise (sometimes called Arrow securities). Thus, agents can buy consumption contingent on the state, and they finance that by selling their endowment contingent on the state.

The agent's problem in the A-D economy is

$$\begin{aligned} \max_{\{x_s\}_{s=1}^m} \quad & u^i\left(x_1, x_2, \ldots, x_m\right) = \sum_{s=1}^{m} v^i\left(x_s\right)\pi_s \\ s.t. \quad & \sum_{s=1}^{m} p_s x_x^i = \sum_{s=1}^{m} p_s e_s^i. \end{aligned}$$

The first-order condition with respect to x_s is given by

$$\frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i}\pi_s - \mu_i p_s = 0,$$

where μ_i is agent i's Lagrange multiplier on the constraint.

For the planner, the problem is

$$\begin{aligned} \max_{\{x_s^i\}_{s\in\{1,2,\ldots,m\}\, i\in I}} \quad & \sum_{s=1}^{m}\left(\sum_{i\in I}\lambda_i v^i\left(x_s^i\right)\right)\pi_s \\ s.t. \quad & \sum_{i\in I} x_s^i = \bar{e}_s \;\forall s = 1, 2, \ldots, m. \end{aligned}$$

The first-order condition with respect to x_s^i is then

$$\lambda_i \frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i}\pi_s - \gamma_s = 0.$$

Letting $\gamma_s = p_s$, we obtain that

$$p_s = \frac{1}{\mu_i}\frac{\partial v^i\left(g^i\left(\bar{e}_s\right)\right)}{\partial x_s^i}\pi_s = \lambda_i\frac{\partial v^i\left(g^i\left(\bar{e}_s\right)\right)}{\partial x_s^i}\pi_s, \tag{5.4}$$

where we replaced $x_s^i = g^i\left(\bar{e}_s\right)$. The expression above means that state prices, p_s, reflect the probability that the state s is realised, π_s, as well as the scarcity of the aggregate endowment in state s, through the marginal utility of consumption. Notice that state prices are lower if the probability (π_s) is small or if the aggregate endowment in that state ($\bar{e}_s$) is large (i.e. when the goods are relatively plentiful in the state. Recall that v^i is strictly concave so that marginal utility is strictly decreasing).

The equilibrium state price, p_s, is decreasing in consumption—for example, suppose $v^i\left(x\right) = x^{1-\gamma}/1-\gamma$, then marginal utility is $x^{-\gamma}$, therefore marginal utility decreases by $\gamma\%$ when consumption increases by 1%. If one agent, say i, has linear utility, and the others do not, since the agent with linear utility does not mind variability in consumption across states (but others do). In this case, since (5.4) must hold for all agents, including i, who has a constant marginal utility, state prices do not change with consumption. This means that agent i must be providing insurance for all other agents.[18]

5.3 Security Markets

We will study an economy where trade does not occur in contingent markets but in securities markets, where each security pays different dividends in different states.

Let d_{ks} be the payoff of security k in state s and suppose that there are K such securities. Let q_k denote the price of the security k, h_k^i the quantity of security k that agent i purchases, and θ_k^i the endowment of this security by agent i. Implicitly, we assume that agents make a decision in one period but he does not consume anything in that period. Then, the state of the world is realised, and the agent consumes based on the decision in the previous period. In this way, we abstract ourselves from intertemporal issues so that we can focus on the issue of risk sharing. There is no "interest rate" in this setting!

In this case, there are two "budget" constraints for each agent i. The first constraint requires that the value of purchases is limited by the value of sales of the securities:

$$\sum_{k=1}^{K} h_k^i q_k = \sum_{k=1}^{K} \theta_k^i q_k. \tag{5.5}$$

One can think of this as the budget constraint for the first period. But as mentioned, since we do not allow the agents to consume in the first period, consumption does not appear here. The second constraint requires that each agent's consumption must be limited by his endowment plus his payoffs (in each state):

$$x_s^i = \sum_{k=1}^{K} h_k^i d_{ks} + \hat{e}_s^i,\ \forall s = 1, 2, \ldots m, \tag{5.6}$$

where $\hat{e}_s^i$ denotes the endowments of the good in state s for agent i.

The two constraints can be more succinctly written using vector-matrix notation:

$$\mathbf{q}^\intercal \mathbf{h}^i = \mathbf{q}^\intercal \boldsymbol{\theta}^i, \tag{5.7}$$

$$\mathbf{x}^i = D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i, \tag{5.8}$$

[18] See 2016/17 Final (Optimal Risk Sharing). We also cover this with Myerson in Price Theory III.

where

$$\mathbf{q} = \begin{pmatrix} q_1 & q_2 & \cdots & q_K \end{pmatrix}^\intercal,$$
$$\mathbf{h}^i = \begin{pmatrix} h_1^i & h_2^i & \cdots & h_K^i \end{pmatrix}^\intercal,$$
$$\boldsymbol{\theta}^i = \begin{pmatrix} \theta_1^i & \theta_2^i & \cdots & \theta_K^i \end{pmatrix}^\intercal,$$
$$\mathbf{x}^i = \begin{pmatrix} x_1^i & x_2^i & \cdots & x_m^i \end{pmatrix}^\intercal,$$
$$\hat{\mathbf{e}}^i = \begin{pmatrix} \hat{e}_1^i & \hat{e}_2^i & \cdots & \hat{e}_m^i \end{pmatrix}^\intercal,$$

and D is the payoff matrix of the K securities in the m states; i.e.

$$D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1m} \\ d_{21} & d_{22} & \cdots & d_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ d_{K1} & d_{K2} & \cdots & d_{Km} \end{pmatrix}_{K\times m}.$$

We will now compare the two market structures.

The first economy has Arrow-Debreu (A-D), or contingent markets, with A-D prices, or state/contingent prices as described in the previous section. The agent's budget constraint is:

$$\sum_{s=1}^{m} p_s x_s^i = \sum_{s=1}^{m} p_s e_s^i,$$

or, equivalently,

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{p}^\intercal \mathbf{e}^i, \tag{5.9}$$

where $\mathbf{p} = (p_1, p_2, \ldots, p_m)^\intercal$ and $\mathbf{e}^i = \left(e_1^i, e_2^i, \ldots, e_m^i\right)^\intercal$.

In the second economy, which we refer to as the security market economy, budget constraints are given by (5.5) and (5.6). There are two market clearing conditions for this economy: one for securities, and another for goods:

$$\sum_{i\in I} h_k^i = \sum_{i\in I} \theta_k^i,\ \ \forall k = 1, 2, \ldots, K,$$
$$\sum_{i\in I} x_s^i = \sum_{i\in I} \left[\hat{e}_s^i + \left(\sum_{k=1}^{K} d_{ks}\theta_k^i\right)\right],\ \ s = 1, 2, \ldots, m.$$

These can equivalently be written as:

$$\sum_{i\in I} \mathbf{h}^i = \sum_{i\in I} \boldsymbol{\theta}^i, \tag{5.10}$$
$$\sum_{i\in I} \mathbf{x}^i = \sum_{i\in I} \left[\hat{\mathbf{e}}^i + D^\intercal \boldsymbol{\theta}^i\right]. \tag{5.11}$$

We now establish the relationship between budget feasible allocation, $(\mathbf{x}, \mathbf{h})$ in the security market economy, $(\mathbf{x})$ in the A-D economy.

The next definition links the security prices and state prices.

Definition 5.2. *(Price consistency)* We say that security prices $\mathbf{q}$ and payoffs D are *consistent* with state price $\mathbf{p}$, if

$$\begin{aligned} q_k &= \sum_{s=1}^{m} p_s d_{ks},\ \forall k = 1, 2, \ldots, K. \\ \Leftrightarrow \mathbf{q} &= D\mathbf{p}. \end{aligned} \tag{5.12}$$

Definition 5.3. *(Endowment equivalence)* The endowment $\mathbf{e}^i$ and $\left(\hat{\mathbf{e}}^i, \boldsymbol{\theta}^i\right)$ are *equivalent* if

$$\begin{aligned} \hat{e}_s^i + \sum_{k=1}^{K} d_{ks}\theta_k^i &= e_s^i,\ \forall s = 1, 2, \ldots, m. \\ \Leftrightarrow \hat{\mathbf{e}}^i + D^\intercal \boldsymbol{\theta}^i &= \mathbf{e}^i. \end{aligned} \tag{5.13}$$

Proposition 5.1. *Assume that prices $\mathbf{q}$ and payoffs D are consistent with state prices $\mathbf{p}$ (i.e. (5.12) holds), and that the endowments are equivalent (i.e. (5.13) holds).*

(i) If $(\mathbf{x}, \mathbf{h})$ is budget feasible in the security market economy, then $\mathbf{x}$ is budget feasible in the A-D economy; i.e.

$$\mathbf{x}^i = D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i,\ \mathbf{q}^\intercal \mathbf{h}^i = \mathbf{q}^\intercal \boldsymbol{\theta}^i \Rightarrow \mathbf{p}^\intercal \mathbf{x}^i = \mathbf{p}^\intercal \mathbf{e}^i$$

(ii) If $\mathbf{x}$ is budget feasible in the A-D economy, then it must be budget feasible in the security market economy, provided that D has full rank; i.e.

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{p}^\intercal \mathbf{e}^i,\ D \textit{ has full rank} \Rightarrow \exists \mathbf{h}^i :\ \mathbf{x}^i = D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i,\ \mathbf{q}^\intercal \mathbf{h}^i = \mathbf{q}^\intercal \boldsymbol{\theta}^i.$$

Price consistency means that the price of each security k is given by a linear combination of Arrow-Debreu prices according to the dividends that the security pays in each state. Endowment equivalency relates the goods endowment $\mathbf{e}^i$ in the A-D economy with the goods endowments $\hat{\mathbf{e}}^i$ and securities endowments $\boldsymbol{\theta}^i$ in the security market economy.

Proof. (*Proposition 5.1*) To show (i). Suppose that $(\mathbf{x}, \mathbf{h})$ are budget feasible in the security market economy; repeating (5.7) and (5.8) again,

$$\begin{aligned} \mathbf{x}^i &= D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i, \\ \mathbf{q}^\intercal \mathbf{h}^i &= \mathbf{q}^\intercal \boldsymbol{\theta}^i. \end{aligned}$$

Multiplying both sides of the first expression by $\mathbf{p}^\intercal$ yields

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{p}^\intercal D^\intercal \mathbf{h}^i + \mathbf{p}^\intercal \hat{\mathbf{e}}^i.$$

Using the fact that endowments are equivalent, (5.13), we may replace $\hat{\mathbf{e}}^i$ to obtain

$$\begin{aligned} \mathbf{p}^\intercal \mathbf{x}^i &= \mathbf{p}^\intercal D^\intercal \mathbf{h}^i + \mathbf{p}^\intercal \left(\mathbf{e}^i - D^\intercal \boldsymbol{\theta}^i\right) \\ &= \mathbf{p}^\intercal D^\intercal \left(\mathbf{h}^i - \boldsymbol{\theta}^i\right) + \mathbf{p}^\intercal \mathbf{e}^i. \end{aligned}$$

Using the fact that $\mathbf{p}$ and $\mathbf{q}$ are consistent, (5.12), which implies that $\mathbf{q}^\intercal = (D\mathbf{p})^\intercal = \mathbf{p}^\intercal D^\intercal$:

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{q}^\intercal \left(\mathbf{h}^i - \boldsymbol{\theta}^i\right) + \mathbf{p}^\intercal \mathbf{e}^i.$$

Finally, using the second budget feasibility condition, we know that $\mathbf{q}^\intercal \left(\mathbf{h}^i - \boldsymbol{\theta}^i\right) = 0$ so that

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{p}^\intercal \mathbf{e}^i.$$

To show (ii). Notice that, since D has full rank, D is invertible, and $D^\intercal$ is also invertible. Hence, we can write

$$\begin{aligned}
\mathbf{x}^i &= \mathbf{x}^i - \hat{\mathbf{e}}^i + \hat{\mathbf{e}}^i \\
&= D^\intercal \underbrace{(D^\intercal)^{-1}\left(\mathbf{x}^i - \hat{\mathbf{e}}^i\right)}_{:=\mathbf{h}^i} + \hat{\mathbf{e}}^i \\
&= D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i.
\end{aligned}$$

We now claim that such $\mathbf{h}^i$ is affordable, which would be the case if we can show that $\mathbf{q}^\intercal \mathbf{h}^i = \mathbf{q}^\intercal \boldsymbol{\theta}^i$. Since $\mathbf{x}^i$ is budget feasible, it must be that (see (5.9)),

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{p}^\intercal \mathbf{e}^i.$$

Using the fact that $\mathbf{p}$ and $\mathbf{q}$ are consistent, (5.12), and the fact that D is invertible, we can write

$$\mathbf{p} = D^{-1}\mathbf{q} \Rightarrow \mathbf{p}^\intercal = \mathbf{q}^\intercal \left(\mathbf{D}^{-1}\right)^\intercal$$

Thus,

$$\mathbf{p}^\intercal \mathbf{x}^i = \mathbf{q}^\intercal \left(\mathbf{D}^{-1}\right)^\intercal \mathbf{x}^i = \mathbf{q}^\intercal \left(\mathbf{D}^{-1}\right)^\intercal \mathbf{e}^i = \mathbf{p}^\intercal \mathbf{e}^i.$$

Using the fact that endowments are equivalent, (5.13), we may replace $\mathbf{e}^i$ to obtain:

$$\begin{aligned}
\mathbf{q}^\intercal \left(\mathbf{D}^{-1}\right)^\intercal \mathbf{x}^i &= \mathbf{q}^\intercal \left(\mathbf{D}^{-1}\right)^\intercal \left(\hat{\mathbf{e}}^i + D^\intercal \boldsymbol{\theta}^i\right) \\
\Leftrightarrow \mathbf{q}^\intercal \left(\mathbf{D}^\intercal\right)^{-1} \mathbf{x}^i &= \mathbf{q}^\intercal \left(\mathbf{D}^\intercal\right)^{-1} \left(\hat{\mathbf{e}}^i + D^\intercal \boldsymbol{\theta}^i\right) \\
&= \mathbf{q}^\intercal \left(\mathbf{D}^\intercal\right)^{-1} \hat{\mathbf{e}}^i + \mathbf{q}^\intercal \boldsymbol{\theta}^i.
\end{aligned}$$

Since $\mathbf{x}^i = D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i$,

$$\begin{aligned}
\mathbf{q}^\intercal \left(\mathbf{D}^{-1}\right)^\intercal \left(D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i\right) &= \mathbf{q}^\intercal \left(\mathbf{D}^\intercal\right)^{-1} \hat{\mathbf{e}}^i + \mathbf{q}^\intercal \boldsymbol{\theta}^i \\
\Leftrightarrow \mathbf{q}^\intercal \mathbf{h}^i + \mathbf{q}^\intercal \left(\mathbf{D}^\intercal\right)^{-1} \hat{\mathbf{e}}^i &= \mathbf{q}^\intercal \left(\mathbf{D}^\intercal\right)^{-1} \hat{\mathbf{e}}^i + \mathbf{q}^\intercal \boldsymbol{\theta}^i \\
\Leftrightarrow \mathbf{q}^\intercal \mathbf{h}^i &= \mathbf{q}^\intercal \boldsymbol{\theta}^i.
\end{aligned}$$

■

Definition 5.4. *(Incomplete and complete markets)* If the payoff matrix does not have full rank, we refer to the economy as one with *incomplete markets*. If D has full rank, we refer to the economy as having *complete markets*.

Exercise 5.2. Exercises:

▷ What is the minimum number of securities needed to have complete markets?

The markets are complete if and only if D, which is a $K \times m$ matrix, has full rank. So the minimum number of securities is $K = m$.

▷ How are the budget sets in in the AD and security market economies with prices related by (5.12) if markets are incomplete? Which is a larger set?

Since $\mathbf{x}$ that is budget feasible in A-D economy is budget feasible in the security market economy if and only if D has full rank, but that $(\mathbf{x}, \mathbf{h})$ that is budget feasible in the security market economy is always budget feasible in the A-D economy, it must be that the budget set in the A-D economy is larger.

The following establishes the relationship between equilibrium in the security market with that of the A-D economy.

Proposition 5.2. *Assume that the endowments $\mathbf{e}^i$ and $\left(\hat{\mathbf{e}}^i, \boldsymbol{\theta}^i\right)$ are equivalent.*

(i) If $\left(\mathbf{x}^i, \mathbf{h}^i\right)$ clears the markets in the security markets in the security market economy, then $\mathbf{x}^i$ clears the markets in the A-D economy.

(ii) Assume also that D has full rank. If $\mathbf{x}^i$ clears the market in the A-D economy, then $\left(\mathbf{x}^i, \mathbf{h}^i\right)$ clears the market in the security market economy.

Proof. (Proposition 5.2) To show (i). Suppose that $\left(\mathbf{x}^i, \mathbf{h}^i\right)$ clears the market, which means (see (5.10)):

$$\sum_{i \in I} \mathbf{x}^i = \sum_{i \in I} \left(\hat{\mathbf{e}}^i + D^\intercal \boldsymbol{\theta}^i\right).$$

Since endowments are equivalent, (5.13), then

$$\sum_{i \in I} \mathbf{x}^i = \sum_{i \in I} \left(\hat{\mathbf{e}}^i + D^\intercal \boldsymbol{\theta}^i\right) = \sum_{i \in I} \mathbf{e}^i.$$

Hence, the goods market clears in the A-D economy.

To show (ii). We go through the steps in reverse to show that the goods market clears in the security market economy. It remains to show that the security market clears, i.e. (5.10). Since $\mathbf{h}^i$ must satisfy the budget constraint:

$$\mathbf{x}^i = D^\intercal \mathbf{h}^i + \hat{\mathbf{e}}^i,$$

where we already showed that

$$\mathbf{h}^i = \left(D^\intercal\right)^{-1} \left(\mathbf{x}^i - \hat{\mathbf{e}}^i\right).$$

Now, use the fact that endowments are equivalent, (5.13), to replace $\hat{\mathbf{e}}^i$:

$$\begin{aligned}
\mathbf{h}^i &= \left(D^\intercal\right)^{-1} \left(\mathbf{x}^i - \left(\mathbf{e}^i - D^\intercal \boldsymbol{\theta}^i\right)\right) \\
&= \left(D^\intercal\right)^{-1} \left(\mathbf{x}^i - \mathbf{e}^i\right) + \boldsymbol{\theta}^i.
\end{aligned}$$

Summing across consumers gives that

$$\sum_{i \in I} \mathbf{h}^i = \left(D^\intercal\right)^{-1} \sum_{i \in I} \left(\mathbf{x}^i - \mathbf{e}^i\right) + \sum_{i \in I} \boldsymbol{\theta}^i.$$

Since $\mathbf{x}^i$ clears the market in the A-D economy, $\sum_{i \in I} \left(\mathbf{x}^i - \mathbf{e}^i\right) = 0$ so that:

$$\sum_{i \in I} \mathbf{h}^i = \sum_{i \in I} \boldsymbol{\theta}^i. \qquad \blacksquare$$

Recall that state prices are given by, (5.4),

$$p_s = \frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i}\frac{\pi_s}{\mu_i}$$

Thus, if $(\mathbf{q}, \mathbf{D})$ and $\mathbf{p}$ are consistent, we can write the price of any security k, q_k, as

$$q_k = \sum_{s=1}^{m} p_s d_{ks} = \sum_{s=1}^{m} \frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i}\frac{\pi_s}{\mu_i} d_{ks}. \tag{5.14}$$

Exercise 5.3. Exercises:

▷ Are the equilibrium allocations of a security market economy with complete market necessarily Pareto optimal?

Yes.

▷ Can the equilibrium allocations of a security economy be Pareto optimal if the security markets are not complete?

The endowments could correspond to a Pareto optimal allocation (e.g. the initial allocation is Pareto optimal). But this does not hold generally.

5.4 The Tilde Economy

Recall that in the A-D economy, agent's utility was given by

$$u^i\left(x_1^i, x_2^i, \ldots, x_m^i\right) = \sum_{s=1}^{m} v^i\left(x_s^i\right)\pi_s, \tag{5.15}$$

so that utility depended upon the level of consumption in each state s, x_s. The agent maximised above subject to

$$\sum_{s=1}^{m} p_s x_s^i = \sum_{s=1}^{m} p_s e_s^i.$$

Consider instead the following analysis of the security market economy. We will use objects with tildes to denote the A-D economy that corresponds to the security market economy. In this economy, we define the utility as a function of the portfolio shares, so that $\tilde{L} = \mathbb{R}^K$. The utility is given by

$$\tilde{u}^i\left(\tilde{x}_1^i, \tilde{x}_2^i, \ldots, \tilde{x}_K^i\right) \coloneqq u^i\left(\underbrace{\sum_{k=1}^{K} d_{k1}\tilde{x}_k^i + e_1^i}_{=x_1^i}, \ldots, \underbrace{\sum_{k=1}^{K} d_{ks}\tilde{x}_k^i + e_s^i}_{=x_s^i}, \ldots, \underbrace{\sum_{k=1}^{K} d_{km}\tilde{x}_k^i + e_m^i}_{=x_m^i}\right),$$

where $\tilde{x}_k^i$ has the interpretation of the number of shares bought or sold of security k by agent i. Notice that x_s is the sum of the payoffs from holding of the securities by the individual plus the endowments in state s. The budget constraint is:

$$\sum_{k=1}^{K} \tilde{p}_k \tilde{x}_k^i = \sum_{k=1}^{K} \tilde{p}_k \tilde{e}_k^i,$$

where

$$\tilde{\mathbf{e}}^i = \boldsymbol{\theta}^i;$$

i.e. endowments are in terms of the shares of the securities. In this economy, feasible allocations are defined relative to the rearrangements of securities across agents.

Letting $\tilde{\mu}_i$ denote agent i's Lagrange multiplier on the budget constraint, the first-order condition with respect to x is

$$\sum_{s=1}^{m} \frac{\partial u^i}{\partial x_s} \frac{\partial x_s}{\partial \tilde{x}_k^i} - \tilde{\mu}_i \tilde{p}_k = 0.$$

Since $\partial x_s / \partial \tilde{x}_k^i = d_{ks}$ and given the form of u^i as in (5.15), above simplifies to

$$\tilde{p}_k = \frac{1}{\tilde{\mu}_i} \sum_{s=1}^{m} \frac{\partial u^i}{\partial x_s} d_{ks} = \sum_{s=1}^{m} \frac{\partial v^i}{\partial x_s^i} \frac{\pi_s}{\tilde{\mu}_i} d_{ks}.$$

Observe that this corresponds to (5.14). In other words, $\tilde{p}_k$ corresponds to q_k from above.

Exercise 5.4. Exercises.

▷ Show that if u^i is strictly increasing, then the First Welfare Theorem hold for this economy.

It suffices to show that $\tilde{u}^i$ is strictly increasing in $\tilde{x}_k^i$'s. An increase in $\tilde{x}_k^i$ must strictly increase at least some x_s^i's (if not, then dividends for security k is zero in all states). Since u^i is strictly increasing, then the desired result follows. Note that agents maximise $\tilde{u}^i \left(\tilde{x}_1^i, \tilde{x}_2^i, \ldots, \tilde{x}_K^i\right)$ subject to the budget constraint $\sum_{k=1}^{K} \tilde{p}_k \tilde{x}_k^i = \sum_{k=1}^{K} \tilde{p}_k \tilde{e}_k^i$—same set up as in the usual general equilibrium case so the First Welfare Theorem will hold.

▷ Show that if u^i are strictly quasiconcave, and $\tilde{e}^i > 0$, the second welfare theorem holds too.

We assume that $\tilde{x}_k$'s is an element of some convex set $\tilde{X}^i$. Then, it suffices to show that $\tilde{u}^i$ is strictly quasiconcave in $\tilde{x}_k^i$'s. That u^i is strictly quasiconcave means that, for any $\mathbf{x} \neq \mathbf{y}$,

$$u^i \left(t\mathbf{x} + (1-t)\mathbf{y}\right) > \min\left\{u^i(\mathbf{x}), u^i(\mathbf{y})\right\}, \forall t \in (0,1).$$

Note that

$$\begin{aligned} t\mathbf{x} + (1-t)\mathbf{y} &= t\left(\sum_{k=1}^{K} d_{k1}\tilde{x}_k^i + e_1^i, \ldots, \sum_{k=1}^{K} d_{km}\tilde{x}_k^i + e_m^i\right) \\ &\quad (1-t)\left(\sum_{k=1}^{K} d_{k1}\tilde{y}_k^i + e_1^i, \ldots, \sum_{k=1}^{K} d_{km}\tilde{y}_k^i + e_m^i\right) \\ &= \left(\sum_{k=1}^{K} d_{k1}\left[t\tilde{x}_k^i + (1-t)\tilde{y}_k^i\right] + e_1^i, \ldots, \sum_{k=1}^{K} d_{km}\left[t\tilde{x}_k^i + (1-t)\tilde{y}_k^i\right] + e_m^i\right) \end{aligned}$$

It follows that

$$\begin{aligned} \tilde{u}^i\left(t\tilde{\mathbf{x}}^i + (1-t)\tilde{\mathbf{y}}^i\right) = u^i\left(t\mathbf{x} + (1-t)\mathbf{y}\right) &> \min\left\{u^i(\mathbf{x}), u^i(\mathbf{y})\right\} \\ &= \min\left\{\tilde{u}^i(\tilde{\mathbf{x}}), \tilde{u}^i(\tilde{\mathbf{y}})\right\}. \end{aligned}$$

That is, $\tilde{u}^i$ is strictly quasiconcave. Thus, the Second Welfare Theorem holds.

▷ Assume that D does not have full rank. Show that there could be Pareto optimal allocations in the corresponding A-D economy that are not equilibrium in this economy (consider the case, for example, where there are: only one asset, two agents, no aggregate uncertainty, identical strictly concave utilities u^i, two states with equal probabilities and symmetric endowments). How is this possible in view of the previous results about incomplete markets? (Hint: are the set of feasible allocations in the tilde economy with K goods the same as in the original A-D economy with m goods?)

Let $K = 1$, $I = 2$, $m = 2$, $\bar{e} = \bar{e}_1 = \bar{e}_2$, $u \equiv u^i$ for $i = 1, 2$ is strictly concave, $\pi_1 = \pi_2 = 1/2$, and $e_s^1 = e_s^2 = \bar{e}_s/2$ for $s = 1, 2$. Since everything is symmetric, Pareto optimal allocation in the A-D economy is simply

$$x_s^i = \frac{\bar{e}}{2}, \ s = 1, 2,$$

D is a $K \times m$ matrix, which, in this case, is a 1×2 matrix. Since there are fewer securities than the number of states, markets are necessarily incomplete. In this case, utility function and the budget constraints in the tilde economy are

$$\begin{aligned} \tilde{u}^i\left(\tilde{x}^i\right) &= u^i\left(\underbrace{\left(d_1 + d_2\right)\tilde{x}^i + \frac{\bar{e}}{2}}_{=x_1^i}, \underbrace{\left(d_1 + d_2\right)\tilde{x}^i + \frac{\bar{e}}{2}}_{=x_2^i}\right), \\ \tilde{p}\tilde{x}^i &= \tilde{p}\tilde{e}^i = \tilde{p}\theta^i. \end{aligned}$$

From the budget constraint, it follows that $\tilde{x}^i = \theta^i$. ???

5.5 Asset Prices and the "Equity Premium"

Consider an economy with one good, m states, and complete markets. Assume that all u^i are given by expected utility with strictly concave v^i and that all agents have the same subjective probabilities π_s for all states. We are interested in understanding the price of two securities. Security $k = 1$ is a riskless bond, i.e.

$$d_{1s} = 1, \ \forall s = 1, 2, \ldots, m.$$

Security $k = 2$ is the *aggregate stock*, which is similar to a stock and pays

$$d_{2s} = \bar{e}_s, \ \forall s = 1, 2, \ldots, m;$$

i.e. it pays a dividend equal to the aggregate endowment. We are interested in the *risk premium*, defined as the difference between the expected return on the aggregate stock and the return on the riskless bond.

The expected (gross) return of any security k is denoted as r_k and defined as

$$1 + r_k := \frac{\sum_{s=1}^m d_{ks}\pi_s}{q_k} = \frac{\mathbb{E}\left[d_k\right]}{q_k};$$

i.e. it is given by the expected payoffs divided by its prices. We can define the risk premium as

$$RP_k := \frac{1 + r_k}{1 + r_1}.$$

Using Theorem 5.1, we can describe the equilibrium allocations by $x_s^i = g^i\left(\bar{e}_s\right)$ where g^i is a strictly increasing function.

Then, recall that a competitive equilibrium allocation satisfies the first-order condition, (5.4),

$$p_s = \frac{\partial v^i\left(x_s^i\right)}{\partial x_s^i}\frac{\pi_s}{\mu_i} = \frac{1}{\mu_i}\frac{\partial v^i\left(g_i\left(\bar{e}_s\right)\right)}{\partial x_s^i}\pi_s, \tag{5.16}$$

where μ_i the Lagrangian multiple on the budget constraint from the agent i's maximisation problem.

Let $\mathbf{q}$ and D be consistent with state prices $\mathbf{p}$ so that (see (5.12)):

$$q_k = \sum_{s=1}^{m} p_s d_{ks}.$$

In particular, given the payoff structure for securities $k=1$ and $k=2$:

$$q_1 = \sum_{s=1}^{m} p_s,\ q_2 = \sum_{s=1}^{m} p_s \bar{e}_s.$$

Combining these expressions with (5.16), and recalling that $\bar{e}_s$ and d_{ks} can be interpreted as random variables:

$$\begin{aligned}
q_1 &= \frac{1}{\mu_i}\sum_{s=1}^{m}\frac{\partial v^i\left(g_i\left(\bar{e}_s\right)\right)}{\partial x_s^i}\pi_s = \frac{1}{\mu_i}\mathbb{E}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i}\right],\\
q_2 &= \frac{1}{\mu_i}\sum_{s=1}^{m}\frac{\partial v^i\left(g_i\left(\bar{e}_s\right)\right)}{\partial x_s^i}\bar{e}_s\pi_s = \frac{1}{\mu_i}\mathbb{E}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i}\bar{e}\right],\\
q_k &= \frac{1}{\mu_i}\sum_{s=1}^{m}\frac{\partial v^i\left(g_i\left(\bar{e}_s\right)\right)}{\partial x_s^i}d_{ks}\pi_s = \frac{1}{\mu_i}\mathbb{E}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i}d_k\right],\ \forall k>2,
\end{aligned}$$

where q_k represents the price of some security $k>2$ that pay dividends according to $d_{k1},\ldots,d_{km}$ with no particular structure.

These imply that

$$\begin{aligned}
1+r_1 &= \mu_i\frac{1}{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}\right]} \because \mathbb{E}\left[d_1\right]=1\\
1+r_2 &= \mu_i\frac{\mathbb{E}\left[\bar{e}\right]}{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}\bar{e}\right]} \because \mathbb{E}\left[d_2\right]=\mathbb{E}\left[\bar{e}\right]\\
1+r_k &= \mu_i\frac{\mathbb{E}\left[d_k\right]}{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}d_k\right]} \because \mathbb{E}\left[d_2\right]=\mathbb{E}\left[d_k\right],
\end{aligned}$$

Thus, risk premium for a generic security k is given by

$$RP_k = \frac{1+r_k}{1+r_1} = \frac{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}\right]\mathbb{E}\left[d_k\right]}{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}d_k\right]}. \tag{5.17}$$

Recall that, for any random variable Y and Z:

$$\begin{aligned}\mathrm{Cov}\left(Y,Z\right) &= \mathbb{E}\left[\left(Y-\mathbb{E}\left[Y\right]\right)\left(Z-\mathbb{E}\left[Z\right]\right)\right]\\ &= \mathbb{E}\left[YZ\right]-\mathbb{E}\left[Y\right]\mathbb{E}\left[Z\right]\\ \Leftrightarrow \mathbb{E}\left[YZ\right] &= \mathbb{E}\left[Y\right]\mathbb{E}\left[Z\right]+\mathrm{Cov}\left(Y,Z\right).\end{aligned}$$

Hence,

$$\mathbb{E}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i}d_k\right]=\mathbb{E}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i}\right]\mathbb{E}\left[d_k\right]+\mathrm{Cov}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i},d_k\right].$$

We can then write RP_k as

$$\begin{aligned}RP_k=\frac{1+r_k}{1+r_1} &= \frac{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}\right]\mathbb{E}\left[d_k\right]}{\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}\right]\mathbb{E}\left[d_k\right]+\mathrm{Cov}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i},d_k\right]}\\ &= \frac{1}{1+\mathrm{Cov}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i},d_k\right]/\mathbb{E}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}\right]\mathbb{E}\left[d_k\right]}.\end{aligned}$$

This implies that

$$\begin{aligned}\frac{1+r_k}{1+r_1}\geq 1 &\Leftrightarrow r_k\geq r_1\\ &\Leftrightarrow \mathrm{Cov}\left[\frac{\partial v^i\left(g_i\left(\bar{e}\right)\right)}{\partial x^i},d_k\right]\leq 0\end{aligned}$$

since marginal utilities and expected dividends are strictly positive.

The inequality above means that the return on security k, r_k, is higher when the (expected) payoff/dividends from the security is negatively correlated with the marginal utility of the consumer. Recall that marginal utility is decreasing so that agents get a lower marginal utility when aggregate consumption is higher. Thus, the return on a security is higher if it pays out more in states when the aggregate endowment is lower. This is because agents are risk averse—agents prefer assets that can better insure them against the risk from aggregate endowment.

Lemma 5.1. *Let f be a positive function that is strictly decreasing in X. Assume that X has a strictly positive variance. Show that*

$$\mathrm{Cov}\left[X,f\left(X\right)\right]<0.$$

Proof. Recall that

$$\begin{aligned}\mathrm{Cov}\left(X,Y-b\right) &= \mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)\left(Y-b-\mathbb{E}\left[Y-b\right]\right)\right]\\ &= \mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)\left(Y-\mathbb{E}\left[Y\right]\right)\right]\\ &= \mathrm{Cov}\left(X,Y\right),\end{aligned}$$

where b is a constant. Thus, let $\tilde{f}(X) = f(X) - \mathbb{E}[f(X)]$. Then $\mathbb{E}\left[\tilde{f}(X)\right] = 0$ so that

$$\begin{aligned}
\text{Cov}[Xf(X)] &= \text{Cov}\left[X\tilde{f}(X)\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}[X])\tilde{f}(X)\right] \\
&= \int_{-\infty}^{\infty} (X - \mathbb{E}[X])\tilde{f}(X) F(dX) \\
&= \int_{-\infty}^{\mathbb{E}[X]} \underbrace{(X - \mathbb{E}[X])}_{<0}\tilde{f}(X) F(dX) \\
&\quad + \int_{\mathbb{E}[X]}^{\infty} \underbrace{(X - \mathbb{E}[X])}_{>0}\tilde{f}(X) F(dX).
\end{aligned}$$

Notice that since $\tilde{f}(X)$ is decreasing and it is formed by subtracting a constant from $f(X)$, then $\tilde{f}(X_1) > \tilde{f}(X_0)$ for any $X_1 > X_0$. Thus, the "weights" placed on the negative parts are greater than the weights placed on the positive. Hence, $\text{Cov}[Xf(X)] < 0$. ■

Notice that

$$\text{Cov}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}, \bar{e}\right] = \text{Cov}[f(\bar{e}), \bar{e}],$$

where $f(\cdot)$ is a positive and decreasing function (since it is the marginality utility). Thus, it is immediate by the Lemma that, if $\bar{e}$ has a strictly positive variance, then

$$\text{Cov}\left[\frac{\partial v^i(g_i(\bar{e}))}{\partial x^i}, \bar{e}\right] < 0 \Leftrightarrow \frac{1+r_2}{1+r_1} > 1.$$

That is, the risk premium on the aggregate stock is strictly positive.

5.5.1 Example

Now suppose that

$$v^i(x) := \begin{cases} \frac{x^{1-\gamma}}{1-\gamma} & \gamma \neq 1 \\ \ln(x) & \gamma = 1 \end{cases}.$$

The planner's problem is given by

$$\begin{aligned}
\max_{\{x_s^i\}} \quad & \sum_{i \in I} \lambda_i \left(\sum_{s=1}^{m} v^i\left(x_s^i\right)\pi_s\right) \\
s.t. \quad & \sum_{i \in I} x_s^i = \bar{e}_s, \ \forall s = 1, 2, \ldots, m.
\end{aligned}$$

The Lagrangian is then

$$\mathcal{L} = \sum_{i \in I} \lambda_i \left(\sum_{s=1}^{m} v^i\left(x_s^i\right)\pi_s\right) + \sum_{s=1}^{m} \mu_s \left(\bar{e}_s - \sum_{i \in I} x_s^i\right).$$

The first-order condition with respect to x_s^i is

$$\begin{aligned}\lambda_i \left(x_s^i\right)^{-\gamma} \pi_s &= \mu_s \\ \Rightarrow x_s^i &= \left(\frac{\mu_s}{\lambda_i \pi_s}\right)^{-\frac{1}{\gamma}} \\ &= \lambda_i^{\frac{1}{\gamma}} \left(\frac{\mu_s}{\pi_s}\right)^{-\frac{1}{\gamma}}.\end{aligned}$$

Summing across agents, we obtain that

$$\begin{aligned}\sum_{i\in I} x_s^i = \bar{e}_s &= \sum_{i\in I} \lambda_i^{\frac{1}{\gamma}} \left(\frac{\mu_s}{\pi_s}\right)^{-\frac{1}{\gamma}} \\ &= \left(\frac{\mu_s}{\pi_s}\right)^{-\frac{1}{\gamma}} \left(\sum_{i\in I} \lambda_i^{\frac{1}{\gamma}}\right) \\ \Rightarrow \left(\frac{\mu_s}{\pi_s}\right)^{-\frac{1}{\gamma}} &= \bar{e}_s \left(\sum_{i\in I} \lambda_i^{\frac{1}{\gamma}}\right)^{-1}.\end{aligned}$$

Substituting into x_s^i gives

$$x_s^i = \frac{\lambda_i^{\frac{1}{\gamma}}}{\sum_{i\in I} \lambda_i^{\frac{1}{\gamma}}} \bar{e}_s = \delta^i \bar{e}_s,$$

where we notice that $\delta^i > 0$ and $\sum_{i\in I} \delta^i = 1$. Thus, here, the function g^i is linear with no intercept.

Then,

$$\frac{\partial v^i \left(g_i \left(\bar{e}_s\right)\right)}{\partial x_s^i} = \left(x_s^i\right)^{-\gamma} = \left(\delta^i \bar{e}_s\right)^{-\gamma}.$$

Hence, using (5.17), we realise that

$$\begin{aligned}\frac{1+r_2}{1+r_1} &= \frac{\mathbb{E}\left[\left(\delta^i \bar{e}\right)^{-\gamma}\right] \mathbb{E}\left[\bar{e}\right]}{\mathbb{E}\left[\left(\delta^i \bar{e}\right)^{-\gamma} \bar{e}\right]} \\ &= \frac{\mathbb{E}\left[\bar{e}^{-\gamma}\right] \mathbb{E}\left[\bar{e}\right]}{\mathbb{E}\left[\bar{e}^{-\gamma} \bar{e}\right]}.\end{aligned}$$

Proposition 5.3. *Suppose X is log normally distributed. Specifically,* $\ln X \sim N\left(\mu, \sigma^2\right)$. *Then,*

$$\mathbb{E}\left[X^{-\gamma}\right] = \exp\left[-\gamma\mu + \frac{1}{2}\gamma^2\sigma^2\right].$$

Proof. Define $x = \ln X$. Then,

$$\begin{aligned}
\mathbb{E}\left[X^{-\gamma}\right] &= \mathbb{E}\left[\exp\left(\ln\left(X^{-\gamma}\right)\right)\right] = \mathbb{E}\left[\exp\left[-\gamma x\right]\right] \\
&= \int_{-\infty}^{\infty} \exp\left[-\gamma x\right]\phi\left(x\right)dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} \exp\left[-\gamma x\right]\exp\left[-\frac{\left(x-\mu\right)^2}{2\sigma^2}\right]dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} \exp\left[-\frac{2\sigma^2\gamma x + x^2 + \mu^2 - 2x\mu}{2\sigma^2}\right]dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} \exp\left[-\frac{x^2 + 2\left(-\mu+\sigma^2\gamma\right)x + \mu^2}{2\sigma^2}\right]dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} \exp\left[-\frac{\left(x+\left(-\mu+\sigma^2\gamma\right)\right)^2 - \left(-\mu+\sigma^2\gamma\right)^2 + \mu^2}{2\sigma^2}\right]dx \\
&= \exp\left[-\frac{-\left(\mu-\sigma^2\gamma\right)^2+\mu^2}{2\sigma^2}\right]\underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left[-\frac{\left(x-\left(\mu-\sigma^2\gamma\right)\right)^2}{2\sigma^2}\right]dx}_{=1} \\
&= \exp\left[-\frac{-\mu^2-\sigma^4\gamma^2+2\mu\sigma^2\gamma+\mu^2}{2\sigma^2}\right] \\
&= \exp\left[-\gamma\mu + \frac{1}{2}\gamma^2\sigma^2\right].
\end{aligned}$$

■

Now suppose that the aggregate endowment is log-normally distributed (instead of having m states with probabilities π_s); i.e.

$$\ln \bar{e} \sim N\left(\bar{\mu},\sigma^2\right).$$

Then, using the above result, we know that

$$\begin{aligned}
\mathbb{E}\left[\bar{e}^{-\gamma}\right] &= \exp\left[-\gamma\bar{\mu} + \gamma^2\frac{1}{2}\sigma^2\right] \\
\mathbb{E}\left[\bar{e}^{-\gamma}\bar{e}\right]\mathbb{E}\left[\bar{e}^{1-\gamma}\right] &= \exp\left[\left(1-\gamma\right)\bar{\mu} + \left(1-\gamma\right)^2\frac{1}{2}\sigma^2\right] \\
\mathbb{E}\left[\bar{e}\right] &= \exp\left[\bar{\mu}+\frac{1}{2}\sigma^2\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{1+r_2}{1+r_1} &= \frac{\exp\left[-\gamma\bar{\mu}+\gamma^2\frac{1}{2}\sigma^2\right]\exp\left[\bar{\mu}+\frac{1}{2}\sigma^2\right]}{\exp\left[\left(1-\gamma\right)\bar{\mu}+\left(1-\gamma\right)^2\frac{1}{2}\sigma^2\right]} \\
&= \frac{\exp\left[\left(1-\gamma\right)\bar{\mu}+\left(1+\gamma^2\right)\frac{1}{2}\sigma^2\right]}{\exp\left[\left(1-\gamma\right)\bar{\mu}+\left(1-\gamma\right)^2\frac{1}{2}\sigma^2\right]} \\
&= \exp\left[\left(1+\gamma^2\right)\frac{1}{2}\sigma^2 - \left(1-\gamma\right)^2\frac{1}{2}\sigma^2\right] \\
&= \exp\left[\gamma\sigma^2\right].
\end{aligned}$$

Hence, for small $\gamma\sigma^2$,

$$r_2 - r_1 \simeq \ln\left(\frac{1+r_2}{1+r_1}\right) = \gamma\sigma^2.$$

To help interpret this result, define as w, the proportional insurance premium; i.e. for a random variable X and utility unction v, w is defined as the solution to:

$$v\left(\left(1-w\right)\mathbb{E}\left[X\right]\right) = \mathbb{E}\left[v\left(X\right)\right].$$

The interpretation is that the consumer is equally happy with a fraction $(1-w)$ of the expected of value of X than with the random variable X. Suppose that $\ln X \sim N\left(\bar{\mu}, \sigma^2\right)$ and we have v as before. We wish to derive w.

Notice

$$\begin{aligned}
\mathbb{E}\left[v\left(X\right)\right] &= \frac{\mathbb{E}\left[X^{1-\gamma}\right]}{1-\gamma} = \frac{\exp\left[\left(1-\gamma\right)\bar{\mu} + \left(1-\gamma\right)^2 \frac{1}{2}\sigma^2\right]}{1-\gamma}, \\
v\left(\left(1-w\right)\mathbb{E}\left[X\right]\right) &= \frac{\left(1-w\right)^{1-\gamma}\left(\exp\left[\bar{\mu} + \frac{1}{2}\sigma^2\right]\right)^{1-\gamma}}{1-\gamma}.
\end{aligned}$$

Equating the two gives

$$\begin{aligned}
\left(1-w\right)^{1-\gamma} &= \frac{\exp\left[\left(1-\gamma\right)\bar{\mu} + \left(1-\gamma\right)^2 \frac{1}{2}\sigma^2\right]}{\exp\left[\left(1-\gamma\right)\left(\bar{\mu} + \frac{1}{2}\sigma^2\right)\right]} \\
&= \exp\left[\left(1-\gamma\right)\bar{\mu} + \left(1-\gamma\right)^2 \frac{1}{2}\sigma^2 - \left(1-\gamma\right)\left(\bar{\mu} + \frac{1}{2}\sigma^2\right)\right] \\
&= \exp\left[\left(\left(1-2\gamma+\gamma^2\right) - \left(1-\gamma\right)\right)\frac{1}{2}\sigma^2\right] \\
&= \exp\left[\gamma\left(1-\gamma\right)\frac{1}{2}\sigma^2\right] \\
\Rightarrow 1-w &= \exp\left[\frac{1}{2}\gamma\sigma^2\right].
\end{aligned}$$

Hence, for small $\gamma\sigma^2$,

$$w \simeq \frac{1}{2}\gamma\sigma^2.$$

6 Risk Aversion and Portfolio Choice

6.1 Coefficients of Risk Aversion

Definition 6.1. *(Arrow-Pratt Coefficient of Absolute Risk Aversion).* This coefficient is a measure of the curvature of the utility function around the point x, and is given by

$$ra(x) = -\frac{u''(x)}{u'(x)}.$$

The higher the coefficient, the greater is the curvature and, hence, the more *risk averse* the agent is.

Definition 6.2. (Coefficient of Relative Risk Aversion).

$$rra(x) = -\frac{u''(x)}{u'(x)}x.$$

Example 6.1. Arrow-Pratt coefficient of absolute risk aversion for various utility functions.

▷ Linear utility:

$$\begin{aligned} u(x) = ax + b \Rightarrow ra(x) &= -\frac{0}{a} = 0 \\ &\Rightarrow rra(x) = 0. \end{aligned}$$

▷ Log utility:

$$\begin{aligned} u(x) = \ln x \Rightarrow ra(x) &= -\frac{-\frac{1}{x^2}}{\frac{1}{x}} = \frac{1}{x} \\ &\Rightarrow rra(x) = 1. \end{aligned}$$

▷ Constant Relative Risk Aversion:

$$\begin{aligned} u(x) = \frac{x^{1-\gamma}-1}{1-\gamma} \Rightarrow ra(x) &= -\frac{-\gamma x^{-\gamma-1}}{x^{-\gamma}} = \frac{\gamma}{x} \\ &\Rightarrow rra(x) = \gamma. \end{aligned}$$

This utility function is a generalisation of log-utility:

$$\begin{aligned} \lim_{\gamma\to 1}\frac{x^{1-\gamma}-1}{1-\gamma} &= \lim_{\gamma\to 1}\frac{e^{(1-\gamma)\ln x}-1}{1-\gamma} \\ \text{(L'Hôpital)} &= \lim_{\gamma\to 1}\frac{-(\ln x)\,e^{(1-\gamma)\ln x}}{-1} \\ &= \ln x. \end{aligned}$$

▷ Constant Absolute Risk Aversion:

$$u(x) = -\frac{1}{a}e^{-ax} \Rightarrow ra(x) = -\frac{-ae^{-ax}}{e^{-ax}} = a$$
$$\Rightarrow rra(x) = ax.$$

Notice that ra is constant for this type of utility function as the name suggests.

6.2 Risk Aversion and Insurance Premium

6.2.1 (Absolute) insurance premium

Definition 6.3. (*Absolute insurance premium*) Absolute insurance premium p is the maximum amount that an agent is willing to pay to avoid a risk $\tilde{x}$ (a random variable); i.e.

$$u(\mathbb{E}[x] - p) = \mathbb{E}[u(\tilde{x})], \tag{6.1}$$

where p is the premium and $\tilde{x}$ is the risk. The size of p depends on the willingness of the agent to bear risk, as well as the size of the risk.

Proposition 6.1. *(Absolute risk premium) Suppose that the risk is small. Then, the absolute risk premium p is given by*

$$p = \frac{1}{2}\left(-\frac{u''(\bar{x})}{u'(\bar{x})}\right)\sigma^2, \tag{6.2}$$

where $-u''(\bar{x})/u'(\bar{x})$ measures the agent's willingness to bear absolute risk, σ^2 measures the size of the risk, and the utility function is evaluated at the expected value of the risk; i.e. $\bar{x} = \mathbb{E}[\tilde{x}]$.

Proof. First-order expansion of the left-hand side of (6.1) around $\bar{x}$ yields,

$$\begin{aligned} u(\bar{x} - p) &\approx u(\bar{x}) + u'(\bar{x})(\bar{x} - p - \bar{x}) \\ &= u(\bar{x}) - u'(\bar{x})p; \end{aligned}$$

A second-order Taylor expansion of the right-hand side of (6.1) (ignoring the expectation operator for now) around $\bar{x}$,

$$u(\tilde{x}) = u(\bar{x}) + u'(\bar{x})(\tilde{x} - \bar{x}) + \frac{u''(\bar{x})}{2}(\tilde{x} - \bar{x})^2.$$

Together, they give

$$\begin{aligned} u(\bar{x}) - u'(\bar{x})p &= \mathbb{E}\left[u(\bar{x}) + u'(\bar{x})(\tilde{x} - \bar{x}) + \frac{u''(\bar{x})}{2}(\tilde{x} - \bar{x})^2\right] \\ &= u(\bar{x}) + u'(\bar{x})\underbrace{(\mathbb{E}[\tilde{x}] - \bar{x})}_{=0} + \frac{u''(\bar{x})}{2}\underbrace{\mathbb{E}\left[(\tilde{x} - \bar{x})^2\right]}_{=\sigma^2} \\ \Rightarrow -u'(\bar{x})p &= \frac{u''(\bar{x})}{2}\sigma^2 \\ \Rightarrow p &= \frac{1}{2}\left(-\frac{u''(\bar{x})}{u'(\bar{x})}\right)\sigma^2. \end{aligned}$$

■

6.2.2 (Relative) insurance premium

Let us examine the insurance premium for a "proportional risk".

Definition 6.4. (Relative insurance premium) Relative insurance premium is the the maximum proportion of certain consumption, $\bar{x}$, that an agent is willing to pay to avoid a risk $\tilde{x} = \bar{x}(1+\varepsilon)$; i.e.

$$u((1-\rho)\bar{x}) = \mathbb{E}[u(\bar{x}(1+\varepsilon))], \tag{6.3}$$

where $\mathbb{E}[\varepsilon] = 0$ and $\mathbb{E}\left[\varepsilon^2\right] = \sigma_\varepsilon^2$.

Proposition 6.2. *(Relative insurance premium) Suppose that the risk is small. Then, the relative insurance premium ρ is given by*

$$\rho = \frac{1}{2}\left(-\frac{u''(\bar{x})\bar{x}}{u'(\bar{x})}\right)\sigma_\varepsilon^2, \tag{6.4}$$

where $-u''(\bar{x})\bar{x}/u'(\bar{x})$ measures the agent's willingness to bear relative risk, σ_ε^2 measures the size of the risk, and the utility function is evaluated at the expected value of the risk; i.e. $\bar{x} = \mathbb{E}[\tilde{x}]$.

Proof. We can relate the absolute and relative risk premia, p and ρ, using the left-hand side of (6.1) and (6.3),

$$\begin{aligned} \bar{x} - p &= (1-\rho)\bar{x} \\ \Rightarrow p &= \bar{x}\rho. \end{aligned}$$

We can also relate the size of risk using the right-hand side of (6.1) and (6.3) while noting that, here, we are interested in equating the variance of the terms inside the utility functions (since we take a second-order Taylor expansion):

$$\begin{aligned} \mathrm{Var}[\tilde{x}] &= \mathrm{Var}[\bar{x}(1+\varepsilon)] \\ \Rightarrow \sigma^2 &= \bar{x}^2\sigma_\varepsilon^2. \end{aligned}$$

Substituting these into the expression for p, (6.2),

$$\begin{aligned} \bar{x}\rho &= \frac{1}{2}\left(-\frac{u''(\bar{x})}{u'(\bar{x})}\right)\bar{x}^2\sigma_\varepsilon^2 \\ \Rightarrow \rho &= \frac{1}{2}\left(-\frac{u''(\bar{x})}{u'(\bar{x})}\bar{x}\right)\sigma_\varepsilon^2. \end{aligned}$$

Of course, we can also simply brute for by using Taylor expansion as before. First-order Taylor expansion of the left-hand side of (6.3) around $\bar{x}$ gives

$$\begin{aligned} u((1-\rho)\bar{x}) &\approx u(\bar{x}) + u'(\bar{x})((1-\rho)\bar{x} - \bar{x}) \\ &= u(\bar{x}) - u'(\bar{x})\bar{x}\rho. \end{aligned}$$

Second-order Taylor expansion of the right-hand side of (6.3) yields (ignoring the expectation operator for now):

$$\begin{aligned} u\left(\bar{x}\left(1+\varepsilon\right)\right) &= u\left(\bar{x}\right)+u'\left(\bar{x}\right)\left(\bar{x}\left(1+\varepsilon\right)-\bar{x}\right)+\frac{u''\left(\bar{x}\right)}{2}\left(\bar{x}\left(1+\varepsilon\right)-\bar{x}\right)^2 \\ &= u\left(\bar{x}\right)+u'\left(\bar{x}\right)\bar{x}\varepsilon+\frac{u''\left(\bar{x}\right)}{2}\bar{x}^2\varepsilon^2. \end{aligned}$$

Together, we therefore have

$$\begin{aligned} u\left(\bar{x}\right)-u'\left(\bar{x}\right)\bar{x}\rho &= \mathbb{E}\left[u\left(\bar{x}\right)+u'\left(\bar{x}\right)\bar{x}\varepsilon+\frac{u''\left(\bar{x}\right)}{2}\bar{x}^2\varepsilon^2\right] \\ &= u\left(\bar{x}\right)+u'\left(\bar{x}\right)\bar{x}\mathbb{E}\left[\varepsilon\right]+\frac{u''\left(\bar{x}\right)}{2}\bar{x}^2\mathbb{E}\left[\varepsilon^2\right] \\ -u'\left(\bar{x}\right)\bar{x}\rho &= \frac{u''\left(\bar{x}\right)}{2}\bar{x}^2\sigma_\varepsilon^2 \\ \Rightarrow \rho &= \frac{1}{2}\left(-\frac{u''\left(\bar{x}\right)}{u'\left(\bar{x}\right)}\bar{x}\right)\sigma_\varepsilon^2. \end{aligned}$$

■

6.3 Certainty Equivalence

A concept closely related to the insurance premium is the certainty equivalent of a risk $\tilde{x}$.

Definition 6.5. A certainty equivalent of risk $\tilde{x}$, denoted $c_e\left(\tilde{x}\right)$, is given by

$$u\left(c_e\right)=\mathbb{E}\left[u\left(\tilde{x}\right)\right].$$

Hence, c_e is the sure (deterministic) amount of consumption that will be equivalent to a given risk $\tilde{x}$. To draw parallel with (6.1),

$$c_e=\bar{x}-p=\bar{x}\left(1-\rho\right).$$

Example 6.2. (*How risk averse are you?*) Consider the following game. Suppose you make \$1,000k per year, and assume that you face the following lottery: with probability 1/2 you will win \$10k, and with probability 1/2, you lose \$10k. What will be the certainty equivalent amount for this lottery? What is the implied relative risk aversion?

Suppose you are willing to pay \$1,000 to avoid this risk. Then, you're certainty equivalent is

$$c_e=\$1,000\text{k}-\$1\text{k}=\$999\text{k}.$$

To calculate your relative risk aversion: recall that

$$\begin{aligned} c_e &= \bar{x}\left(1-\rho\right), \\ \rho &= \frac{1}{2}\left(raa\right)\sigma_\varepsilon^2. \end{aligned}$$

Solving for raa gives

$$\begin{aligned} raa &= \frac{2}{\sigma_\varepsilon^2}\rho = \frac{2}{\sigma_\varepsilon^2}\left(1-\frac{c_e}{\bar{x}}\right) \\ &= \frac{2}{(\$10\text{k}/\$1,000\text{k})^2}\left(1-\frac{\$999\text{k}}{\$1,000\text{k}}\right) \\ &= 20. \end{aligned}$$

Now suppose that your preference had a relative risk coefficient of 1 (i.e. log preferences), then what would be the certainty equivalent then? In this case, we would have

$$\begin{aligned} raa = 1 &= \frac{2}{(\$10\text{k}/\$1,000\text{k})^2}\left(1-\frac{c_e}{\$1,000\text{k}}\right) \\ \Rightarrow c_e &= \$999,950. \end{aligned}$$

That is, for you to have a relative risk aversion of one, you must be willing to pay only \$50 to avoid the risk.

6.4 Insurance Premium for large risks

So far, we focused on examining insurance premium for small risks and found that there is a simple formula given by (6.2). More generally (with larger risks), greater aversion to risk implies a higher premium but there is no simple formula. Below, we analyse some special cases which can be solved analytically. More generally, moments of $\tilde{x}$ higher than variance may matter.

Example 6.3. *(Quadratic utility)* Suppose $\mathbb{E}[\tilde{x}] = \mu$, $\text{Var}[\tilde{x}] = \sigma^2$ and utility is quadratic:

$$u(x) = x - \frac{\alpha}{2}x^2.$$

Substituting into (6.1) yields

$$\begin{aligned} u(\mu - p) &= \mathbb{E}[u(\tilde{x})] \\ \Rightarrow (\mu - p) - \frac{\alpha}{2}(\mu - p)^2 &= \mathbb{E}\left[\tilde{x} - \frac{\alpha}{2}\tilde{x}^2\right] \\ &= \mu - \frac{\alpha}{2}\underbrace{\left(\text{Var}[\tilde{x}] + \mathbb{E}[\tilde{x}]^2\right)}_{=\mathbb{E}[\tilde{x}^2]} \\ &= \mu - \frac{\alpha}{2}\left(\sigma^2 + \mu^2\right) \\ \Rightarrow -p - \frac{\alpha}{2}\left(\mu^2 - 2\mu p + p^2\right) &= -\frac{\alpha}{2}\left(\sigma^2 + \mu^2\right) \\ \Rightarrow 0 &= -\frac{\alpha}{2}p^2 - (1 - \alpha\mu)p + \frac{\alpha}{2}\sigma^2. \end{aligned}$$

Using the quadratic formula,

$$p\left(\sigma^2\right) = \frac{(1-\alpha\mu) \pm \sqrt{(1-\alpha\mu)^2 + \alpha^2\sigma^2}}{-\alpha}.$$

The positive root is then given by

$$p\left(\sigma^2\right) = \frac{(1-\alpha\mu) - \sqrt{(1-\alpha\mu)^2 + \alpha^2\sigma^2}}{-\alpha} > 0.$$

To see how this look like for small σ^2, we can use Taylor expansion and compute

$$p\left(\sigma^2\right) = p(0) + p'(0)\,\sigma^2 + o\left(\sigma^2\right),$$

where $o\left(\sigma^2\right)$ means of order smaller than σ^2; i.e.

$$\lim_{\sigma^2 \to 0} \frac{o\left(\sigma^2\right)}{\sigma^2} = 0.$$

Note that

$$\begin{aligned} p(0) &= 0, \\ p'\left(\sigma^2\right) &= \frac{1}{2\alpha}\alpha^2 \frac{1}{\sqrt{(1-\alpha\mu)^2 + \alpha^2\sigma^2}} \\ \Rightarrow p'(0) &= \frac{1}{2}\frac{\alpha}{1-\alpha\mu} \end{aligned}$$

so that

$$p\left(\sigma^2\right) \approx \frac{1}{2}\left(\frac{\alpha}{1-\alpha\mu}\right)\sigma^2.$$

Given (6.2), we realise that

$$ra(\mu) = -\frac{u''(\mu)}{u'(\mu)} = \frac{\alpha}{1-\alpha\mu}.$$

Example 6.4. *(Constant Absolute Risk Aversion and normally distributed risk).* Suppose

$$\begin{aligned} \tilde{x} &\sim N\left(\mu, \sigma^2\right), \\ u(x) &= -\frac{1}{\lambda}\exp(-\lambda x),\ \lambda > 0. \end{aligned}$$

Then, from Example 6.1, we know that

$$ra(x) = -\frac{u''(x)}{u'(x)} = \lambda,\ \forall x.$$

Substituting into (6.1) yields

$$\begin{aligned} u(\mu - p) &= \mathbb{E}\left[u(\tilde{x})\right] \\ \Rightarrow -\frac{1}{\lambda}\exp(-\lambda(\mu - p)) &= -\frac{1}{\lambda}\mathbb{E}\left[\exp(-\lambda\tilde{x})\right]. \\ \Rightarrow \exp(-\lambda(\mu - p)) &= \mathbb{E}\left[\exp(-\lambda\tilde{x})\right]. \end{aligned} \tag{6.5}$$

Since $\tilde{x} \sim N\left(\mu, \sigma^2\right)$, using Proposition 5.3,[19]

$$\mathbb{E}\left[\exp\left[-\lambda\tilde{x}\right]\right] = \exp\left[-\mu\lambda + \frac{1}{2}\lambda^2\sigma^2\right].$$

Thus, we can simplify (6.5) to

$$\begin{aligned}\exp\left(-\lambda\left(\mu - p\right)\right) &= \exp\left[-\mu\lambda + \frac{\sigma^2\lambda^2}{2}\right]\\ \Rightarrow -\left(\mu - p\right) &= -\mu + \frac{\sigma^2\lambda}{2}\\ \Rightarrow p &= \frac{1}{2}\left(\lambda\right)\sigma^2.\end{aligned}$$

Example 6.5. *(Constant Relative Risk Aversion and Log-Normal risk).* Suppose

$$\begin{aligned}\log\tilde{x} &\sim N\left(\mu, \sigma^2\right),\\ u\left(x\right) &= \frac{x^{1-\gamma}}{1-\gamma}, \gamma > 0.\end{aligned}$$

Then, from Example 6.1, we know that

$$rra\left(x\right) = -\frac{u''\left(x\right)}{u'\left(x\right)}x = \gamma, \forall x.$$

Substituting into (6.3) yields

$$\begin{aligned}u\left(\left(1-\rho\right)\bar{x}\right) &= \mathbb{E}\left[u\left(\tilde{x}\right)\right]\\ \Rightarrow \frac{\left(\left(1-\rho\right)\bar{x}\right)^{1-\gamma}}{1-\gamma} &= \mathbb{E}\left[\frac{\tilde{x}^{1-\gamma}}{1-\gamma}\right]\\ \Rightarrow \left(\left(1-\rho\right)\bar{x}\right)^{1-\gamma} &= \mathbb{E}\left[\tilde{x}^{1-\gamma}\right]. \end{aligned} \tag{6.6}$$

Recall from the previous example that, if x is normally distributed, then

$$\mathbb{E}\left[\exp\left[-\lambda x\right]\right] = \exp\left[-\mu\lambda + \frac{\sigma^2\lambda^2}{2}\right]. \tag{6.7}$$

Now suppose that $x = \log\tilde{x}$, and setting $\lambda = -1$, then

$$\bar{x} = \mathbb{E}\left[\tilde{x}\right] = \mathbb{E}\left[\exp\left(-\left(-1\right)\left(\log\tilde{x}\right)\right)\right] = \exp\left[\mu + \frac{\sigma^2}{2}\right].$$

Note also that

$$\log\left(\tilde{x}^{1-\gamma}\right) = \left(1-\gamma\right)\log\tilde{x} \sim N\left(\left(1-\gamma\right)\mu, \left(1-\gamma\right)^2\sigma^2\right).$$

Thus, setting $x = \log\tilde{x}$ and $\lambda = \left(1-\gamma\right)$ in (6.7),

$$\begin{aligned}\mathbb{E}\left[\tilde{x}^{1-\gamma}\right] &= \mathbb{E}\left[\exp\left(-\left(-\left(1-\gamma\right)\right)\left(\log\tilde{x}\right)\right)\right]\\ &= \exp\left[\left(1-\gamma\right)\mu + \frac{\sigma^2\left(1-\gamma\right)^2}{2}\right].\end{aligned}$$

[19] Note that $\exp\left[-\lambda\tilde{x}\right] = \exp\left[\ln\left(X^{-\gamma}\right)\right]$, where $x = \ln X$. So $\mathbb{E}\left[\exp\left[-\lambda\tilde{x}\right]\right] = \mathbb{E}\left[X^{-\gamma}\right]$ as we had in Proposition 5.3.

We can therefore rewrite (6.6) as

$$\begin{aligned}
\left((1-\rho)\exp\left[\mu+\frac{\sigma^2}{2}\right]\right)^{1-\gamma} &= \exp\left[(1-\gamma)\mu+\frac{\sigma^2(1-\gamma)^2}{2}\right] \\
\Rightarrow (1-\rho)^{1-\gamma}\exp\left[(1-\gamma)\mu+\frac{(1-\gamma)\sigma^2}{2}\right] &= \exp\left[(1-\gamma)\mu+\frac{\sigma^2(1-\gamma)^2}{2}\right] \\
\Rightarrow (1-\rho)\exp\left[\frac{\sigma^2}{2}\right] &= \exp\left[\frac{\sigma^2(1-\gamma)}{2}\right] \\
\Rightarrow (1-\rho) &= \exp\left[-\frac{\sigma^2\gamma}{2}\right] \\
\Rightarrow \log(1-\rho) &= -\frac{1}{2}(\gamma)\sigma^2 \\
\Rightarrow \rho &\cong \frac{1}{2}(\gamma)\sigma^2,
\end{aligned}$$

where the last line uses that $\log(1+y)\cong y$ for $y=-\rho$ small.

6.5 Arrow-Pratt Theorem

Theorem 6.1. *(Arrow-Pratt) Let u and v be utility functions. The following statements are equivalent.*

(i) If u is an increasing and concave transformation of v; i.e. there exists a function f such that

$$u(x)=f(v(x)),\forall x,$$

and,

$$\begin{aligned} f'(\cdot) &> 0, \\ f''(\cdot) &< 0. \end{aligned}$$

(ii) u has a higher insurance premium than v; i.e. for all random variables $\tilde{x}$, the insurance premium $p_u(\tilde{x})$, $p_v(\tilde{x})$ corresponding to the utility functions u and v are:

$$p_u(\tilde{x})>p_v(\tilde{x}).$$

(iii) The absolute risk aversion coefficient of u is higher than that of v everywhere; i.e.

$$-\frac{u''(x)}{u'(x)}>-\frac{v''(x)}{v'(x)},\forall x.$$

Proof. It suffices to show that: (i) implies (ii), (ii) implies (iii), and (iii) implies (i).

First, to see how (i) implies (ii). Since f is strictly concave and increasing, by Jensen's Inequality,

$$\mathbb{E}[u(\tilde{x})]=\mathbb{E}[f(v(\tilde{x}))]<f(\mathbb{E}[v(\tilde{x})]).$$

The definition of p_u, $\mathbb{E}\left[u\left(\tilde{x}\right)\right] = u\left(\bar{x} - p_u\left(\tilde{x}\right)\right)$, so that we can write

$$u\left(\bar{x} - p_u\left(\tilde{x}\right)\right) < f\left(\mathbb{E}\left[v\left(\tilde{x}\right)\right]\right).$$

Recall the definition of p_v, $\mathbb{E}\left[v\left(\tilde{x}\right)\right] = v\left(\bar{x} - p_v\left(\tilde{x}\right)\right)$, and applying f to both sides yields

$$\begin{aligned} f\left(\mathbb{E}\left[v\left(\tilde{x}\right)\right]\right) &= f\left(v\left(\bar{x} - p_v\left(\tilde{x}\right)\right)\right) \\ &= u\left(\bar{x} - p_v\left(\tilde{x}\right)\right). \end{aligned}$$

Hence, we have

$$u\left(\bar{x} - p_u\left(\tilde{x}\right)\right) < u\left(\bar{x} - p_v\left(\tilde{x}\right)\right).$$

Finally, since u is strictly increasing, above implies that

$$p_u\left(\tilde{x}\right) > p_v\left(\tilde{x}\right).$$

Second, to see how (ii) implies (iii). Consider a random variable $\tilde{x}$ with small variance centred around $\bar{x}$. Since we can choose the random variable with arbitrarily small variance, we can use the result for small risks:

$$\begin{aligned} p_u\left(\tilde{x}\right) &= \frac{1}{2}\left(-\frac{u''\left(\bar{x}\right)}{u'\left(\bar{x}\right)}\right)\sigma^2, \\ p_v\left(\tilde{x}\right) &= \frac{1}{2}\left(-\frac{v''\left(\bar{x}\right)}{v'\left(\bar{x}\right)}\right)\sigma^2. \end{aligned}$$

By assumption, $p_u\left(\tilde{x}\right) > p_v\left(\tilde{x}\right)$, so that it must be the case that

$$-\frac{u''\left(\bar{x}\right)}{u'\left(\bar{x}\right)} > -\frac{v''\left(\bar{x}\right)}{v'\left(\bar{x}\right)}.$$

Finally, we show that (iii) implies (i). Define

$$f\left(w\right) \coloneqq u\left(v^{-1}\left(w\right)\right).$$

Then, if we let $w = v\left(x\right)$,

$$f\left(v\left(x\right)\right) = u\left(v^{-1}\left(v\left(x\right)\right)\right) = u\left(x\right), \ \forall x.$$

So we show the first part of (i). We also need to show that f is strictly increasing and strictly concave. Let us differentiate the definition of f with respect to w,[20] then

$$f'(w) = u'\left(v^{-1}(w)\right)\frac{1}{v'\left(v^{-1}(w)\right)} > 0,$$

where the inequality follows from the fact that $u', v' > 0$. Differentiating the expression above again gives

$$\begin{aligned} f''(w) &= u''\left(v^{-1}(w)\right)\left(\frac{1}{v'\left(v^{-1}(w)\right)}\right)^2 - u'\left(v^{-1}(w)\right)\left(\frac{1}{v'\left(v^{-1}(w)\right)}\right)^2\frac{v''\left(v^{-1}(w)\right)}{v'\left(v^{-1}(w)\right)} \\ &= -u'\left(v^{-1}(w)\right)\left(\frac{1}{v'\left(v^{-1}(w)\right)}\right)^2\left[-\frac{u''\left(v^{-1}(w)\right)}{u'\left(v^{-1}(w)\right)} - \left(-\frac{v''\left(v^{-1}(w)\right)}{v'\left(v^{-1}(w)\right)}\right)\right]. \end{aligned}$$

By assumption $-u''/u' > -v''/v'$, $u', v' > 0$, it follows that $f''(w) < 0$. ■

6.6 Portfolio Choice Problem

We now consider a one-period problem of an investor in which he has already decided how much to invest, W, but he needs to choose how to invest the money; i.e. what assets to buy or sell. The investor has access to a menu of $1, 2, \ldots, N$ risky assets and one riskless asset denoted by 0. The risky assets are assumed to have a random gross return $\tilde{R}_i$ (for $i = 1, 2, \ldots, N$) and the riskless asset has a gross return of $\bar{\mu}$. We denote by w_i the fraction of W allocated to each asset $i = 0, 1, 2 \ldots, N$. There are S number of possible states with the probability of state s being realised given by $\pi(s)$. Once the portfolio decisions are made, the returns are realised and the investor's wealth becomes $\tilde{W}$, which is given by

$$\begin{aligned} \tilde{W} &= W\underbrace{\left(1 - \sum_{i=1}^{N} w_i\right)}_{=w_0}\bar{\mu} + W\sum_{i=1}^{N} w_i\tilde{R}_i \\ &= W\left[\sum_{i=1}^{N} w_i\tilde{R}_i - \sum_{i=1}^{N} w_i\bar{\mu} + \bar{\mu}\right] \\ &= W\left[\sum_{i=1}^{N} w_i\left(\tilde{R}_i - \bar{\mu}\right) + \bar{\mu}\right]. \end{aligned}$$

[20] *(Differentiating inverse functions)* By Leibniz's rule, we know that

$$\frac{dx}{dy}\frac{dy}{dx} = \frac{dx}{dx} = 1 \Rightarrow \frac{dx}{dy} = \frac{1}{dy/dx}.$$

Let $y = f(x)$ and suppose that $f^{-1}(y) = x$ exists. Then,

$$\begin{aligned} \frac{dy}{dx} &= \frac{d}{dx}(f(x)) = f'(x) = f'\left(f^{-1}(y)\right), \\ \frac{dx}{dy} &= \frac{d}{dy}\left(f^{-1}(y)\right) = \left(f^{-1}\right)'(y). \end{aligned}$$

Hence,

$$\frac{dx}{dy} = \frac{1}{dy/dx} \Rightarrow \left(f^{-1}\right)'(y) = \frac{1}{f'\left(f^{-1}(y)\right)}.$$

In this case, we want to differentiate $v^{-1}(w)$, where $v = f$ and $w = f$, hence,

$$\left(v^{-1}\right)'(w) = \frac{1}{v'\left(v^{-1}(w)\right)}$$

We assume that the investor chooses the weights w_i's to maximise expected utility so that the problem is:

$$\max_{\{w_i\}_{i=1}^N} \quad \mathbb{E}\left[u\left(\tilde{W}\right)\right] = \sum_{s=1}^{S} u\left(W\left[\sum_{i=1}^{N} w_i\left(R_{i,s} - \bar{\mu}\right) + \bar{\mu}\right]\right)\pi_s.$$

The first-order condition is then

$$\mathbb{E}\left[u'\left(\tilde{W}\right)\left(\tilde{R}_i - \bar{\mu}\right)\right] = 0,\ \forall i = 1, 2, \ldots, N,$$

which is equivalent to

$$u'\left(W\left[\sum_{i=1}^{N} w_i\left(R_{i,s} - \bar{\mu}\right) + \bar{\mu}\right]\right)\left(R_{j,s} - \bar{\mu}\right)\pi_s = 0,\ \forall s = 1, 2, \ldots, S,\ \forall j = 1, 2, \ldots N.$$

We now proceed to examine the second-order condition for this problem, as well as examine the sufficient (and necessary) conditions for the uniqueness of the solution. To show that the objective function is concave, we use the following two properties:

Lemma 6.1. *(Properties of concavity).*

(i) Additive of concavity: let $f(x)$ and $g(x)$ be (strictly) concave, then $h(x) = f(x) + g(x)$ is also (strictly) concave.

(ii) Preservation of concavity under strictly increasing and strictly concave transformation: let $f(x)$ and $g(x)$ be strictly concave and $f(x)$ be an strictly increasing function, then $h(x) = f(g(x))$ is also strictly concave.

Proof. (i) Recall that, if $f(x)$ and $g(x)$ are two twice continuously differentiable functions, then

$$f''(x), g''(x) < 0.$$

Differentiating $h(x)$ yields

$$\begin{aligned} h'(x) &= f'(x) + g'(x) > 0 \\ h''(x) &= f''(x) + g''(x) < 0. \end{aligned}$$

Hence, h is strictly concave.

(ii) Differentiating h again yields

$$\begin{aligned} h'(x) &= f'(g(x))\, g'(x) > 0 \\ h''(x) &= f''(g(x))\left(g'(x)\right)^2 + f'(g(x))\, g''(x) < 0 \end{aligned}$$

Thus, once again, h is strictly concave. ■

We now apply the Lemma above to our portfolio problem. First, define realisation of wealth in each state.

$$W\left(w_1, w_2, \ldots, w_N | s\right) \coloneqq W\left[\sum_{i=1}^{N} w_i\left(R_{i,s} - \bar{\mu}\right) + \bar{\mu}\right].$$

Clearly, $W\left(w_1, w_2, \ldots, w_N | s\right)$ is concave (weakly, since it is in fact linear). Then, $u\left(W\left(w_1, w_2, \ldots, w_N | s\right)\right)$ is also concave by property (ii). Then, by property (i), the objective function

$$\mathbb{E}\left[u\left(\tilde{W}\right)\right] = \sum_{s=1}^{S} u\left(W\left[\sum_{i=1}^{N} w_i\left(R_{i,s} - \bar{\mu}\right) + \bar{\mu}\right]\right)\pi_s$$

is also concave in $(w_1, w_2, \ldots, w_N)$. That is, the second-order condition for this problem is satisfied.

6.6.1 One risky asset case

In case there is only risky asset (i.e. $N = 1$), we can denote the share of wealth devoted to the (sole) risky asset as w, so that the share devoted to riskless asset is $1-w$. We also denote the return on the risky asset as by $\tilde{R}$. Then, the investor's problem is given by

$$\max_{w} \mathbb{E}\left[u\left(W\left[w\left(\tilde{R} - \bar{\mu}\right) + \bar{\mu}\right]\right)\right],$$

and the first-order condition is

$$W\mathbb{E}\left[u'\left(W\left[w_u^*\left(\tilde{R} - \bar{\mu}\right) + \bar{\mu}\right]\right)\left(\tilde{R} - \bar{\mu}\right)\right] = 0,$$

where w_u^* is the optimal allocation of wealth devoted to the risky asset given utility function u.

The following propositions show that: (i) the investor will always invest in a risky asset with a higher return than the risk-free return, independent of the degree of his risk aversion; and (ii) the more risk averse the agent is, the smaller the size of the investment in the risky asset.

Proposition 6.3. *The investor will always invest in a risky asset with a higher return than the risk-free return, no matter how risk averse the investor is. That is,*

$$w_u^* > 0 \Leftrightarrow \mathbb{E}\left[\tilde{R}\right] > \bar{\mu}.$$

Proposition 6.4. *Suppose $\mathbb{E}\left[\tilde{R}\right] > \bar{\mu}$, u is strictly concave, and that $N = 1$ (i.e. there is only one risky asset). Suppose u is more risk averse than v; i.e.*

$$-\frac{u''(x)}{u'(x)} > -\frac{v''(x)}{v'(x)}, \forall x > 0.$$

Then, the proportion of invested into the risky asset, w_u and w_v corresponding to u and v respectively, are such that

$$w_u < w_v.$$

Proof. (*Proposition 6.3*) Define $f(w)$ to be the derivative of $\mathbb{E}\left[u\left(\tilde{W}\right)\right]$ with respect to w:

$$\begin{aligned} f(w) &:= \frac{\partial}{\partial w}\mathbb{E}\left[u\left(\tilde{W}\right)\right] = \frac{\partial}{\partial w}\mathbb{E}\left[u\left(W\left[w\left(\tilde{R} - \bar{\mu}\right) + \bar{\mu}\right]\right)\right] \\ &= W\mathbb{E}\left[u'\left(W\left[w\left(\tilde{R} - \bar{\mu}\right) + \bar{\mu}\right]\right)\left(\tilde{R} - \bar{\mu}\right)\right]. \end{aligned}$$

Note that $f(w)$ is a strictly decreasing function:

$$\begin{aligned} f'(w) &= \frac{\partial^2}{\partial w^2}\mathbb{E}\left[u\left(\tilde{W}\right)\right] \\ &= W^2\mathbb{E}\left[u''\left(W\left[w\left(\tilde{R} - \bar{\mu}\right) + \bar{\mu}\right]\right)\left(\tilde{R} - \bar{\mu}\right)^2\right] < 0, \end{aligned}$$

since $u'' < 0$. Evaluating $f(w)$ at $w = 0$, then

$$\begin{aligned} f(0) &= \left.\frac{\partial}{\partial w}\mathbb{E}\left[u\left(\tilde{W}\right)\right]\right|_{w=0} \\ &= W\mathbb{E}\left[u'\left(W\bar{\mu}\right)\left(\tilde{R}-\bar{\mu}\right)\right] \\ &= Wu'\left(W\bar{\mu}\right)\mathbb{E}\left[\left(\tilde{R}-\bar{\mu}\right)\right]. \end{aligned}$$

Since $W, u' > 0$, then $\mathbb{E}\left[\tilde{R}-\bar{\mu}\right] > 0$ implies that $f(0) > 0$. Since $f(w)$ is strictly decreasing, then $f(w) > 0$, for all $w \leq 0$. Since first-order condition is that $f(w) = 0$, it follows that the optimal weight on risky asset, w^*, has to be positive. ∎

Before proving Proposition 6.4, we first prove the following Lemma.

Lemma 6.2. *Suppose u is more risk averse than v. Then, for any x and y such that $y > x$,*

$$\frac{u'(y)}{u'(x)} < \frac{v'(y)}{v'(x)}.$$

Proof. We integrate $\log u'(x)$ between x and y:

$$\begin{aligned} \log u'(y) &= \log u'(x) + \int_x^y \frac{d}{dz}\log u'(z)\,dz \\ &= \log u'(x) + \int_x^y \frac{u''(z)}{u'(z)}dz \\ \Rightarrow \log\frac{u'(y)}{u'(x)} &= \int_x^y \frac{u''(z)}{u'(z)}dz \\ \Rightarrow \frac{u'(y)}{u'(x)} &= \exp\left[-\int_x^y \frac{-u''(z)}{u'(z)}dz\right]. \end{aligned}$$

Similarly,

$$\frac{v'(y)}{v'(x)} = \exp\left[-\int_x^y \frac{-v''(z)}{v'(z)}dz\right].$$

Since u is more risk averse than v (i.e. $-u''/u' > -v''/v'$),

$$\frac{u'(y)}{u'(x)} < \frac{v'(y)}{v'(x)}.$$

∎

Proof. (*Proposition 6.4*) The idea of the proof is to use the first-order conditions divided by $u'(W\bar{\mu})$ and the previous property for returns $R_s > \bar{\mu}$ and $R_s \leq \bar{\mu}$ separately. First, note that

$$\arg\max_w \mathbb{E}\left[u\left(\tilde{W}\right)\right] = \arg\max_w \mathbb{E}\left[\frac{u\left(\tilde{W}\right)}{u'\left(W\bar{\mu}\right)}\right]$$

since $u'(W\bar{\mu})$ is a constant. The first-order condition for the right-hand side is

$$0 = \frac{\partial}{\partial w}\mathbb{E}\left[\frac{u\left(\tilde{W}\right)}{u'\left(W\bar{\mu}\right)}\right] = \mathbb{E}\left[\frac{u'\left(\tilde{W}_{w_u^*}\right)\left(\tilde{R}-\bar{\mu}\right)}{u'\left(W\bar{\mu}\right)}\right].$$

We can see that the first-order condition is sufficient for a maximum since:

$$\frac{\partial^2}{\partial w^2}\mathbb{E}\left[\frac{u\left(\tilde{W}\right)}{u'\left(W\bar{\mu}\right)}\right] = \mathbb{E}\left[\frac{u''\left(\tilde{W}\right)\left(\tilde{R}-\bar{\mu}\right)^2}{u'\left(W\bar{\mu}\right)}\right] < 0$$

since $u'' < 0$ by assumption.

We can split the states into those such that $R_s > \bar{\mu}$ and $R_s \leq \bar{\mu}$ so that the fist-order condition becomes

$$\mathbb{E}\left[\frac{u'\left(\tilde{W}_{w_u^*}\right)\left(\tilde{R}-\bar{\mu}\right)}{u'\left(W\bar{\mu}\right)}\right] = \sum_{s:R_s>\bar{\mu}} \frac{u'\left(\tilde{W}_{w_u^*,s}\right)}{u'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right)\pi_s + \sum_{s:R_s\leq\bar{\mu}} \frac{u'\left(\tilde{W}_{w_u^*,s}\right)}{u'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right)\pi_s,$$

where we recall that $\tilde{W}_{w,s} = W\left[w_u^*\left(R_s-\bar{\mu}\right)+\bar{\mu}\right]$. Similarly, for v:

$$\mathbb{E}\left[\frac{v'\left(\tilde{W}_{w_v^*}\right)\left(\tilde{R}-\bar{\mu}\right)}{v'\left(W\bar{\mu}\right)}\right] = \sum_{s:R_s>\bar{\mu}} \frac{v'\left(\tilde{W}_{w_v^*,s}\right)}{v'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right)\pi_s + \sum_{s:R_s\leq\bar{\mu}} \frac{v'\left(\tilde{W}_{w_v^*,s}\right)}{v'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right)\pi_s.$$

Since $\mathbb{E}\left[\tilde{R}\right] > \bar{\mu}$ by assumption, we know that $w_u^* > 0$ from Proposition 6.3. Hence,

$$\begin{aligned}
R_s > \bar{\mu} \Rightarrow \tilde{W}_{w,s} &= W\left[w_u^*\underbrace{\left(R_s-\bar{\mu}\right)}_{>0}+\bar{\mu}\right] > W\bar{\mu}, \\
R_s < \bar{\mu} \Rightarrow \tilde{W}_{w,s} &= W\left[w_u^*\underbrace{\left(R_s-\bar{\mu}\right)}_{<0}+\bar{\mu}\right] < W\bar{\mu}, \\
R_s = \bar{\mu} \Rightarrow \tilde{W}_{w,s} &= W\left[w_u^*\underbrace{\left(R_s-\bar{\mu}\right)}_{=0}+\bar{\mu}\right] = W\bar{\mu}.
\end{aligned}$$

Using the Lemma, we obtain

$$\begin{aligned}
\frac{v'\left(\tilde{W}_{w,s}\right)}{v'\left(W\bar{\mu}\right)} &> \frac{u'\left(\tilde{W}_{w,s}\right)}{u'\left(W\bar{\mu}\right)},\ \forall R_s > \bar{\mu}, \\
\frac{v'\left(\tilde{W}_{w,s}\right)}{v'\left(W\bar{\mu}\right)} &< \frac{u'\left(\tilde{W}_{w,s}\right)}{u'\left(W\bar{\mu}\right)},\ \forall R_s < \bar{\mu}, \\
1 = \frac{v'\left(\tilde{W}_{w,s}\right)}{v'\left(W\bar{\mu}\right)} &= \frac{u'\left(\tilde{W}_{w,s}\right)}{u'\left(W\bar{\mu}\right)},\ \forall R_s = \bar{\mu}
\end{aligned}$$

so that

$$\begin{aligned}
\frac{v'\left(\tilde{W}_{w,s}\right)}{v'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right) &> \frac{u'\left(\tilde{W}_{w,s}\right)}{u'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right),\ \forall R_s > \bar{\mu}, \\
\frac{v'\left(\tilde{W}_{w,s}\right)}{v'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right) &> \frac{u'\left(\tilde{W}_{w,s}\right)}{u'\left(W\bar{\mu}\right)}\left(R_s-\bar{\mu}\right),\ \forall R_s < \bar{\mu},
\end{aligned}$$

where the inequality for the second line follows since $R_s - \bar{\mu} < 0$. So we realise that for any state such that $R_s > \bar{\mu}$ and for any state $R_s < \bar{\mu}$, the left-hand side is greater than the right-hand side. Moreover, the two sides are equal in any state such that $R_s = \bar{\mu}$.[21] Thus,

$$\mathbb{E}\left[\frac{v'\left(\tilde{W}_w\right)\left(\tilde{R}-\bar{\mu}\right)}{v'\left(W\bar{\mu}\right)}\right] > \mathbb{E}\left[\frac{u'\left(\tilde{W}_w\right)\left(\tilde{R}-\bar{\mu}\right)}{u'\left(W\bar{\mu}\right)}\right] = 0.$$

So, w_u^* that satisfies the first-order condition with u (given by the right-hand side) would not satisfy the first-order condition with v. In particular, the first-order condition with v evaluated at w_u^* is strictly positive. Since the second-order condition is negative, this means that $w_v^* > w_u^*$ in order for the first-order condition with v evaluated at $w_v^*$to equal zero. ■

6.6.2 Special cases

We have so far analysed the portfolio decision of a risk-averse agent with general (concave) utility u. Now, we consider two special cases in which the agents only care about the expected value and the variance of their consumption/wealth, which occurs in two important cases:

- ▷ u is quadratic and the returns follow an arbitrary distribution;
- ▷ u is concave and the returns (i.e. wealth) are normally distributed (more generally, symmetrically distributed).

Example 6.6. (*Quadratic utility with arbitrarily distributed returns*) Suppose

$$u\left(x\right) = x - \frac{\alpha}{2}x^2,\ \forall x \in \left(0, \frac{1}{\alpha}\right),$$

where $\alpha > 0$. Assume that $\tilde{W}$ has support $(-\infty, 1/\alpha)$ and

$$\mathbb{E}\left[\tilde{W}\right] = \mu,\ \text{Var}\left[\tilde{W}\right] = \sigma^2.$$

In this case, the expected value of utility is a function of just the mean and the variance:

$$\begin{aligned}
\mathbb{E}\left[u\left(\tilde{W}\right)\right] &= \mathbb{E}\left[\tilde{W} - \frac{\alpha}{2}\tilde{W}^2\right] \\
&= \mathbb{E}\left[\tilde{W}\right] - \frac{\alpha}{2}\mathbb{E}\left[\tilde{W}^2\right] = \mathbb{E}\left[\tilde{W}\right] - \frac{\alpha}{2}\left(\text{Var}\left[\tilde{W}\right] + \mathbb{E}\left[\tilde{W}\right]^2\right) \\
&= \mu - \frac{\alpha}{2}\left(\sigma^2 + \mu^2\right) = V\left(\mu, \sigma\right).
\end{aligned}$$

Observe that

$$\begin{aligned}
\frac{\partial V}{\partial \mu} &= 1 - \alpha\mu > 0, \\
\frac{\partial V}{\partial \sigma} &= -\alpha\sigma < 0,
\end{aligned}$$

where the first inequality follows from the fact that $\tilde{W}$ has support $(-\infty, 1/\alpha)$ so that $\mu = \mathbb{E}\left[\tilde{W}\right] < 1/\alpha$. Hence, the objective function $V\left(\mu, \sigma\right)$ is increasing in μ and decreasing in σ.

In (σ, μ) space, the indifference curves are upward sloping (to be on the same utility, a higher μ, which would mean greater utility, requires a higher variance to offset). Utility is increasing in the North-West direction (for a fixed μ, a lower

[21] Remember,

$$\mathbb{E}\left[\frac{u'\left(\tilde{W}_w\right)\left(\tilde{R}-\bar{\mu}\right)}{u'\left(W\bar{\mu}\right)}\right] = \sum_{s=1}^{m}\frac{u'\left(\tilde{W}_{w,s}\right)}{u'\left(W\bar{\mu}\right)}\left(R_s - \bar{\mu}\right)\pi_s.$$

variance is preferred; for a fixed σ, a higher μ is preferred).

Example 6.7. *(Normally distributed returns)* Consider the case where wealth $\tilde{W}$ is normally distributed so that $\tilde{W} \sim N\left(\mu, \sigma^2\right)$ and u is concave. Then, we can express $\tilde{W} = \mu + \sigma Z$, where $Z \sim N(0,1)$.

$$\begin{aligned}\mathbb{E}\left[u\left(\tilde{W}\right)\right] &= \mathbb{E}\left[u\left(\mu+\sigma z\right)\right] \\ &= \int_{-\infty}^{+\infty} u\left(\mu+\sigma z\right)\phi\left(z\right)dz = V\left(\mu,\sigma\right),\end{aligned}$$

where $\phi(z)$ is the standard normal density function. Once again, we can express the objective function as a function of just μ and σ. Note that, if the returns of N assets are normally distributed, then for any weight w, $\tilde{W}$ is also normally distributed (remember that sum of normally distributed random variables are normal).

As before, we will find that V is increasing in μ and decreasing in σ^2.

To show that V is increasing in μ,

$$\begin{aligned}\frac{\partial V\left(\mu,\sigma\right)}{\partial\mu} &= \frac{\partial}{\partial\mu}\mathbb{E}\left[u\left(\mu+\sigma z\right)\right] \\ &= \mathbb{E}\left[\frac{\partial}{\partial\mu}u\left(\mu+\sigma z\right)\right] \\ &= \int_{-\infty}^{+\infty} u'\left(\mu+\sigma z\right)\phi\left(z\right)dz > 0\end{aligned}$$

since $u' > 0$ and $\phi(z) > 0$.

To show that V is decreasing in σ,

$$\begin{aligned}\frac{\partial V\left(\mu,\sigma\right)}{\partial\sigma} &= \frac{\partial}{\partial\sigma}\mathbb{E}\left[u\left(\mu+\sigma z\right)\right] \\ &= \mathbb{E}\left[\frac{\partial}{\partial\sigma}u\left(\mu+\sigma z\right)\right] \\ &= \int_{-\infty}^{+\infty} u'\left(\mu+\sigma z\right)z\phi\left(z\right)dz \\ &= \int_{-\infty}^{0} u'\left(\mu+\sigma z\right)z\phi\left(z\right)dz + \int_{0}^{+\infty} u'\left(\mu+\sigma z\right)z\phi\left(z\right)dz.\end{aligned}$$

Note that, for any function $g(z)$,

$$\int_{-\infty}^{0} g\left(z\right)dz = \int_{0}^{+\infty} g\left(-z\right)dz. \tag{6.8}$$

Let $g(z) = u'(\mu+\sigma z)\, z\phi(z)$, then,

$$\begin{aligned}g\left(-z\right) &= u'\left(\mu-\sigma z\right)\left(-z\right)\phi\left(-z\right) \\ &= u'\left(\mu-\sigma z\right)\left(-z\right)\phi\left(z\right)\end{aligned}$$

where the last equality follows from the fact that ϕ is symmetric. We can then write (6.8) as

$$\int_{-\infty}^{0} u'\left(\mu+\sigma z\right)z\phi\left(z\right)dz = -\int_{0}^{+\infty} u'\left(\mu-\sigma z\right)z\phi\left(z\right)dz.$$

Then,

$$\begin{aligned}\frac{\partial V(\mu,\sigma)}{\partial\sigma} &= -\int_0^{+\infty} u'(\mu-\sigma z)\, z\phi(z)\, dz + \int_0^{+\infty} u'(\mu+\sigma z)\, z\phi(z)\, dz \\ &= \int_0^{+\infty} \left[u'(\mu+\sigma z) - u'(\mu-\sigma z)\right] z\phi(z)\, dz < 0,\end{aligned}$$

where the inequality follows from the fact: (i) $u'(\mu+\sigma z) < u'(\mu-\sigma z)$ since u is concave; (ii) that we integrate over z in the positive domain; and (iii) $\phi(z) > 0$.

As before, in (σ,μ) space, the indifference curves are upward sloping and utility increases in the South-East direction.

Part IV

Optimal Control and Dynamic Programming

Overview

In this part, we introduce methods to solving discrete and continuous time dynamic optimisations problems. There are two main methods: optimal control, which writing the dynamic problem as a sequence problem, and and dynamic programming, which involves writing the dynamic problem as a recursive problem (Bellman equation). We focus on the deterministic case and in this setting, the two methods are essentially interchangeable.

For both discrete and continuous time models, we study the speed of convergence to the steady state, and then study the local dynamics and stability of the optimal decision rules (i.e. solution to the dynamic optimisation problem).

7 Euler Equations and Transversality Conditions for Dynamic Problems

In this section, we show the conditions under which Euler equations and Transversality Conditions are necessary and sufficient for a path to be optimal. We then introduce the Maximum Principle and its relation to the classical variational approach used to derive the Euler equations in continuous time. Finally, we will show how to obtain the continuous time version of the neoclassical growth model from its discrete time version counterpart.

7.1 Discrete time

7.1.1 State formulation

The elements of a dynamic programming problem are

$$\{X, \Gamma, F, \beta\},$$

where

- X is the set of states. We typically use x to denote the current state and y to denote the next-period state.
- $\Gamma : X \to X$ is the correspondence describing the feasibility constraints. That is, for each $x \in X$, $\Gamma(x)$ gives the set of feasible values for the state variable next period if the current state is x. Its graph is given by

$$\mathrm{Gr}(\Gamma) \coloneqq \left\{(y, x) \in X^2 : x \in X, y \in \Gamma(x)\right\}.$$

- $F : \mathrm{Gr}(\Gamma) \to \mathbb{R}$ is the period-return function.
- $\beta \in (0, 1)$ is the discount factor.

The *sequence problem* is then defined as follows:

$$\begin{aligned} V^*(x_0) \coloneqq \max_{\{x_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \\ s.t. \quad & x_{t+1} \in \Gamma(x_t), \ \forall t \geq 0, \\ & x_0 \text{ given.} \end{aligned}$$

7.1.2 Control-state formulation

A related notation distinguishes between controls, u_t, and states, x_t. In this notation, the sequence problem is described by $\{X, U, h, g, \beta\}$, where

- U is the set of feasible controls,
- h is the period-return function,
- g is the law of motion of the state.

The sequence problem is then defined as

$$
\begin{aligned}
V^*(x_0) := \max_{\{u_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t h(x_t, u_t) \\
s.t. \quad & x_{t+1} = g(x_t, u_t), \\
& u_t \in U, \\
& x_0 \text{ given.}
\end{aligned}
$$

In this set up, rather than choosing the state directly, the agent chooses the control variable u_t, which, together with the current state (x_t), determines the next period's state (x_{t+1}).

To see how this control-state notation is equivalent to the previous notation, set

$$
\begin{aligned}
F(x, y) &= \max_u \{h(x, u) : u \in U, y = g(x, u)\}, \\
\Gamma(x) &= \{y : \exists u \in U \, s.t. \, y = g(x, u)\}.
\end{aligned}
$$

7.1.3 Euler equations (EE) and Transversality conditions (TC)

Assume that $X \in \mathbb{R}^m$, F is C^1 and $\beta \in (0, 1)$.

Definition 7.1. *(Euler Equations in discrete time)* The path $\{x_{t+1}\}_{t=0}^{\infty}$ satisfies EE if

$$F_y(x_t, x_{t+1}) + \beta F_x(x_{t+1}, x_{t+2}) = 0, \ \forall t \geq 0.$$

If x is a vector, then EE must hold for each dimension of x.

Definition 7.2. *(Transversality Condition in discrete time)* The path $\{x_{t+1}\}_{t=0}^{\infty}$ satisfies TC if

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

Note that $\cdot$ is the inner product.

Remark 7.1. To remember where the Euler equation comes from, note that the first-order condition of the sequence problem with respect to x_{t+1} is

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

Dividing through by β^t gives the Euler equation.

7.1.4 Sufficiency

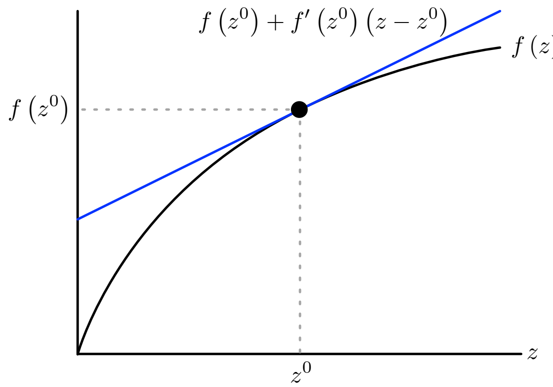
We now show that EE and TC are sufficient for optimality if the problem is convex.

Proposition 7.1. *Assume that F is concave in (x, y), $F_x\left(x_t^*, x_{t+1}^*\right) \geq 0$ and $X = \mathbb{R}_+^m$. Then, if $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ satisfies EE and TC, the path $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ is optimal.*

Proof. We use the fact that, if f is concave the tangency line at any given point always lies above f; i.e.

$$f(z) \leq f\left(z^0\right) + f'\left(z^0\right)\left(z - z^0\right), \ \forall z. \tag{7.1}$$

See also the figure below.



Now, take an arbitrary path that has the same initial condition as the optimal path: i.e. take $\{x_{t+1}\}_{t=0}^{\infty}$ with $x_0 = x_0^*$. By the assumption that $X = \mathbb{R}_+^m$, $x_{t+1} \geq 0$ for all t. We wish to show that $F\left(x_t^*, x_{t+1}^*\right)$ is greater than $F\left(x_t, x_{t+1}\right)$ across all periods; i.e. we want to show that:

$$\lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\left[F\left(x_t, x_{t+1}\right) - F\left(x_t^*, x_{t+1}^*\right)\right] \leq 0.$$

Note that we can rearrange (7.1),

$$f(z) - f\left(z^0\right) \leq f'\left(z^0\right)\left(z - z^0\right).$$

Then, letting $z^0 = x^*$,

$$F\left(x_t, x_{t+1}\right) - F\left(x_t^*, x_{t+1}^*\right) \leq F_x\left(x_t^*, x_{t+1}^*\right)\left(x_t - x_t^*\right) + F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right).$$

That is,

$$\begin{aligned}
&\lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\left[F\left(x_t, x_{t+1}\right) - F\left(x_t^*, x_{t+1}^*\right)\right] \\
&\leq \lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\left[F_x\left(x_t^*, x_{t+1}^*\right)\left(x_t - x_t^*\right) + F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right)\right].
\end{aligned}$$

Expanding the summation, we have that

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \left[F_x\left(x_t^*, x_{t+1}^*\right)\left(x_t - x_t^*\right) + F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right)\right] \\
&= F_x\left(x_0^*, x_1^*\right)\left(x_0 - x_0^*\right) + F_y\left(x_0^*, x_1^*\right)\left(x_1 - x_1^*\right) \\
&\qquad + \beta\left(F_x\left(x_1^*, x_2^*\right)\left(x_1 - x_1^*\right) + F_y\left(x_1^*, x_2^*\right)\left(x_2 - x_2^*\right)\right) \\
&\qquad + \beta^2\left(F_x\left(x_2^*, x_3^*\right)\left(x_2 - x_2^*\right) + F_y\left(x_2^*, x_3^*\right)\left(x_3 - x_3^*\right)\right) \\
&\qquad + \cdots \\
&\qquad + \beta^t\left(F_x\left(x_t^*, x_{t+1}^*\right)\left(x_t - x_t^*\right) + F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right)\right) \\
&\qquad + \beta^{t+1}\left(F_x\left(x_{t+1}^*, x_{t+2}^*\right)\left(x_{t+1} - x_{t+1}^*\right) + F_y\left(x_{t+1}^*, x_{t+2}^*\right)\left(x_{t+2} - x_{t+2}^*\right)\right) \\
&\qquad + \cdots \\
&\qquad + \beta^T\left(F_x\left(x_T^*, x_{T+1}^*\right)\left(x_T - x_T^*\right) + F_y\left(x_T^*, x_{T+1}^*\right)\left(x_{T+1} - x_{T+1}^*\right)\right).
\end{aligned}$$

By assumption, $x_0 = x_0^*$ so that the first term is zero. Rewriting the expression gives

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \left[F_x\left(x_t^*, x_{t+1}^*\right)\left(x_t - x_t^*\right) + F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right)\right] \\
&= \left[F_y\left(x_0^*, x_1^*\right)\left(x_1 - x_1^*\right) + \beta\left(F_x\left(x_1^*, x_2^*\right)\left(x_1 - x_1^*\right)\right)\right] \\
&\qquad + \beta\left[F_y\left(x_1^*, x_2^*\right)\left(x_2 - x_2^*\right) + \beta F_x\left(x_2^*, x_3^*\right)\left(x_2 - x_2^*\right)\right] \\
&\qquad + \cdots \\
&\qquad + \beta^t\left[F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right) + \beta F_x\left(x_{t+1}^*, x_{t+2}^*\right)\left(x_{t+1} - x_{t+1}^*\right)\right] \\
&\qquad + \cdots \\
&\qquad + \beta^T F_y\left(x_T^*, x_{T+1}^*\right)\left(x_{T+1} - x_{T+1}^*\right),
\end{aligned}$$

where we realise that the terms inside the square brackets are all zero by the EE. Thus, the expression above simplifies to

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \left[F_x\left(x_t^*, x_{t+1}^*\right)\left(x_t - x_t^*\right) + F_y\left(x_t^*, x_{t+1}^*\right)\left(x_{t+1} - x_{t+1}^*\right)\right] \\
&= \beta^T F_y\left(x_T^*, x_{T+1}^*\right)\left(x_{T+1} - x_{T+1}^*\right).
\end{aligned}$$

From EE, we know that $F_y\left(x_t, x_{t+1}\right) = -\beta F_x\left(x_{t+1}, x_{t+2}\right)$, hence we can write

$$\begin{aligned}
&\beta^T F_y\left(x_T^*, x_{T+1}^*\right)\left(x_{T+1} - x_{T+1}^*\right) \\
&= -\beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right)\left(x_{T+1} - x_{T+1}^*\right).
\end{aligned}$$

By assumption, we know that $x_{T+1} \geq 0$ and $F_x\left(x_{T+1}^*, x_{T+2}^*\right) \geq 0$ so that

$$\begin{aligned}
& \beta^T F_y\left(x_T^*, x_{T+1}^*\right)\left(x_{T+1}-x_{T+1}^*\right) \\
= & -\beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right)\left(x_{T+1}-x_{T+1}^*\right) \\
= & \underbrace{-\beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right) x_{T+1}}_{\leq 0}+\beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right) x_{T+1}^* \\
\leq & \beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right) x_{T+1}^* .
\end{aligned}$$

By assumption, TC holds so that

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \sum_{t=0}^T \beta^t\left[F\left(x_t, x_{t+1}\right)-F\left(x_t^*, x_{t+1}^*\right)\right] \\
\leq & \lim_{T \rightarrow \infty} \beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right) x_{T+1}^*=0 .
\end{aligned}$$

■

7.1.5 Necessity

We now prove that EE and TC are necessary conditions for the path to be optimal.

Proposition 7.2. *Assume that F is C^1 and $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ is optimal. Then, $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ satisfies EE and TC.*

Proof. We will consider adding perturbations to the optimal path $\left\{x_t^*\right\}_{t=1}^{\infty}$, denoted by ε. Let

$$x_t(\alpha, \varepsilon)=x_t^*+\alpha \varepsilon_t, \ \forall t \geq 0$$

for $\alpha \in \mathbb{R}$ and $\varepsilon=\left\{\varepsilon_t\right\}_{t=0}^{\infty}$ with $\varepsilon_t \in \mathbb{R}^m$ and $\varepsilon_0=0$ (again, we must start from the same point). Since $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ is optimal, it must be that

$$\begin{aligned}
V^*\left(x_0\right)=v(0) & \coloneqq \lim_{T \rightarrow \infty} \sum_{t=0}^T \beta^t F\left(x_t(0, \varepsilon), x_{t+1}(0, \varepsilon)\right) \\
& \geq v(\alpha) \coloneqq \lim_{T \rightarrow \infty} \sum_{t=0}^T \beta^t F\left(x_t(\alpha, \varepsilon), x_{t+1}(\alpha, \varepsilon)\right)
\end{aligned}$$

for any α, ε such that $x_{t+1}(\alpha, \varepsilon) \in \Gamma\left(x_t(\alpha, \varepsilon)\right), \ \forall t \geq 0$ (i.e. perturbed path must still be feasible).

Since $\alpha=0$ maximises v, if v is differentiable, it must be that

$$\frac{\partial v(0)}{\partial \alpha}=0 .$$

Assuming that the limits involved in the derivative (with respect to α) and in the summation (with respect to T) can be exchanged, we obtain that (since $\partial x_t(\alpha, \varepsilon) / \partial \alpha=\varepsilon_t$),

$$\frac{\partial v(0)}{\partial \alpha}=\lim_{T \rightarrow \infty} \sum_{t=0}^T \beta^t\left[F_x\left(x_t^*, x_{t+1}^*\right) \varepsilon_t+F_y\left(x_t^*, x_{t+1}^*\right) \varepsilon_{t+1}\right] .$$

Consider the summation separately,

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \left[F_x\left(x_t^*, x_{t+1}^*\right)\varepsilon_t + F_y\left(x_t^*, x_{t+1}^*\right)\varepsilon_{t+1}\right] \\
&= F_x\left(x_0^*, x_1^*\right)\varepsilon_0 + F_y\left(x_0^*, x_1^*\right)\varepsilon_1 \\
&\quad + \beta\left[F_x\left(x_1^*, x_2^*\right)\varepsilon_1 + F_y\left(x_1^*, x_2^*\right)\varepsilon_2\right] \\
&\quad + \beta^2\left[F_x\left(x_2^*, x_3^*\right)\varepsilon_2 + F_y\left(x_2^*, x_3^*\right)\varepsilon_3\right] \\
&\quad + \cdots \\
&\quad + \beta^t\left[F_x\left(x_t^*, x_{t+1}^*\right)\varepsilon_t + F_y\left(x_t^*, x_{t+1}^*\right)\varepsilon_{t+1}\right] \\
&\quad + \beta^{t+1}\left[F_x\left(x_{t+1}^*, x_{t+2}^*\right)\varepsilon_{t+1} + F_y\left(x_{t+1}^*, x_{t+2}^*\right)\varepsilon_{t+2}\right] \\
&\quad + \cdots \\
&\quad + \beta^{T-1}\left[F_x\left(x_{T-1}^*, x_T^*\right)\varepsilon_{T-1} + F_y\left(x_{T-1}^*, x_T^*\right)\varepsilon_T\right] \\
&\quad + \beta^T\left[F_x\left(x_T^*, x_{T+1}^*\right)\varepsilon_T + F_y\left(x_T^*, x_{T+1}^*\right)\varepsilon_{T+1}\right].
\end{aligned}$$

By the assumption that $\varepsilon_0 = 0$, the first term is zero. Then, we can write

$$\begin{aligned}
\frac{\partial v(0)}{\partial \alpha} &= \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \left[F_x\left(x_t^*, x_{t+1}^*\right)\varepsilon_t + F_y\left(x_t^*, x_{t+1}^*\right)\varepsilon_{t+1}\right] \\
&= \lim_{T\to\infty} \sum_{t=0}^{T-1} \beta^t \left(F_y\left(x_t^*, x_{t+1}^*\right) + \beta F_x\left(x_{t+1}^*, x_{t+2}^*\right)\right)\varepsilon_{t+1} + \lim_{T\to\infty} \beta^T F_y\left(x_T^*, x_{T+1}^*\right)\varepsilon_{T+1}.
\end{aligned}$$

Consider the case where $\varepsilon_s = 0$ for all s except at time $t+1$. In this case, $x_{t+1}(\alpha, \varepsilon)$ will be feasible if $\left(x_{t+1}^*, x_t^*\right) \in \operatorname{int}(\operatorname{Gr}(\Gamma))$ (if it was on the boundary, then perturbing would lead to $x_{t+1}(\alpha, \varepsilon)$ that is not feasible). Also, assume that v is differentiable and the limits can be interchanged. Then, it must be that

$$\frac{\partial v(0)}{\partial \alpha} = \left[F_y\left(x_t^*, x_{t+1}^*\right) + \beta F_x\left(x_{t+1}^*, x_{t+2}^*\right)\right]\varepsilon_{t+1} = 0.$$

Since this must hold for any ε_{t+1} in the neighbourhood of 0 such that $x_{t+1}(\alpha, \varepsilon)$ is feasible, it must be that the term inside the square brackets sum to zero; i.e.

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

Since this must be true for any $t+1$, we obtain the EE and that

$$0 = \frac{\partial v(0)}{\partial \alpha} = \lim_{T\to\infty} \beta^T F_y\left(x_T^*, x_{T+1}^*\right)\varepsilon_{T+1}.$$

Using the EE that we've already established, we know $F_y\left(x_T^*, x_{T+1}^*\right) = -\beta F_x\left(x_{T+1}^*, x_{T+2}^*\right)$ so that

$$0 = \frac{\partial v(0)}{\partial \alpha} = -\lim_{T\to\infty} \beta^{T+1} F_x\left(x_{T+1}^*, x_{T+2}^*\right)\varepsilon_{T+1}.$$

Finally, if $\varepsilon_{T+1} = -x^*_{T+1}$ is feasible then

$$0 = \frac{\partial v(0)}{\partial \alpha} = \lim_{T\to\infty} \beta^T F_x\left(x^*_T, x^*_{T+1}\right) x^*_T.$$

That is, TC must hold. ■

7.1.6 Steady state

Definition 7.3. (*Steady state*) $\bar{x}$ is a steady state if it is a solution to

$$F_y(\bar{x},\bar{x}) + \beta F_x(\bar{x},\bar{x}) = 0.$$

Remark 7.2. Concavity of F does not imply that steady states are unique. Concavity ensures unique solution *for any given* initial condition. However, there may be many steady states depending on where you start.

Exercise 7.1. For what kind of problems is $x_{t+1} = \bar{x}$ for $t \geq 0$ optimal if $x_0 = \bar{x}$?

Solution. Under the assumptions of Proposition 7.1.

7.1.7 EE as a second-order difference equation

EE is, in fact, a second-order difference equation; define $x_{t+2} = \psi(x_{t+1}, x_t)$ so that EE can be written as

$$F_y(x_t, x_{t+1}) + \beta F_x(x_{t+1}, \psi(x_{t+1}, x_t)) = 0.$$

To be able to solve for ψ, we need F_x to change with ψ (so that the problem is dynamic). For example, if $f(x,y) = h(x) + g(x)$, then the initial condition on y does not affect x so that the problem is not dynamic.

Exercise 7.2. Assume that F is C^2. What condition will suffice to uniquely define ψ?

Solution. Rewrite

$$F_y(x_t, x_{t+1}) = -\beta F_x(x_{t+1}, \psi(x_{t+1}, x_t)).$$

For ψ to be uniquely defined, we want to ensure that for any given (x_t, x_{t+1}), $x_{t+2} = \psi(x_{t+1}, x_t)$ is uniquely determined by the equation above. So, fix (x_t, x_{t+1}) so that the left-hand side is a constant, then if $F_{xy} > 0$ (or $F_{xy} < 0$), then the right-hand side is strictly increasing (decreasing) in x_{t+2} so that it crosses the left-hand side at most once. To ensure existence, we can impose Inada conditions on F_x.

Exercise 7.3. Write down convexity conditions on X, F, Γ so that the dynamic problem has, at most, one solution.

Solution. $X \in \mathbb{R}^m$, F is C^1 and $\beta \in (0,1)$, $F(x,y)$ is strictly concave in (x,y), $F(x,y)$ is strictly increasing in x, and Γ is convex and increasing in x (i.e. $x' > x \Rightarrow \Gamma(x) \subseteq \Gamma(x')$). check

Definition 7.4. (*Shooting algorithm*) Given x_0, select x_1 arbitrary. Generate a sequence $\{x\}$ using $x_{t+1} = \psi(x_{t+1}, x_t)$ for all $t \geq 2$. Compute if the limit of this sequence satisfies TC for the arbitrary choice of x_1. If not, try a different one.

Exercise 7.4. For what type of problems do the shooting algorithm work? Why does it work?

Solution. If the dynamic system has a unique solution and the system is convergent. Since starting from any initial condition, we will eventually converge to the steady state if we are on the saddle path. check

7.2 Continuous Time

The problem now is to choose the derivative of the state with respect to time, $\dot{x}(t)$, for each period. The elements of a dynamic programming problem are

$$\{X, \Gamma, F, \beta\},$$

where

- $\triangleright$ X is the set of states;
- $\triangleright$ $\Gamma : X \to X$ is the correspondence describing the feasibility constraints, which gives the set of feasible values for the state variable next period if the current state is x. Its graph is given by

$$\mathrm{Gr}(\Gamma) \coloneqq \left\{(\dot{x}, x) \in X^2 : x \in X, \dot{x} \in \Gamma(x)\right\};$$

- $\triangleright$ $F(x, \dot{x}) : \mathrm{Gr}(\Gamma) \to \mathbb{R}$ is the period-return function;
- $\triangleright$ $\rho \in (0, 1)$ is the discount rate.

The sequence problem can then be defined as follows:

$$\begin{aligned} V^*(x_0) \coloneqq \max_{\{\dot{x}(t)\}_{t=0}^{\infty}} \quad & \lim_{T\to\infty} \int_0^T e^{-\rho t} F(x(t), \dot{x}(t))\, dt \\ s.t. \quad & \dot{x}(t) \in \Gamma(x(t)), \forall t \geq 0, \\ & x_0 \text{ given.} \end{aligned}$$

The continuous time versions of the Euler equation and the Transversality conditions are given below.

Definition 7.5. *(EE in continuous time)* The path $\{\dot{x}(t)\}_{t=0}^{\infty}$ satisfies EE if

$$\begin{aligned} & F_x(x(t), \dot{x}(t)) + \rho F_{\dot{x}}(x(t), \dot{x}(t)) \\ & = F_{\dot{x}x}(x(t), \dot{x}(t))\, \dot{x}(t) + F_{\dot{x}\dot{x}}(x(t), \dot{x}(t))\, \ddot{x}(t), \ \forall t \geq 0. \end{aligned} \tag{7.2}$$

Loosely speaking, the continuous time counterparts to x_t, x_{t+1}, x_{t+1} are $x(t)$, $\dot{x}(t)$ and $\ddot{x}(t)$.

Definition 7.6. *(TC in continuous time)* The path $\{\dot{x}(t)\}_{t=0}^{\infty}$ satisfies TC if

$$\lim_{T\to\infty} e^{-\rho T} F_{\dot{x}}(x(T), \dot{x}(T)) \cdot x(T) = 0.$$

Remark 7.3. Problem Set 4, Q2 goes through how to derive the continuous-time version of EE from the discrete time version.

7.2.1 Sufficiency

Proposition 7.3. *Assume that $F(x, \dot{x})$ is concave in $(x, \dot{x})$, $F_{\dot{x}} \leq 0$, $\mathrm{Gr}(\Gamma)$ is convex,[a] that the optimal path $x^*(t)$ is interior, and $X = \mathbb{R}_+^m$. Then, if $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ satisfies EE and TC, the path $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ is optimal.*

[a] Convexity is required to ensure existence of a solution.

Proof. As in the discrete case, we wish to show that

$$\lim_{T\to\infty}\int_0^T e^{-\rho t}\left(F\left(x\left(t\right),\dot{x}\left(t\right)\right)-F\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\right)dt\leq 0.$$

As before, we may write

$$\begin{aligned} F\left(x\left(t\right),\dot{x}\left(t\right)\right)-F\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right) &\leq F_x\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\left(x\left(t\right)-x^*\left(t\right)\right)\\ &\quad+F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\left(\dot{x}\left(t\right)-\dot{x}^*\left(t\right)\right). \end{aligned}$$

Hence,

$$\begin{aligned} &\lim_{T\to\infty}\int_0^T e^{-\rho t}\left(F\left(x\left(t\right),\dot{x}\left(t\right)\right)-F\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\right)dt\\ &\leq\lim_{T\to\infty}\int_0^T e^{-\rho t}F_x\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\left(x\left(t\right)-x^*\left(t\right)\right)dt\\ &\quad+\lim_{T\to\infty}\int_0^T e^{-\rho t}F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\left(\dot{x}\left(t\right)-\dot{x}^*\left(t\right)\right)dt. \end{aligned}$$

Using integration by parts on the second integral:

$$\begin{aligned} &\int_0^T e^{-\rho t}F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\left(\dot{x}\left(t\right)-\dot{x}^*\left(t\right)\right)\\ &=\left[e^{-\rho t}F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\left(x\left(t\right)-x^*\left(t\right)\right)\right]_0^T-\int_0^T\mathcal{X}\left(x\left(t\right)-x^*\left(t\right)\right)dt,\\ &=e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)\left(x\left(T\right)-x^*\left(T\right)\right)-\int_0^T\mathcal{X}\left(x\left(t\right)-x^*\left(t\right)\right)dt, \end{aligned}$$

where we use the fact that $x\left(0\right)=x^*\left(0\right)$ and

$$\begin{aligned} \mathcal{X}&=\frac{d}{dt}\left[e^{-\rho t}F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\right]\\ &=-\rho e^{-\rho t}F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\\ &\quad+e^{-\rho t}F_{\dot{x}x}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\dot{x}^*\left(t\right)+e^{-\rho t}F_{\dot{x}\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\ddot{x}^*\left(t\right)\\ &=e^{-\rho t}\left[-\rho F_{\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)+F_{\dot{x}x}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\dot{x}^*\left(t\right)+F_{\dot{x}\dot{x}}\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\ddot{x}^*\left(t\right)\right] \end{aligned}$$

Suppressing some arguments,

$$\begin{aligned}
&\lim_{T\to\infty}\int_0^T e^{-\rho t}\left(F_x^*\left(x-x^*\right)+F_{\dot{x}}^*\left(\dot{x}-\dot{x}^*\right)\right)dt \\
&\leq \int_0^\infty e^{-\rho t}F_x^*\left(x-x^*\right)dt-\int_0^\infty e^{-\rho t}\left[-\rho F_{\dot{x}}^*+F_{\dot{x}x}\dot{x}^*+F_{\dot{x}\dot{x}}\ddot{x}^*\right]\left(x-x^*\right)dt \\
&\quad +\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)\left(x\left(T\right)-x^*\left(T\right)\right) \\
&=\int_0^\infty e^{-\rho t}\left[\underbrace{F_x^*+\rho F_{\dot{x}}^*-F_{\dot{x}x}\dot{x}^*-F_{\dot{x}\dot{x}}\ddot{x}^*}_{=0\because EE}\right]\left(x-x^*\right)dt+\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)\left(x\left(T\right)-x^*\left(T\right)\right) \\
&=\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)\left(x\left(T\right)-x^*\left(T\right)\right) \\
&=\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)x\left(T\right)-\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)x^*\left(T\right).
\end{aligned}$$

Since $F_{\dot{x}}\leq 0$ and $x\left(t\right)\geq 0$, then

$$\begin{aligned}
&\lim_{T\to\infty}\int_0^T e^{-\rho t}\left(F_x^*\left(x-x^*\right)+F_{\dot{x}}^*\left(\dot{x}-\dot{x}^*\right)\right)dt \\
&\leq \lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)x\left(T\right)-\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)x^*\left(T\right) \\
&\leq -\lim_{T\to\infty}e^{-\rho T}F_{\dot{x}}\left(x^*\left(T\right),\dot{x}^*\left(T\right)\right)x^*\left(T\right)=0,
\end{aligned}$$

where the last equality follows from the Transversality Condition. Thus,

$$\lim_{T\to\infty}\int_0^T e^{-\rho t}\left(F\left(x\left(t\right),\dot{x}\left(t\right)\right)-F\left(x^*\left(t\right),\dot{x}^*\left(t\right)\right)\right)dt\leq 0. \qquad \blacksquare$$

7.2.2 Necessity

Proposition 7.4. *Assume that F is C^1 and that the optimal path is interior; i.e.*

$$\left(x\left(t\right),\dot{x}\left(t\right)\right)\in\mathrm{Int}\left(\mathrm{Gr}\left(\Gamma\right)\right),\forall t\geq 0.$$

Then, the optimal path $\left\{x_{t+1}^*\right\}_{t=0}^\infty$ satisfies EE and TC.

Proof. Consider the variation path

$$x\left(\alpha,\varepsilon\right)\left(t\right)=x\left(t\right)+\alpha\varepsilon\left(t\right),$$

where $\alpha\in\mathbb{R}$ and $\varepsilon\left(t\right)$ is a differentiable function from $\mathbb{R}_+$ to $\mathbb{R}^m$ with $\varepsilon\left(0\right)=0$. Define the value of the variational path as:

$$\begin{aligned}
v\left(\alpha\right)&\coloneqq\lim_{T\to\infty}\int_0^T e^{-\rho t}F\left(x\left(\alpha,\varepsilon\right)\left(t\right),\dot{x}\left(\alpha,\varepsilon\right)\left(t\right)\right)dt \\
&=\lim_{T\to\infty}\int_0^T e^{-\rho t}F\left(x\left(t\right)+\alpha\varepsilon\left(t\right),\dot{x}\left(t\right)+\alpha\dot{\varepsilon}\left(t\right)\right)dt.
\end{aligned}$$

If the variation path is feasible—i.e. $\dot{x}\left(\alpha,\varepsilon\right)\left(t\right)\in\Gamma\left(x\left(\alpha,\varepsilon\right)\left(t\right)\right)$ for all $t\geq 0$, since $\left(x\left(t\right),\dot{x}\left(t\right)\right)$ is the optimal path, it must be that

$$v(0) \geq v(\alpha).$$

Assuming that v is differentiable and that we can interchange derivative of the integral as the integral of the derivative:

$$\begin{aligned}\frac{\partial v(0)}{\partial \alpha} = \lim_{T\to\infty} \int_0^T e^{-\rho t} F_x(x(t), \dot{x}(t))\,\varepsilon(t)\,dt \\ + \lim_{T\to\infty} \int_0^T e^{-\rho t} F_{\dot{x}}(x(t), \dot{x}(t))\,\dot{\varepsilon}(t)\,dt.\end{aligned}$$

To find the impact of the variational path ε for each t, we use integration by parts on the second term:

$$\begin{aligned}&\int_0^T e^{-\rho t} F_{\dot{x}}(x(t), \dot{x}(t))\,\dot{\varepsilon}(t)\,dt \\ &= \left[e^{-\rho t} F_{\dot{x}}(x(t), \dot{x}(t))\,\varepsilon(t)\right]_0^T - \int_0^T \mathcal{X}\varepsilon(t)\,dt \\ &= e^{-\rho T} F_{\dot{x}}(x(T), \dot{x}(T))\,\varepsilon(T) - \int_0^T \mathcal{X}\varepsilon(t)\,dt\end{aligned}$$

where we use the fact that $\varepsilon(0) = 0$ and (assuming $x(t)$ is C^2),

$$\begin{aligned}\mathcal{X} = &-\rho e^{-\rho t} F_{\dot{x}}(x(t), \dot{x}(t)) \\ &+ e^{-\rho t} F_{\dot{x}x}(x(t), \dot{x}(t))\,\dot{x}(t) + e^{-\rho t} F_{\dot{x}\dot{x}}(x(t), \dot{x}(t))\,\ddot{x}(t).\end{aligned}$$

Omitting some arguments, we therefore have that

$$\begin{aligned}0 = \frac{\partial v(0)}{\partial \alpha} = \int_0^\infty e^{-\rho t} F_x \varepsilon dt - \int_0^T e^{-\rho t}\left[-\rho F_{\dot{x}} + F_{\dot{x}x}\dot{x} + F_{\dot{x}\dot{x}}\ddot{x}\right]\varepsilon dt \\ + \lim_{T\to\infty} e^{-\rho T} F_{\dot{x}}(x(T), \dot{x}(T))\,\varepsilon(T).\end{aligned}$$

As in the discrete time case, take ε such that $\varepsilon(t) \neq 0$ and zero otherwise. Then, the expression above becomes

$$0 = e^{-\rho t}\left[F_x + \rho F_{\dot{x}} - F_{\dot{x}x}\dot{x} - F_{\dot{x}\dot{x}}\ddot{x}\right]\varepsilon.$$

Since this has to hold for any feasible values of $\varepsilon(t)$, it must be that the term inside the brackets is zero. Rearranging and writing the arguments in full, we obtain the EE:

$$F_x(x(t), \dot{x}(t)) + \rho F_{\dot{x}}(x(t), \dot{x}(t)) = F_{\dot{x}x}(x(t), \dot{x}(t))\,\dot{x}(t) - F_{\dot{x}\dot{x}}(x(t), \dot{x}(t))\,\ddot{x}(t).$$

Note that the proof is heuristic since we assumed that ε is differentiable but, for the Euler equations, we use a function ε that was discontinuous. A rigorous proof needs to approximate this discontinuous case using a smooth function.

Take a sequence x that satisfies EE then,

$$0 = \frac{\partial v(0)}{\partial \alpha} = \lim_{T\to\infty} e^{-\rho T} F_{\dot{x}}(x(T), \dot{x}(T))\,\varepsilon(T).$$

If $\varepsilon(T) = -x(T)$ is feasible, we obtain the TC:

$$0 = \lim_{T\to\infty} e^{-\rho T} F_{\dot{x}}(x(T), \dot{x}(T))\, x(T). \qquad \blacksquare$$

7.2.3 Steady state

Definition 7.7. (*Steady state*) In the steady state $\dot{x}(t) = \ddot{x}(t) = 0$. Substituting this into the EE characterises the steady state $\bar{x}$:

$$\begin{aligned} F_x(\bar{x}, 0) + \rho F_{\dot{x}}(\bar{x}, 0) &= F_{\dot{x}x}(\bar{x}, 0)\, 0 + F_{\dot{x}\dot{x}}(\bar{x}, 0)\, 0 \\ \Rightarrow F_x(\bar{x}, 0) &= -\rho F_{\dot{x}}(\bar{x}, 0). \end{aligned}$$

7.3 The Maximum Principle: Hamiltonian

We use the control-state formulation. For this, we have the instantaneous return function h that depends on the state vector $x \in X \subseteq \mathbb{R}^m$ and a control vector $u \in U \subseteq \mathbb{R}^n$ (m need not equal n). The problem is

$$\begin{aligned} V^*(x_0) \coloneqq \max_{\{u(t)\}_{t=0}^{\infty}} \quad & \lim_{T\to\infty} \int_0^T e^{-\rho t} h(x(t), u(t))\, dt \\ s.t. \quad & \dot{x}(t) = g(x(t), u(t)), \ \forall t \geq 0 \\ & u(t) \in U, \ \forall t \geq 0, \\ & x_0 \text{ given.} \end{aligned}$$

We will study a procedure to obtain necessary and sufficient conditions for an optimum. This requires some regularity conditions.

Let λ be a vector on $\mathbb{R}^m$ of *co-state* variable and H be the *current-value Hamiltonian* function defined as

$$H(x, u, \lambda) := h(x, u) + \lambda g(x, u).$$

The following conditions are necessary (under regularity assumptions) and sufficient (under regularity and convexity assumptions) for the path of x and u to be optimal:

$$\begin{aligned} H_u(x(t), u(t), \lambda(t)) &= 0, \\ \rho\lambda(t) - H_x(x(t), u(t), \lambda(t)) &= \dot{\lambda}(t), \\ g(x(t), u(t)) &= \dot{x}(t), \end{aligned}$$

for all $t \geq 0$.
The state variable(s), x, has an initial value of x_0 and the co-state variable(s), $\lambda(t)$, have a boundary condition—the Transversality Condition—given by

$$\lim_{T\to\infty} e^{-\rho T}\lambda(T)\, x(T) = 0.$$

The initial value of the co-state variable, $\lambda(0)$, is not predetermined and it has to be solved as part of the system.

The interpretation of co-state is that $e^{-\rho t}\lambda(t)$ is the marginal value at time zero of an infinitesimal increase in the state x at time t.

The first condition says that the derivative of the Hamiltonian with respect to the control is zero—this gives the optimal choice of u (one for each control variable). To interpret the second condition, recall that

$$rP = D + \dot{P},$$

can be interpreted as saying that P is the present value of D with interest rate r (D is the dividends, $\dot{P}$ is capital gain). Here, we see that $\lambda(t)$ is the present value of H_x given discount factor ρ. The condition gives the (shadow) marginal value of a unit of x. Finally, the last condition is the feasibility condition.

7.3.1 Heuristic proof

The Lagrangian Let us form the Lagrangian, using $e^{-\rho t}\lambda(t)$ for the multiplier of $\dot{x}(t) = g(x(t), u(t))$:

$$\mathcal{L}(x, u, \lambda) = \lim_{T\to\infty}\left(\int_0^T e^{-\rho t} h(x(t), u(t))\, dt + \int_0^T e^{-\rho t}\lambda(t)\left[g(x(t), u(t)) - \dot{x}(t)\right] dt\right).$$

We want to maximise $\mathcal{L}$ with respect to x and u, and minimise with respect to λ.[22] First, consider the following term, and use integration by parts to obtain:

$$\int_0^T e^{-\rho t}\lambda(t)\,\dot{x}(t)\, dt = \left[e^{-\rho t}\lambda(t)\, x(t)\right]_0^T - \int_0^T \left[-\rho e^{-\rho t}\lambda(t) + e^{-\rho t}\dot{\lambda}(t)\right] x(t)\, dt.$$

[22] Recall that Lagrangian method is an max-min problem.

We therefore have

$$\begin{aligned}\mathcal{L}(x,u,\lambda) = &\lim_{T\to\infty}\int_0^T e^{-\rho t} h(x(t),u(t))\,dt \\ &+ \lim_{T\to\infty}\int_0^T e^{-\rho t}\left[\lambda(t)\,g(x(t),u(t)) - \rho\lambda(t)\,x(t) + \dot{\lambda}(t)\,x(t)\right]dt \\ &- \lim_{T\to\infty}\left[e^{-\rho t}\lambda(t)\,x(t)\right]_0^T.\end{aligned}$$

Maximising with respect to x Since $\mathcal{L}(x,u,\lambda)$ has to be maximised by x, the first-order conditions with respect to $x(t)$ gives

$$\frac{d\mathcal{L}(x,u,\lambda)}{dx(t)} = e^{-\rho t}\left[h_x(x(t),u(t)) + \lambda(t)\,g_x(x(t),u(t)) - \rho\lambda(t) + \dot{\lambda}(t)\right] = 0.$$

Since $e^{-\rho t} > 0$, it follows that

$$\dot{\lambda}(t) = \rho\lambda(t) - h_x(x(t),u(t)) - \lambda(t)\,g_x(x(t),u(t)).$$

Note that

$$\begin{aligned}\frac{dH(x,u,\lambda)}{dx(t)} &= H_x(x(t),u(t),\lambda(t)) \\ &= h_x(x(t),u(t)) + \lambda(t)\,g_x(x(t),u(t))\end{aligned}$$

so that we can rewrite $\dot{\lambda}(t)$ as

$$\dot{\lambda}(t) = \rho\lambda(t) - H_x(x(t),u(t),\lambda(t)). \tag{7.3}$$

Maximising with respect to u The first-order condition with respect to $u(t)$ is

$$\frac{d\mathcal{L}(x,u,\lambda)}{du(t)} = e^{-\rho t}\left[h_u(x(t),u(t)) + \lambda(t)\,g_u(x(t),u(t))\right] = 0.$$

Since

$$\begin{aligned}\frac{dH(x,u,\lambda)}{du(t)} &= H_u(x(t),u(t),\lambda(t)) \\ &= h_u(x(t),u(t)) + \lambda(t)\,g_u(x(t),u(t)),\end{aligned}$$

we therefore have that

$$H_u(x(t),u(t),\lambda(t)) = 0. \tag{7.4}$$

7.3.2 Relating to the classical EE analysis

To see the relation with the classical EE analysis, consider the special case:

$$\begin{aligned}u &= \dot{x}, \\ F(x,u) &= h(x,u), \\ g(x,u) &= u.\end{aligned}$$

In this case,

$$H(x,u,\lambda) = F(x,\dot{x}) + \lambda \dot{x}$$

so that (7.4) becomes

$$\begin{aligned} H_u(x,u,\lambda) = 0 &\Rightarrow F_{\dot{x}}(x(t),\dot{x}(t)) + \lambda(t) = 0 \\ &\Rightarrow \lambda(t) = -F_{\dot{x}}(x(t),\dot{x}(t)) \end{aligned} \tag{7.5}$$

and

$$H_x(x,u,\lambda) = F_x(x(t),\dot{x}(t)).$$

Therefore, (7.3) becomes

$$\dot{\lambda}(t) = \rho\lambda(t) - F_x(x(t),\dot{x}(t)).$$

Combining the two we get that

$$\dot{\lambda}(t) = -\rho F_{\dot{x}}(x(t),\dot{x}(t)) - F_x(x(t),\dot{x}(t)).$$

Thus, $-\dot{\lambda}(t)$ is the right-hand side of the EE in (7.2). To confirm, differentiating (7.5) with respect to t gives

$$\dot{\lambda}(t) = -\frac{d}{dt}F_{\dot{x}}(x(t),\dot{x}(t)) = -F_{\dot{x}x}(x(t),\dot{x}(t))\dot{x}(t) - F_{\dot{x}\dot{x}}(x(t),\dot{x}(t))\ddot{x}(t).$$

Thus, we have the EE as in (7.2):

$$\begin{aligned} &F_x(x(t),\dot{x}(t)) + \rho F_{\dot{x}}(x(t),\dot{x}(t)) \\ &= F_{\dot{x}x}(x(t),\dot{x}(t))\dot{x}(t) + F_{\dot{x}\dot{x}}(x(t),\dot{x}(t))\ddot{x}(t). \end{aligned}$$

7.4 Neoclassical model

7.4.1 Discrete time

In the discrete-time neoclassical growth model, output can be used for consumption and investment: $C_t + I_t = G(k_t, 1)$; and the next-period capital is given by: $k_{t+1} = k_t(1-\delta) + I_t$. $G(\cdot,\cdot)$ is a neoclassical constant returns production function and δ is the depreciation rate. Combining the two gives the law of motion

$$\begin{aligned} \Rightarrow C_t &= G(k_t,1) + k_t(1-\delta) - k_{t+1} \\ &= f(k_t) - k_{t+1}, \end{aligned}$$

where we defined $f(k) = G(k,1) + (1-\delta)k$. The consumer then chooses an infinite sequence of k_{t+1} to maximise his utility (in any given period t, k_t is given). Thus, the neoclassical growth model takes the following from

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

In setting the lower bound for k_{t+1} to be zero, we are implicitly assuming that capital can be dismantled at no cost and consumed. If capital is irreversible, then the lower bound for k_{t+1} should be set to k_t.

We can fit the problem above into our general notation:

$$F\left(x,y\right)=U\left(f\left(x\right)-y\right),$$
$$\Gamma\left(x\right)=\left[0,f\left(x\right)\right].$$

EE and TC We have

$$F_x\left(x,y\right)=U'\left(f\left(x\right)-y\right)f'\left(x\right),$$
$$F_y\left(x,y\right)=-U'\left(f\left(x\right)-y\right)$$

so that the Euler equations for the neoclassical growth model is given by:

$$-U'\left(f\left(k_t\right)-k_{t+1}\right)+\beta U'\left(f\left(k_{t+1}\right)-k_{t+2}\right)f'\left(k_{t+1}\right)=0,\ \forall t\geq 0.$$

The Transversality Condition is given by

$$\lim_{t\to\infty}\beta^t U'\left(f\left(k_t\right)-k_{t+1}\right)f'\left(k_t\right)k_t=0.$$

Steady state The steady state solves

$$\beta U'\left(f\left(\bar{k}\right)-\bar{k}\right)f'\left(\bar{k}\right)=U'\left(f\left(\bar{k}\right)-\bar{k}\right)$$
$$\Leftrightarrow f'\left(\bar{k}\right)=\frac{1}{\beta}.$$

7.4.2 Continuous time

In the continuous-time neoclassical growth model, law of motion for capital is given by

$$\dot{k}\left(t\right)=f\left(k\right)-c\left(t\right)-\delta k\left(t\right),$$

where we implicitly assume unit inelastic supply of labour as in the discrete time case. Unlike in the discrete-time case above f is the production function gross of depreciation. The consumer then chooses an infinite path of $\dot{k}\left(t\right)$ to maximise his utility. Thus, the neoclassical growth model now takes the following from

$$\begin{aligned} V^*\left(k\left(0\right)\right)\coloneqq \max_{\left\{\dot{k}(t)\right\}_{t=0}^{\infty}} & \int_0^\infty e^{-\rho t}U\left(f\left(k\right)-\delta k\left(t\right)-\dot{k}\left(t\right)\right) \\ s.t. \quad & \dot{k}\in\mathbb{R}, \\ & k\left(0\right)\ \text{given}. \end{aligned}$$

We can fit the problem above into our general notation:

$$F\left(k,\dot{k}\right)=U\left(f\left(k\right)-\delta k-\dot{k}\right),$$
$$\Gamma\left(k\right)=\mathbb{R}.$$

EE and TC Note

$$\begin{aligned}
F_k &= \left(f'(k) - \delta\right) U'\left(f(k) - \delta k - \dot{k}\right), \\
F_{\dot{k}} &= -U'\left(f(k) - \delta k - \dot{k}\right), \\
F_{\dot{k}k} &= -\left(f'(k) - \delta\right) U''\left(f(k) - \delta k - \dot{k}\right), \\
F_{\dot{k}\dot{k}} &= U''\left(f(k) - \delta k - \dot{k}\right).
\end{aligned}$$

The EE is therefore given by

$$\begin{aligned}
&\left(f'(k) - \delta\right) U' - \rho U' = -\left(f'(k) - \delta\right) U''\dot{k} + U''\ddot{k} \\
\Leftrightarrow\ &\left(f'(k) - \delta - \rho\right) U' = -\left(\left(f'(k) - \delta\right)\dot{k} - \ddot{k}\right) U''.
\end{aligned} \tag{7.6}$$

Steady state The steady state is given by

$$\left(f'(k) - \delta - \rho\right) U' = 0 \Rightarrow f'(k) = \rho + \delta.$$

7.4.3 Deriving the continuous time EE from the discrete time version

Let Δ denote the length of time between periods when the state is determined. Decisions are taken whenever the state is determined at times 0, Δ, 2Δ, ... The sequence of state is then

$$\left\{x_{\Delta(s+1)}\right\}_{s=0}^{\infty} = \left\{x_0, x_\Delta, x_{2\Delta}, \ldots\right\},$$

where x_0, as usual, is given. We define the discount factor β for an interval Δ using (the instantaneous) discount rate ρ as

$$\beta = \frac{1}{1 + \Delta\rho}.$$

We let U denote the instantaneous utility from consuming c_t amount of consumption. So, during an interval of length Δ in which consumption is given by $c_{s\Delta}$, the total utility is given by

$$\Delta U\left(c_{s\Delta}\right).$$

Market clearing condition must hold at all times (including during the interval Δ). Instantaneous market clearing condition is thus

$$c_s + i_s = f\left(k_s\right),$$

where i_s denotes investment. Market clearing condition for the interval of length Δ from period $t\Delta$ is thus

$$\Delta c_{s\Delta} + \Delta i_{s\Delta} = \Delta f\left(k_{s\Delta}\right).$$

Law of motion for capital is

$$k_{s\Delta+\Delta} = k_{(s+1)\Delta} = \Delta i_{s\Delta} + k_{s\Delta}\left(1 - \Delta\delta\right),$$

where δ is the instantaneous depreciation rate. Thus, we can write the neoclassical growth model: as

$$\begin{aligned}
\max_{\{c_t, i_t\}_{t=0}^{\infty}} \quad & \sum_{s=0}^{\infty} \left(\frac{1}{1+\Delta\rho} \right)^s \Delta U\left(c_{s\Delta}\right), \\
s.t. \quad & \Delta c_t + \Delta i_t = \Delta f\left(k_t\right), \ \forall t \geq 0, \\
& k_{t+\Delta} = \Delta i_t + k_t\left(1-\Delta\delta\right), \ \forall t \geq 0, \\
& k_0 \text{ given},
\end{aligned}$$

where $t = s\Delta$ for some integer s. In other words, we may write the problem in an equivalent manner as

$$\begin{aligned}
\max_{\{c_{i\Delta}, i_{i\Delta}\}_{i=0}^{\infty}} \quad & \sum_{s=0}^{\infty} \left(\frac{1}{1+\Delta\rho} \right)^s \Delta U\left(c_{s\Delta}\right), \\
s.t. \quad & \Delta c_{s\Delta} + \Delta i_{s\Delta} = \Delta f\left(k_{s\Delta}\right), \ \forall s \geq 0, \\
& k_{(s+1)\Delta} = \Delta i_{s\Delta} + k_{s\Delta}\left(1-\Delta\delta\right), \ \forall s \geq 0, \\
& k_0 \text{ given},
\end{aligned}$$

Letting $\Delta = 1$, the problem reduces to the standard one.

Continuous-time version of the law of motion for capital Fix Δ. First, we eliminate Δi_t from the constraint to obtain the law of motion for capital:

$$k_{t+\Delta} = \Delta f\left(k_t\right) + k_t\left(1-\Delta\delta\right) - \Delta c_t. \tag{7.7}$$

Rearranging above and dividing through by Δ:

$$\frac{k_{t+\Delta} - k_t}{\Delta} = f\left(k_t\right) - \delta k_t - c_t. \tag{7.8}$$

Taking limits as $\Delta \downarrow 0$:

$$\lim_{\Delta \downarrow 0} \frac{k_{t+\Delta} - k_t}{\Delta} := \dot{k}\left(t\right) = f\left(k\left(t\right)\right) - \delta k\left(t\right) - c\left(t\right),$$

where we change notation following the convention (capital at time t is written k_t in discrete time and $k\left(t\right)$ in continuous time). This gives the continuous time version of the law of motion for capital.

Continuous-time version of EE Define

$$\dot{k}_t := \frac{k_{t+\Delta} - k_t}{\Delta} \Rightarrow k_{t+\Delta} = \dot{k}_t\Delta + k_t.$$

We can then rewrite (7.8), while noting that $t = s\Delta$, as

$$\begin{aligned}
c_t = c_{s\Delta} &= f\left(k_t\right) - \delta k_t - \frac{k_{t+\Delta} - k_t}{\Delta} \\
&= f\left(k_t\right) - \delta k_t - \dot{k}_t
\end{aligned}$$

Using (7.7), we can then write the problem as

$$\max_{\{\dot{k}_t\}_{t=0}^{\infty}} \quad \sum_{t=0}^{\infty} \left(\frac{1}{1+\Delta\rho} \right)^t \Delta U\left(f\left(k_t\right) - \delta k_t - \dot{k}_t \right).$$

Writing the sequence as

$$
\begin{aligned}
& \cdots + \left(\frac{1}{1+\Delta\rho}\right)^{t} \Delta U\left(f\left(k_t\right) - \delta k_t - \dot{k}_t\right) + \left(\frac{1}{1+\Delta\rho}\right)^{t+1} \Delta U\left(f\left(k_{t+\Delta}\right) - \delta k_{t+\Delta} - \dot{k}_{t+\Delta}\right) \cdots \\
= & \cdots + \left(\frac{1}{1+\Delta\rho}\right)^{t} \Delta U\left(f\left(k_t\right) - \delta k_t - \dot{k}_t\right) + \left(\frac{1}{1+\Delta\rho}\right)^{t+1} \Delta U\left(f\left(\dot{k}_t\Delta + k_t\right) - \delta\left(\dot{k}_t\Delta + k_t\right) - \frac{k_{t+2\Delta} - k_{t+\Delta}}{\Delta}\right) \cdots \\
= & \cdots + \left(\frac{1}{1+\Delta\rho}\right)^{t} \Delta U\left(f\left(k_t\right) - \delta k_t - \dot{k}_t\right) + \left(\frac{1}{1+\Delta\rho}\right)^{t+1} \Delta U\left(f\left(\dot{k}_t\Delta + k_t\right) - \delta\left(\dot{k}_t\Delta + k_t\right) - \frac{k_{t+2\Delta} - \left(\dot{k}_t\Delta + k_t\right)}{\Delta}\right) \cdots \\
= & \cdots + \left(\frac{1}{1+\Delta\rho}\right)^{t} \Delta U\left(f\left(k_t\right) - \delta k_t - \dot{k}_t\right) + \left(\frac{1}{1+\Delta\rho}\right)^{t+1} \Delta U\left(f\left(\dot{k}_t\Delta + k_t\right) - \delta\left(\dot{k}_t\Delta + k_t\right) - \frac{k_{t+2\Delta} - k_t}{\Delta} + \dot{k}_t\right) \cdots
\end{aligned}
$$

The first-order condition with respect to $\dot{k}_t$ is

$$
-\left(\frac{1}{1+\Delta\rho}\right)^{t} \Delta U'\left(c_t\right) + \left(\frac{1}{1+\Delta\rho}\right)^{t+1} \Delta U'\left(c_{t+\Delta}\right)\left(\Delta f'\left(k_{t+\Delta}\right) - \Delta\delta + 1\right) = 0.
$$

Rearranging this gives the EE:

$$
U'\left(c_t\right) = \left(\frac{1}{1+\Delta\rho}\right) U'\left(c_{t+\Delta}\right)\left(\Delta f'\left(k_{t+\Delta}\right) + \left(1 - \Delta\delta\right)\right).
$$

Taylor expansion of $U'\left(c_{t+\Delta}\right)$ around c_t gives

$$
U'\left(c_{t+\Delta}\right) = U'\left(c_t\right) + U''\left(c_t\right)\left(c_{t+\Delta} - c_t\right) + R\left(c_{t+\Delta} - c_t\right), \tag{7.9}
$$

where $R\left(c_{t+\Delta} - c_t\right) = o\left(\Delta\right)$ as $\Delta \to 0$; i.e.

$$
\lim_{\Delta\downarrow 0} \frac{R\left(c_{t+\Delta} - c_t\right)}{\Delta} = 0.
$$

Substituting (7.9) into the EE:

$$
\left(1 + \Delta\rho\right) U'\left(c_t\right) = \left(U'\left(c_t\right) + U''\left(c_t\right)\left(c_{t+\Delta} - c_t\right) + R\left(c_{t+\Delta} - c_t\right)\right)\left(\Delta f'\left(k_{t+\Delta}\right) + \left(1 - \Delta\delta\right)\right).
$$

Collecting $U'\left(c_t\right)$ together

$$
\begin{aligned}
& \left[\left(1 + \Delta\rho\right) - \Delta f'\left(k_{t+\Delta}\right) - \left(1 - \Delta\delta\right)\right] U'\left(c_t\right) \\
& = \left[U''\left(c_t\right)\left(c_{t+\Delta} - c_t\right) + R\left(c_{t+\Delta} - c_t\right)\right]\left[\Delta f'\left(k_{t+\Delta}\right) + \left(1 - \Delta\delta\right)\right].
\end{aligned}
$$

Dividing through by Δ:

$$
\begin{aligned}
& \left[\rho - f'\left(k_{t+\Delta}\right) + \delta\right] U'\left(c_t\right) \\
= & \left(U''\left(c_t\right) \frac{\left(c_{t+\Delta} - c_t\right)}{\Delta} + \frac{R\left(c_{t+\Delta} - c_t\right)}{\Delta}\right)\left(\Delta f'\left(k_{t+\Delta}\right) + \left(1 - \Delta\delta\right)\right).
\end{aligned}
$$

Taking limits as $\Delta \downarrow 0$ of each side

$$
\lim_{\Delta\downarrow 0} \left[\rho - f'\left(k_{t+\Delta}\right) + \delta\right] U'\left(c_t\right) = \left[\rho - f'\left(k\left(t\right)\right) + \delta\right] U'\left(c\left(t\right)\right),
$$

and

$$\lim_{\Delta\downarrow 0}\left(U''(c_t)\frac{(c_{t+\Delta}-c_t)}{\Delta}+\frac{R(c_{t+\Delta}-c_t)}{\Delta}\right)(\Delta f'(k_{t+\Delta})+(1-\Delta\delta))=U''(c(t))\dot{c}(t),$$

where

$$\dot{c}(t):=\lim_{\Delta\downarrow 0}\frac{(c_{t+\Delta}-c_t)}{\Delta}.$$

Combining the two and rearranging, we obtain

$$(f'(k(t))-\delta-\rho)U'(c(t))=-U''(c(t))\dot{c}_t,$$

where

$$\begin{aligned} c(t) &= f(k(t))-\delta k(t)-\dot{k}(t)\\ \Rightarrow \dot{c}(t) &= (f'(k(t))-\delta)\dot{k}(t)-\ddot{k}(t). \end{aligned}$$

Therefore, we have

$$(f'(k(t))-\delta-\rho)U'(c(t))=-U''(c(t))\left[(f'(k(t))-\delta)\dot{k}(t)-\ddot{k}(t)\right].$$

as we had in (7.6).

7.4.4 Hamiltonian

We can also analyse the neoclassical model using the Hamiltonian. The period-return function is $U(c)$ and the law of motion is given by $\dot{k}=f(k)-\delta k-c$ (note that $k(0)$ is given). That is,

$$\begin{aligned} h(k,c) &= U(c),\\ g(k,c) &= f(k)-\delta k-c,\\ \Rightarrow H(k,c,\lambda) &= U(c)+\lambda(f(k)-\delta k-c). \end{aligned}$$

Then,

$$\begin{aligned} H_u=0 &\Rightarrow U'(c)=\lambda,\\ \dot{\lambda}=\rho\lambda-H_x &\Rightarrow \dot{\lambda}=\lambda(\rho-(f'(k)-\delta)),\\ \dot{x}=g &\Rightarrow \dot{k}=f(k)-\delta k-c. \end{aligned}$$

We therefore have the following dynamic equations:

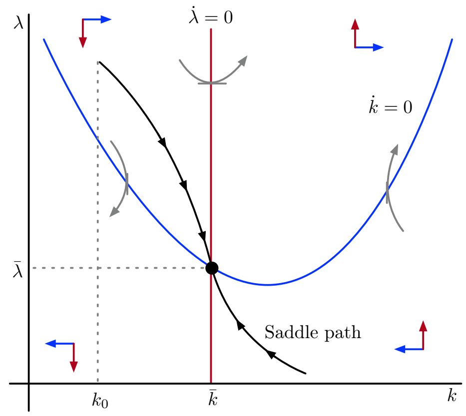
$$\begin{aligned} \dot{\lambda} &= \lambda(\rho-(f'(k)-\delta)),\\ \dot{k} &= f(k)-\delta k-c. \end{aligned}$$

Phase diagram in (k,λ) space To draw the phase diagram in (k,λ) space, note that

- ▷ $\dot{\lambda}=0$: $f'\left(\bar{k}\right)-\delta=\rho$. In (k,λ) space, this is a vertical line. Dynamics: From $\dot{\lambda}=0$ if $k>\bar{k}$, then $f'(k)$ is lower and $-f'(k)$ is higher so that $\dot{\lambda}>0$. And if $k<\bar{k}$ then $\dot{\lambda}<0$.
- ▷ $\dot{k}=0$: $c=f(k)$. Since $\lambda=U'(c)$, we have that $\lambda=U'(f(k))$. Note that $U'>0$ and $U''<0$ so that $U'(\cdot)$ is a strictly decreasing transformation. Thus, when $f(k)-\delta k$ achieves its maximum (at $\hat{k}$ such that $f'\left(\hat{k}\right)=\delta$),

$U'(f(k)-\delta)$ is at its minimum. As $k \to 0$, $f(k) \to 0$ and $U'(f(k)-\delta) \to \infty$. As $k > \hat{k}$, $f\left(\hat{k}\right)$ falls so that λ increases. These observations imply that $\dot{k}=0$ locus is U-shaped in (k,λ) space. Dynamics: If $c > \bar{c}$, then $\dot{k} < 0$ and if $c < \bar{c}$, then $\dot{k} > 0$. Note that $c > \bar{c}$ represents points below the $\dot{k}=0$ locus.

This gives the following phase diagram.



We see that low value of k corresponds to a higher value of λ. Does it make sense? Recall that λ is the marginal value of a unit of k and, at the optimal, marginal benefit from c is equated to λ, the marginal value of capital. In a similar way that higher c implies lower marginal benefit, a higher k implies lower marginal value of capital; i.e. it represents diminishing returns. Note also that a higher k implies higher production, and since c is a normal good (this is due to the fact that we have separable utility function), higher output implies higher income and c increases in every period.

Phase diagram in (k,c) space Since there is a one-to-one mapping between c and λ, we can also draw the phase diagram in (k,c) space. To do so, we can differentiate $H_u=0$ condition with respect to time to obtain $\dot{\lambda}$ in terms of c:

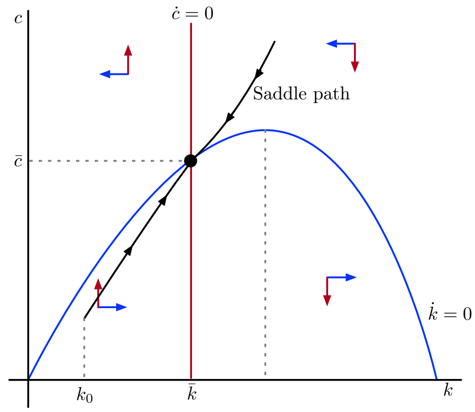
$$U''(c)\,\dot{c} = \dot{\lambda}.$$

Then we can rewrite the second first-order condition as

$$\begin{aligned} U''(c)\,\dot{c} &= U'(c)\left(\rho - \left(f'(k)-\delta\right)\right) \\ \Rightarrow \frac{\dot{c}}{c} &= \frac{1}{-\frac{U''(c)c}{U'(c)}}\left(f'(k)-\delta-\rho\right), \end{aligned}$$

where $-\frac{U''(c)c}{U'(c)}$ is the elasticity of intertemporal substitution, and $\dot{c}/c$ is the percentage change in c (i.e. elasticity). Thus, we see that given $f'(k)$ and ρ, c is given by the preference. Notice also that, if $f'(k)-\delta=\rho$, then c is constant, and if $f'(k)\neq\rho$, then as u becomes linear—i.e. $U''\to 0$—$\dot{c}$ is larger so that you converge more quickly to the steady state.

In fact, speed of convergence here is affected by both the curvature of the utility function as well as the curvature of production function as we see study in the next section.


c
$\dot{c} = 0$
Saddle path
$\bar{c}$
$\dot{k} = 0$
k_0
$\bar{k}$
k

7.5 Summary

7.5.1 Discrete time

State formation

$$\max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F\left(x_t, x_{t+1}\right)$$
$$\text{s.t.} \quad x_{t+1} \in \Gamma\left(x_t\right), \ \forall t \geq 0,$$
$$x_0 \text{ given.}$$

Necessary and sufficient conditions[23]
Euler equation:

$$F_y\left(x_t, x_{t+1}\right) + \beta F_x\left(x_t, x_{t+1}\right) = 0, \ \forall t \geq 0.$$

Transversality condition:

$$\lim_{T\to\infty} \beta^T F_x\left(x_T, x_{T+1}\right) x_T = 0.$$

Control-state formation

$$\max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t h\left(x_t, u_t\right)$$
$$\text{s.t.} \quad x_{t+1} = g\left(x_t, u_t\right), \ \forall t \geq 0,$$
$$u_t \in U$$
$$x_0 \text{ given.}$$

We can return to the state formation by letting

$$F\left(x, y\right) = \max\left\{h\left(x, u\right) : u \in U, y = g\left(x, u\right)\right\},$$
$$\Gamma\left(x\right) = \left\{y : \exists u \in U \ \text{s.t.} \ y = g\left(x, u\right)\right\}.$$

7.5.2 Continuous time

State formation

$$\max_{\{\dot{x}(t)\}_{t=0}^{\infty}} \int_0^{\infty} e^{-\rho t} F\left(x\left(t\right), \dot{x}\left(t\right)\right) dt$$
$$\text{s.t.} \quad \dot{x}\left(t\right) \in \Gamma\left(x\left(t\right)\right), \ \forall t \geq 0,$$
$$x_0 \text{ given.}$$

Necessary and sufficient conditions
Euler equation:

$$F_x + \rho F_{\dot{x}} = F_{\dot{x}x}\dot{x} + F_{\dot{x}\dot{x}}\ddot{x}, \ \forall t \geq 0$$

Transversality condition:

$$\lim_{T\to\infty} e^{-\rho T} F_{\dot{x}}\left(x\left(T\right), \dot{x}\left(T\right)\right) x\left(T\right) = 0.$$

Control-state formation

$$\max_{\{u(t)\}_{t=0}^{\infty}} \int_0^{\infty} e^{-\rho t} h\left(x\left(t\right), u\left(t\right)\right)$$
$$\text{s.t.} \quad \dot{x}\left(t\right) = g\left(x\left(t\right), u\left(t\right)\right), \ \forall t \geq 0,$$
$$u\left(t\right) \in U$$
$$x_0 \text{ given.}$$

Necessary and sufficient conditions
First-order conditions:

$$H\left(x, u, \lambda\right) = h\left(x, u\right) + \lambda g\left(x, u\right),$$
$$H_u\left(x, u, \lambda\right) = 0,$$
$$\dot{\lambda} = \rho\lambda - H_x\left(x, u, \lambda\right),$$
$$\dot{x} = g\left(x, u\right).$$

Transversality condition:

$$\lim_{T\to\infty} e^{-\rho T} \lambda\left(T\right) x\left(T\right) = 0.$$

[23] F concave and C^1.

8 Local stability of optimal paths and speed of convergence

Given an initial state x_0, the solution to a dynamic programming problem completely determines the evolution of the state through time. The purpose of this section is to study the local dynamics and stability of the optimal decision rules in discrete and continuous time models.

8.1 Stability of discrete-time linear dynamic systems of one dimension

8.1.1 Linearisation around the steady state

Let $x_{t+1} = g\left(x_t\right)$ be the optimal decision rule. Using first-order Taylor approximation around $x_t = x^*$ (x^* denotes the steady-state value),

$$x_{t+1} = g\left(x\right) \simeq g\left(x^*\right) + g'\left(x^*\right)\left(x_t - x^*\right).$$

Since $g\left(x^*\right) = x^*$,

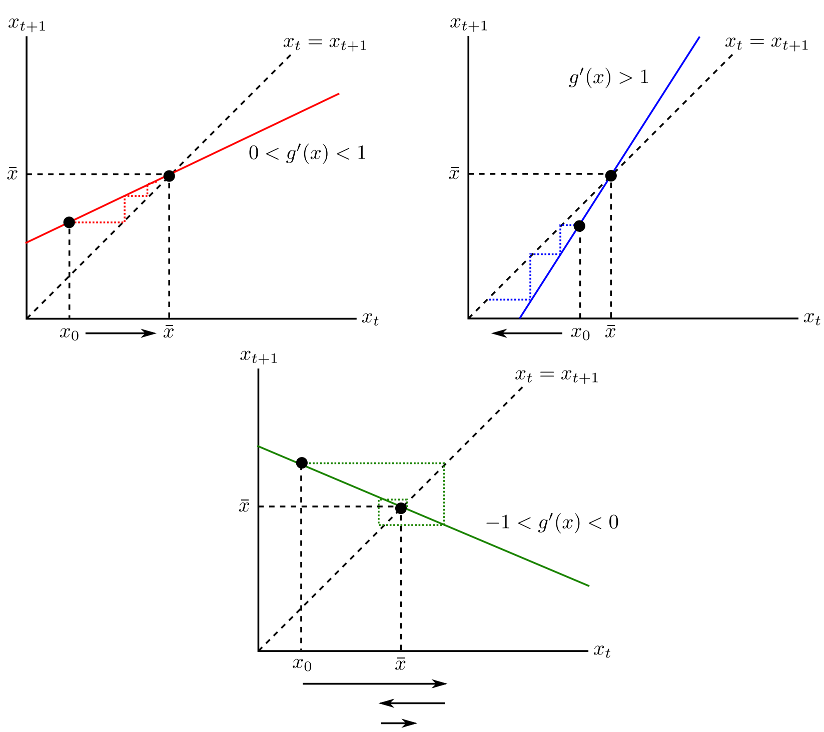
$$x_{t+1} - x^* \simeq g'\left(x^*\right)\left(x_t - x^*\right),$$

8.1.2 Condition for convergence

Suppose that $x_0 < \bar{x}$.

- ▷ If $\left|g'\left(x^*\right)\right| > 1$, then $x_{t+1} - x^* > x_t - x^*$; i.e. x_t diverges.
- ▷ If $\left|g'\left(x^*\right)\right| < 1$, then $x_{t+1} - x^* < x_t - x^*$; i.e. x_t converges.

Hence, for local stability, we require $\left|g'\left(x^*\right)\right| < 1$. See also the figure below.



Remark 8.1. If $-1 < g'\left(x^*\right) < 0$, then we still converge to the steady-state while oscillating above and below x^*. This cannot happen in continuous time case since the movement are infinitesimally small.

8.1.3 Obtaining $g'(x)$

Recall the Euler equation (which must hold for any interior solutions):

$$0 = F_y(x, g(x)) + \beta F_x(g(x), g(g(x))).$$

We want to find out $g'(x^*)$, where x^* is the steady state; i.e. x^* solves

$$F_y(x^*, x^*) + \beta F_x(g^*, g^*) = 0.$$

This would allow us to analyse the dynamics of x_t close to a steady state.

We follow the following steps to obtain $g'(x)$ close to a steady state.

(i) Differentiate the Euler equation with respect to x. This yields a quadratic equation in $g'(x)$.

(ii) Evaluate the resulting quadratic equation at the steady state, x^*.

Differentiating the Euler equation with respect to x yields

$$\begin{aligned} 0 &= F_{yx}(x, g(x)) + F_{yy}(x, g(x))\, g'(x) \\ &\quad + \beta\left[F_{xx}(g(x), g(g(x)))\, g'(x) + F_{xy}(g(x), g(g(x)))\, g'(g(x))\, g'(x)\right]. \end{aligned}$$

Evaluating this derivative at the steady state x^* and using the fact that $g(x^*) = x^*$,

$$\begin{aligned} 0 &= F_{yx}(x^*, x^*) + F_{yy}(x^*, x^*)\, g'(x^*) \\ &\quad + \beta\left[F_{xx}(x^*, x^*)\, g'(x^*) + F_{xy}(x^*, x^*)\, (g'(x^*))^2\right] \\ &= \beta F_{xy}(x^*, x^*)\, (g'(x^*))^2 + (F_{yy}(x^*, x^*) + \beta F_{xx}(x^*, x^*))\, g'(x^*) + F_{yx}(x^*, x^*) \end{aligned} \tag{8.1}$$

Thus, we obtain a quadratic equation in $g'(x^*)$ so that there may potentially be two candidate values for $g'(x^*)$.

Let $\lambda = g'(x^*)$, we can define the quadratic equation:

$$\tilde{Q}(\lambda) \coloneqq \beta F_{xy}\lambda^2 + (F_{yy} + \beta F_{xx})\lambda + F_{yx}. \tag{8.2}$$

The following proposition shows that the roots of $\tilde{Q}(\lambda)$ come in *almost reciprocal pairs*.

Proposition 8.1. *If λ_1 solves $\tilde{Q}(\lambda_1) = 0$, then so does $\lambda_2 = 1/\lambda_1\beta$.*

Proof. Suppose $\tilde{Q}(\lambda_1) = 0$; i.e.

$$\tilde{Q}(\lambda_1) = \beta F_{xy}\lambda_1^2 + (F_{yy} + \beta F_{xx})\lambda_1 + F_{yx} = 0.$$

Consider $\tilde{Q}\left(1/\lambda_1\beta\right)$:

$$\begin{aligned}
\tilde{Q}\left(\frac{1}{\lambda_1\beta}\right) &= \beta F_{xy}\left(\frac{1}{\lambda_1\beta}\right)^2 + \left(F_{yy}+\beta F_{xx}\right)\left(\frac{1}{\lambda_1\beta}\right) + F_{yx} \\
&= F_{xy}\frac{1}{\lambda_1^2\beta} + \left(F_{yy}+\beta F_{xx}\right)\left(\frac{1}{\lambda_1\beta}\right) + F_{yx} \\
&= \frac{1}{\lambda_1^2\beta}\left[F_{xy} + \left(F_{yy}+\beta F_{xx}\right)\lambda_1 + \beta F_{yx}\lambda_1^2\right] \\
[F_{xy}=F_{yx}] &= \frac{1}{\lambda_1^2\beta}\left[\beta F_{xy}\lambda_1^2 + \left(F_{yy}+\beta F_{xx}\right)\lambda_1 + F_{yx}\right] \\
&= \frac{1}{\lambda_1^2\beta}\tilde{Q}\left(\lambda_1\right) = 0.
\end{aligned}$$

■

Therefore, if one root is $|\lambda_1| < 1$, then the other root, λ_2, must be larger than one in absolute value. i.e.

$$|\lambda_2| = \left|\frac{1}{\beta\lambda_1}\right| = \frac{1}{\beta\left|\lambda_1\right|} > 1.$$

As an aside, note that $\lambda_1\lambda_2 = 1/\beta$.

Let x_0 be close to the steady state x^* so that a linear approximation of g is appropriate. Assume we found that the smaller root has absolute value less than one. Now, consider the following sequence of $\{x_{t+1}\}$:

$$x_{t+1} = x^* + g'\left(x^*\right)\left(x_t - x^*\right),\ \forall t \geq 0. \tag{8.3}$$

The sequence, by construction, satisfies the Euler Equations. Since $|g'\left(x^*\right)| < 1$, it converges to the steady state x^* and hence it also satisfies the transversality condition. Thus, if the problem is convex, we have found a solution. If, on the other hand, both roots are bigger than one in absolute value, then we do not know which one describes $g'\left(x^*\right)$, but we do know that the steady state is not locally stable (local since we are approximating).[24]

We now analyse under what circumstances local stability (i.e. $|g'(x^*)| < 1$) can be obtained in a one-dimensional discrete-time setting. Recall from (8.2) that

$$\begin{aligned}
\tilde{Q}\left(\lambda\right) &:= \beta F_{xy}\lambda^2 + \left(F_{yy}+\beta F_{xx}\right)\lambda + F_{yx}, \\
&= F_{xy}\left[\beta\lambda^2 + \underbrace{\left(\frac{F_{yy}+\beta F_{xx}}{F_{xy}}\right)}_{:=b}\lambda + 1\right],
\end{aligned}$$

[24] The argument is heuristic since the Euler Equations are satisfied only approximately for the sequence (8.3). Nevertheless, this approximate solution, by the virtue of the implicit function theorem, can be used to construct an exact solution of the Euler Equations that converges to the steady state in a neighbourhood of x^*. See RMED Chapter 6.

where we used that $F_{xy} = F_{yx}$. Assuming $F_{xy} > 0$, then $\tilde{Q}(\lambda) = 0 \Leftrightarrow Q(\lambda) := \tilde{Q}(\lambda)/F_{xy} = 0$. We solve for the latter:

$$\begin{aligned} Q(0) &= 1 > 0, \\ Q(1) &= 1 + b + \beta, \\ Q\left(\frac{1}{\beta}\right) &= \frac{1}{\beta} + b\frac{1}{\beta} + 1, \\ \frac{\partial Q(\lambda)}{\partial \lambda} &= 2\beta\lambda + b, \\ \frac{\partial^2 Q(\lambda)}{\partial \lambda^2} &= 2\beta > 0. \end{aligned}$$

Observe also that: (i) if F is concave, then $b < 0$; (ii) $Q(\lambda)$ is strictly convex.

The roots of the latter are given by

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4\beta}}{2\beta}.$$

There are two cases

▷ $1 + b + \beta < 0$

$$0 < \lambda_1 < 1 < \lambda_2,$$

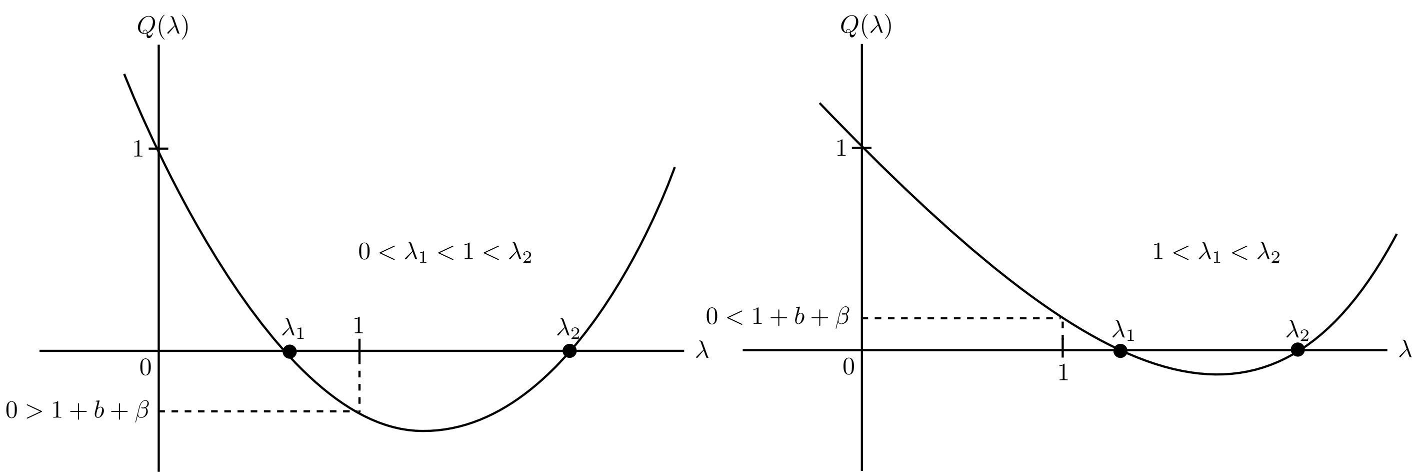
$$\begin{aligned} 0 &< \lambda_1 < 1 < \lambda_2, \\ \lambda_1 &= \frac{-b - \sqrt{b^2 - 4\beta}}{2\beta}. \end{aligned}$$

▷ $1 + b + \beta > 0$:

$$1 < \lambda_1 < \lambda_2.$$

See also the figure below.

$1 + b + \beta < 0$ $1 + b + \beta > 0$



We can also see how the λ_1 changes with the coefficient b. Taking the first case, note that

$$\begin{aligned} \frac{d\lambda_1}{db} &= \frac{-1-\frac{b}{\sqrt{b^2-4\beta}}}{2\beta} = \frac{-1-\frac{b}{\sqrt{b^2-4\beta}}}{2\beta}\frac{\sqrt{b^2-4\beta}}{\sqrt{b^2-4\beta}} \\ &= \frac{-b-\sqrt{b^2-4\beta}}{2\beta}\frac{1}{\sqrt{b^2-4\beta}} = \frac{\lambda_1}{\sqrt{b^2-4\beta}} > 0. \end{aligned}$$

8.1.4 Speed of convergence

The magnitude of $|g'(x^*)|$ describes the speed of convergence. To see this, let $\lambda_1 = g'(x^*)$ such that $|\lambda_1| < 1$.We can write (8.3) as

$$x_t = x^* + \lambda_1 (x_{t-1} - x^*).$$

Backwards substituting yields

$$\begin{aligned} x_t &= x^* + \lambda_1 ((x^* + \lambda_1 (x_{t-2} - x^*)) - x^*) \\ &= x^* + \lambda_1^2 (x_{t-2} - x^*) \\ &= \quad \vdots \\ &= x^* + \lambda_1^t (x_0 - x^*) \\ \Rightarrow x_t - x^* &= \lambda_1^t (x_0 - x^*). \end{aligned}$$

Notice that:

- ▷ $\lim_{t\to\infty} x_t - x^* = 0$ since $\lim_{t\to\infty} \lambda_1^t = 0$ given $|\lambda_1| < 1$;
- ▷ as $|\lambda_1| \to 1$, convergence is slower. The permanent income hypothesis case is one in which $g'(k^*) = 1$. In this case, the convergence speed is so slow that you, in fact, do not move.
- ▷ as $|\lambda_1| \to 0$, convergence is faster. If the cross derivatives, F_{xy} and F_{yx}, are zero (so that the roots are zero), then we would move to the steady state in the next period.

Hence, we realise that the speed of convergence is decreasing in $|\lambda_1|$, or equivalency, it is decreasing in $|g'(x^*)|$.

Exercise 8.1. Assume that $F(x,y)$ is quadratic. Show that, in this case, (8.3) gives the exact solution to the problem. Show that, in the quadratic case, the decision rules are linear.

Solution. If $F(x,y)$ is quadratic, then F_{xx}, F_{yy} and F_{xy} are constants so that linear approximation is in fact exact; i.e. we can find the entire function $g(x)$.

8.1.5 Neoclassical growth model

In the neoclassical model, we have

$$F(x,y) = U(f(x) - y)$$

so that

$$
\begin{aligned}
F_x &= f'(x)\,U'(f(x)-y) \\
F_y &= -U'(f(x)-y) \\
F_{xx} &= f''(x)\,u'(f(x)-y) + [f'(x)]^2\,U''(f(x)-y) \\
F_{yy} &= U''(f(x)-y) \\
F_{xy} &= -f'(x)\,U''(f(x)-y)\,.
\end{aligned}
$$

Recall that the steady-state capital solves $1 = \beta f'(k^*)$. Evaluating these at the steady-state values and substituting into (8.1) gives

$$
\begin{aligned}
0 &= F_{yx} + F_{yy}g' + \beta\left(F_{xx}g' + F_{xy}(g')^2\right) \\
&= -f'U'' + U''g' + \beta\left(\left(f'U' + (f')^2U''\right)g' - f'U''(g')^2\right) \\
&= -f'U'' + \left(U'' + \beta f''Y' + \beta(f')^2U''\right)g' - \beta f'U''(g')^2 \\
&= -U''\left[f' - \left(1 + \beta f''\frac{U'}{U''} + \beta(f')^2\right)g' + \beta f'(g')^2\right] \\
&= -U''\left[\frac{1}{\beta} - \left(1 + \frac{1}{\beta} + \beta f''\frac{U'}{U''}\right)g' + (g')^2\right] \because f' = 1/\beta \\
&= -U''\left[\frac{1}{\beta} - \left(1 + \frac{1}{\beta} + \frac{f''}{f'}\frac{U'}{U''}\right)g' + (g')^2\right] \because \beta = 1/f' \\
&= -U''\left[\frac{1}{\beta} - \left(1 + \frac{1}{\beta} + \left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right)g' + (g')^2\right].
\end{aligned}
$$

In this discrete-time case, we notice that the expression depends on the elasticity of productivity and elasticity of intertemporal substitution.

The quadratic equation is given by

$$
Q(\lambda) = \lambda^2 - \left(1 + \frac{1}{\beta} + \left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right)\lambda + \frac{1}{\beta}.
$$

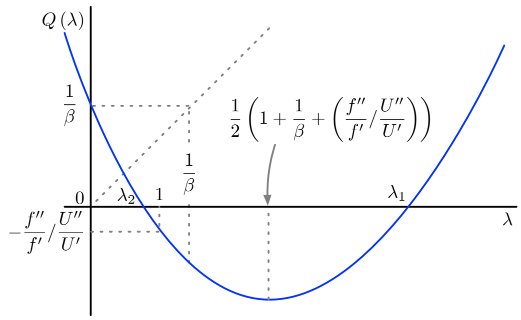
Notice that

$$
\begin{aligned}
Q(0) &= \frac{1}{\beta} > 0, \\
Q(1) &= -\frac{f''}{f'}/\frac{U''}{U'} < 0 \\
Q\left(\frac{1}{\beta}\right) &= -\left(\frac{f''}{f'}/\frac{U''}{U'}\right)\frac{1}{\beta} < 0
\end{aligned}
$$

and $Q(\lambda)$ attains a minimum at $Q'(\lambda^*) = 0$, where

$$
\lambda^* = \frac{1}{2}\left(1 + \frac{1}{\beta} + \left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right) > 1.
$$

See figure below.



Therefore, we have that

$$0 < \lambda_1 < 1 < \frac{1}{\beta} < \lambda_2.$$

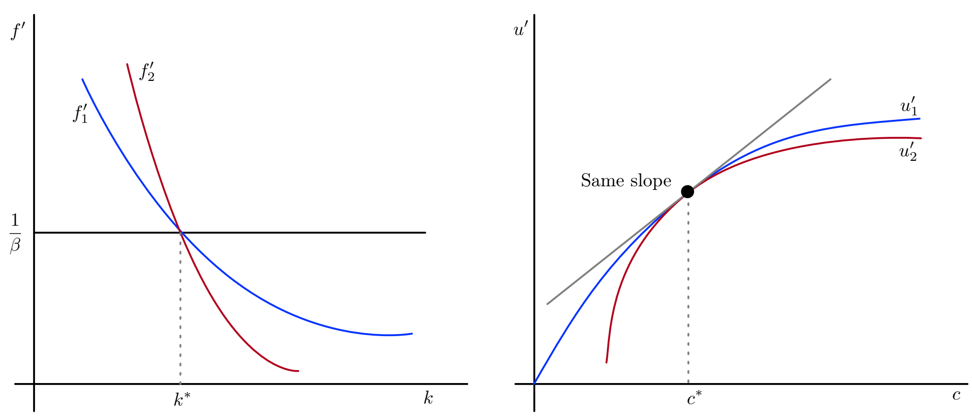
The solution is given by

$$\lambda = \frac{1}{2}\left(\left(1+\frac{1}{\beta}+\left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right) \pm \sqrt{\left(1+\frac{1}{\beta}+\left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right)^2 - \frac{4}{\beta}}\right).$$

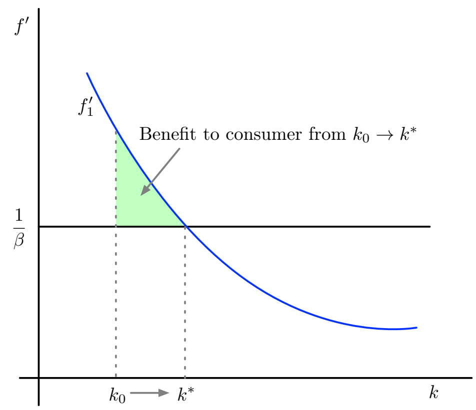
Thus, the root with absolute value less than one is given by

$$\lambda_1 = \frac{1}{2}\left(\left(1+\frac{1}{\beta}+\left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right) - \sqrt{\left(1+\frac{1}{\beta}+\left(\frac{f''}{f'}/\frac{U''}{U'}\right)\right)^2 - \frac{4}{\beta}}\right).$$

Suppose we change the production function and the utility function such that the steady states do not change. That is, we change f''/f' and U''/U' but without changing the slope at the steady state. Then the more curvature there is in U, the more agents prefer to smooth consumption. On the other hand, if u is linear, then agents do not mind consuming today or tomorrow (except for the β discount).



Suppose $k_0 < k^*$, then notice that benefit to the consumer from moving from k_0 to k^* is represented by the shaded area in the figure below (note that to increase capital stock, agents must abstain from consumption). We therefore see that a higher f'' implies a larger gain so that speed of convergence is faster. In case f' is linear, then it has to coincide with $1/\beta$ and the benefit to consumer is zero. If the production is linear but there is curvature in U, it will take forever to converge. Thus, we see that the speed of convergence depends on the "fight" between the curvature of f and u.



8.2 Stability of discrete-time linear dynamic systems of higher dimensions

Let $\mathbf{x} \in \mathbb{R}^n$ and the function $m : \mathbb{R}^n \to \mathbb{R}^n$ define a dynamic system:

$$\mathbf{x}_{t+1} = m(\mathbf{x}_t), \forall t \geq 0.$$

Let $\mathbf{x}^*$ be a steady state; i.e.

$$\mathbf{x}^* = m(\mathbf{x}^*).$$

Consider a first-order approximation of m around $\mathbf{x}^*$:

$$\mathbf{x}_{t+1} = m(\mathbf{x}^*) + m'(\mathbf{x}^*)(\mathbf{x}_t - \mathbf{x}^*).$$

Notice that this analysis is valid globally (i.e. for all $\mathbb{R}^n$) if the system is indeed linear. Alternatively, it is valid in the neighbourhood of the steady state.

Since $\mathbf{x}^* = m(\mathbf{x}^*)$, we can write

$$\begin{aligned} \mathbf{x}_{t+1} - \mathbf{x}^* &= m'(\mathbf{x}^*)(\mathbf{x}_t - \mathbf{x}^*) \\ \Rightarrow \mathbf{y}_{t+1} &= \mathbf{A}\mathbf{y}_t, \end{aligned}$$

where $\mathbf{y}_t = \mathbf{x}_t - \mathbf{x}^*$ and $\mathbf{A}$ is the matrix equal to the Jacobian $m'(\mathbf{x}^*)$. In this notation, the steady state is $\mathbf{y}^* = \mathbf{0}$.

Diagonalising the matrix $\mathbf{A}$, we obtain

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P},$$

where $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues of $\mathbf{A}$, denoted by λ_i (possibly complex), on its diagonal. As the notation already uses $\mathbf{P}$ is invertible and it contains the eigenvectors of $\mathbf{A}$. We can now write the linear system as

$$\mathbf{P}\mathbf{y}_{t+1} = \mathbf{\Lambda}\mathbf{P}\mathbf{y}_t, \ \forall t \geq 0.$$

Define $\mathbf{z}_t \coloneqq \mathbf{P}\mathbf{y}_t$, which is a linear combination of the deviations from the steady state using the eigenvectors $\mathbf{P}$. Since $\mathbf{P}$ is invertible, there is a one-to-one mapping so that each z corresponds to a unique x, and vice versa. We can then write the system as

$$\begin{aligned} \mathbf{z}_{t+1} &= \mathbf{\Lambda}\mathbf{z}_t,\ \forall t \geq 0 \\ \Rightarrow z_{i,t+1} &= \lambda_i z_{it},\ \forall i = 1, 2, \ldots, n,\ \forall t \geq 0. \end{aligned}$$

We can solve this element by element to obtain that

$$z_{i,t} = \lambda_i^t z_{i,0},\ \forall i = 1, 2, \ldots, n,\ \forall t \geq 0.$$

Example 8.1. Suppose we have a system of simultaneous first-order difference equations:

$$\begin{aligned} x_{t+1} &= 4x_t + 2y_t, \\ y_{t+1} &= -x_t + y_t, \end{aligned}$$

with initial values x_0 and y_0. We can write this in matrix form, $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t$:

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}.$$

The eigenvalues of $\mathbf{A}$ are given by k that solves $|\mathbf{A} - k\mathbf{I}| = 0$; i.e.

$$\begin{aligned} \begin{vmatrix} 4-k & 2 \\ -1 & 1-k \end{vmatrix} &= 0 \\ \Rightarrow (4-k)(1-k) + 2 &= 0 \\ \Rightarrow 6 - 5k + k^2 &= 0 \\ \Rightarrow (k-2)(k-3) &= 0. \end{aligned}$$

Thus, eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 3$. The eigenvector corresponding to λ_1 is

$$\begin{aligned} \begin{pmatrix} 4-2 & 2 \\ -1 & 1-2 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} &= 0 \\ \begin{pmatrix} 2x_t + 2y_t \\ -x_t - y_t \end{pmatrix} &= 0 \end{aligned}$$

so that the corresponding eigenvector is $(1, -1)$. For λ_2,

$$\begin{pmatrix} 4-3 & 2 \\ -1 & 1-3 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} x_t + 2y_t \\ -x_t - 2y_t \end{pmatrix} = 0$$

so that the corresponding eigenvector is $(2, -1)$. Hence

$$\begin{aligned}
\mathbf{P}^{-1} &= \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}, \\
\mathbf{\Lambda} &= \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \\
\mathbf{P} &= \frac{1}{1}\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}.
\end{aligned}$$

To verify

$$\mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}\begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ -2 & -3 \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix} = \mathbf{A}.$$

Notice that

$$\begin{aligned}
\mathbf{x}_1 &= \mathbf{A}\mathbf{x}_0, \\
\mathbf{x}_2 &= \mathbf{A}\mathbf{x}_1 \\
&= \mathbf{A}^2\mathbf{x}_0, \\
\vdots &= \vdots \\
\mathbf{x}_t &= \mathbf{A}^t\mathbf{x}_0.
\end{aligned}$$

Since $\mathbf{A} = \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}$,

$$\begin{aligned}
\mathbf{x}_1 &= \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}\mathbf{x}_0, \\
\mathbf{x}_2 &= \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}\mathbf{x}_1, \\
&= \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}\mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}\mathbf{x}_0 \\
&= \mathbf{P}^{-1}\mathbf{\Lambda}^2\mathbf{P}\mathbf{x}_0, \\
\vdots &= \vdots \\
\mathbf{x}_t &= \mathbf{P}^{-1}\mathbf{\Lambda}^t\mathbf{P}\mathbf{x}_0.
\end{aligned}$$

Define $\mathbf{z}_t = \mathbf{P}\mathbf{x}_t$, then

$$\begin{aligned}
\mathbf{z}_t &= \mathbf{\Lambda}^t\mathbf{z}_0 \\
\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} &= \begin{pmatrix} 2^t & 0 \\ 0 & 3^t \end{pmatrix}\begin{pmatrix} z_{1,0} \\ z_{2,0} \end{pmatrix}.
\end{aligned}$$

That is,

$$\begin{aligned}
z_{1,t} &= \lambda_1^t z_{1,0}, \\
z_{2,t} &= \lambda_2^t z_{2,0}.
\end{aligned}$$

The discussion above leads directly to the following important result. To simplify, let us consider the case where all the eigenvalues are real. The proposition says that, if the sequence generated by the dynamic system is to converge to the steady state, then it must be that the initial conditions $\mathbf{x}_0$ belong to a particular linear subspace. The dimension of this subspace is equal to the number of eigenvalues that are larger than one in absolute value ($n-m$ in the notation used above).

Proposition 8.2. *Let λ_i be such that, for $i = 1, 2, \ldots, m$, we have $|\lambda_i| < 1$ and for $i = m+1, m+2, \ldots, n$, we have $|\lambda_i| \geq 1$. Thus, the eigenvalues of $\mathbf{A}$ are ordered so that the first m are smaller than one. Consider the sequence*

$$\mathbf{x}_{t+1} = \mathbf{x}^* + \mathbf{A}\left(\mathbf{x}_t - \mathbf{x}^*\right), \ \forall t \geq 0$$

for some initial condition $\mathbf{x}_0$. Then,

$$\lim_{t\to\infty} \mathbf{x}_t = \mathbf{x}^*,$$

if and only if the initial condition $\mathbf{x}_0$ satisfies

$$\mathbf{x}_0 = \mathbf{P}^{-1}\hat{\mathbf{z}}_0 + \mathbf{x}^*,$$

where $\hat{\mathbf{z}}_0$ is a vector with its $n-m$ last coordinates equal to zero; i.e.

$$\hat{\mathbf{z}}_{i,0} = 0, \ \forall i = m+1, m+2, \ldots, n$$

and where the remaining elements of $\hat{\mathbf{z}}_0$ are arbitrary.

Proof. Recall that

$$z_{i,t} = \lambda_i^t z_{i,0}, \ \forall i = 1, 2, \ldots, n, \ \forall t \geq 0.$$

For any $|\lambda_i| > 1$ with positive initial value, notice that the sequence is exploding. Hence, in order for the system to converge, it must be that for all i with $|\lambda_i| > 1$, $z_{i,0} = 0$. Recall that

$$\begin{aligned}
\mathbf{z}_t = \mathbf{P}\mathbf{y}_t = \mathbf{P}\left(\mathbf{x}_t - \mathbf{x}^*\right) &\Rightarrow \mathbf{P}^{-1}\mathbf{z}_t = \mathbf{x}_t - \mathbf{x}^* \\
&\Rightarrow \mathbf{x}_t = \mathbf{P}^{-1}\mathbf{z}_t + \mathbf{x}^* \\
&\Rightarrow \mathbf{x}_0 = \mathbf{P}^{-1}\mathbf{z}_0 + \mathbf{x}^*.
\end{aligned}$$

Therefore, it must be the case that $\mathbf{x}_0 = \mathbf{P}^{-1}\hat{\mathbf{z}}_0 + \mathbf{x}^*$ for the system to converge. ■

Exercise 8.2. How should the previous statement be modified if λ_i can be complex?

Solution. We only consider the real part of the complex root and the conditions are the same with respect to the real part.

Exercise 8.3. Consider a second-order differential equation:

$$\mathbf{x}_{t+2} = \mathbf{A}_1\mathbf{x}_{t+1} + \mathbf{A}_2\mathbf{x}_t,$$

with $\mathbf{x}_t \in \mathbb{R}^n$ and with initial conditions $\mathbf{x}_0$ and $\mathbf{x}_{-1}$. Define a new variable $\mathbf{X}_t$ in $\mathbb{R}^n$ as follows:

$$\mathbf{X}_t = \begin{pmatrix} \mathbf{x}_t \\ \mathbf{x}_{t-1} \end{pmatrix}.$$

Use $\mathbf{X}_t$ to define a first-order linear difference equation that is equivalent to the previous second-order difference equation. The first-order difference equation is of the form

$$\mathbf{X}_{t+1} = \mathbf{J}\mathbf{X}_t,$$

where the $2n \times 2n$ matrix has four $n \times n$ blocks.

Solution. We can write

$$\mathbf{X}_{t+1} = \begin{pmatrix} \mathbf{x}_{t+2} \\ \mathbf{x}_{t+1} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{t+1} \\ \mathbf{x}_t \end{pmatrix} = \mathbf{J}\mathbf{X}_t.$$

8.3 Stability of continuous-time linear dynamic systems of one dimension

This section complements the derivation of the optimal decision rules for a one-dimensional continuous-time dynamic problem. We wish to characterise the optimal decision rule $\dot{k} = g(k)$ to determine the rate of change of the state as a function of the level of the state. In the counterpart to the difference equation we obtained in the discrete-time case, we obtain a differential equation for the function g in the continuous-time case.

In the continuous-time case, we consider the state representation—i.e. a differential equation with respect to the state k—as opposed to the standard representation of the Euler Equation as a (second-order) differential equation on the state at its derivatives with respect to time.

We then use g to study the local dynamics of k—i.e. dynamics of the state variable k close to the steady state value $\bar{k}$, which we summarise by $g'(\bar{k})$. We will find an algebraic (quadratic) equation for $g'(\bar{k})$ and study sufficient conditions for local stability; i.e. for $g'(\bar{k}) < 0$. We will use the neoclassical growth model as an illustration of the general principle.

Notation: We mainly use k instead of x in the subsection.

8.3.1 General continuous-time framework

We work with the continuous-time model as defined before where we choose the derivative of the state with respect to time, $\dot{x}(t)$, in each time period to maximise:

$$\begin{aligned} V^*(x_0) \coloneqq \max_{\{\dot{x}(t)\}_{t=0}^{\infty}} \quad & \lim_{T\to\infty} \int_0^T e^{-\rho t} F(x(t), \dot{x}(t))\, dt \\ s.t. \quad & \dot{x}(t) \in \Gamma(x(t)), \forall t \geq 0, \\ & x_0 \text{ given.} \end{aligned}$$

Recall that EE in the continuous time case is given by

$$\begin{aligned} & F_x(x(t), \dot{x}(t)) + \rho F_{\dot{x}}(x(t), \dot{x}(t)) \\ & = F_{\dot{x}x}(x(t), \dot{x}(t))\, \dot{x}(t) + F_{\dot{x}\dot{x}}(x(t), \dot{x}(t))\, \ddot{x}(t), \ \forall t \geq 0. \end{aligned} \tag{8.4}$$

To simplify notation, we write EE as

$$H(\ddot{x}(t), \dot{x}(t), x(t)) = 0, \ \forall t \geq 0,$$

where the function is defined in the obvious way.

The TC is given by

$$0 = \lim_{T \to \infty} e^{-\rho T} F_{\dot{x}} \left(x \left(T \right), \dot{x} \left(T \right) \right) x \left(T \right).$$

Time domain We first review the analysis of the problem with respect to time. We are looking for a path of $k(t)$—i.e. capital as a function of time—that:

(i) starts at the initial condition

$$k(0) = k_0;$$

(ii) satisfies the Euler Equations

$$H\left(\ddot{k}(t), \dot{k}(t), k(t)\right) = 0, \ \forall t > 0; \tag{8.5}$$

(iii) converges to the unique steady state (and hence satisfies transversality)

$$k(t) \to \bar{k} \text{ as } t \to \infty.$$

State domain Now we consider the analysis of the problem with respect to the state variable. From this perspective, we are looking for a function $\dot{k} = g(k)$—i.e. the rate of change of capital as a function of the level of capital—that:

(i) solves the Euler Equations (note this is not a function of time here any more):

$$H\left(g'(k) g(k), g(k), k\right) = 0, \ \forall k; \tag{8.6}$$

(ii) converges to the steady state (and hence satisfies transversality):[25]

$$g(k) > 0, \text{ if } k < \bar{k},$$
$$g(k) < 0, \text{ if } k > \bar{k}.$$

To see the equivalence between (8.5) and (8.6), differentiate $\dot{k}$ with respect to t, which gives

$$\ddot{k} = \frac{d\dot{k}}{dt} = \frac{dg(k)}{dt} = g'(k) \frac{dk}{dt} = g'(k) \dot{k} = g'(k) g(k).$$

8.3.2 Linearisation around the steady state

Using Taylor expansion on the (non-linear) law of motion for capital $\dot{k}(t) = g(k(t))$ around $k = \bar{k}$:

$$\dot{k}(t) = g(k) \simeq g\left(\bar{k}\right) + g'(k)\left(k - \bar{k}\right).$$

Since $g\left(\bar{k}\right) = 0$, we obtain the following differential equation:

$$\dot{k} = g'\left(\bar{k}\right)\left(k - \bar{k}\right).$$

[25] This condition can be extended to the m dimensional case:

$$\|g(k)\| \text{ is decreasing in } \left\|k - \bar{k}\right\|.$$

More generally, the condition requires that there exists a function $L : \mathbb{R}^m \to R_+$ with $L(k) = 0$ if and only if $k = 0$ and

$$L(g(k)) g'(k) < 0, \ \forall k.$$

8.3.3 Condition for convergence

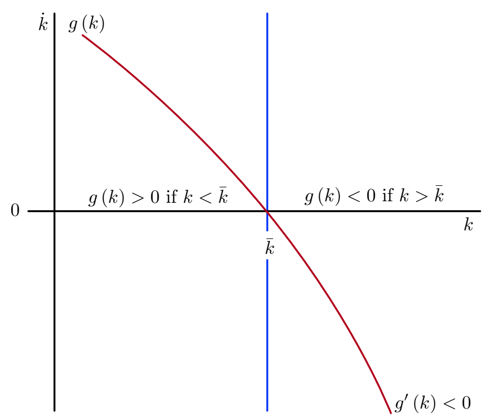
Let us focus on values of k close to the steady state $\bar{k}$. The condition for convergence are

$$g(k) > 0, \text{ if } k < \bar{k},$$
$$g(k) < 0, \text{ if } k > \bar{k},$$

which only need to hold in a neighbourhood of $\bar{k}$. The conditions above simply say that capital must be increasing when below the steady state, and decreasing if above the steady state. Given this interpretation, it follows that we can write the equivalent condition for convergence as

$$g'\left(\bar{k}\right) \equiv \frac{\partial g\left(\bar{k}\right)}{\partial k} < 0. \tag{8.7}$$

See also the figure below



8.3.4 Obtaining $g'(k)$

As in the discrete time case, we use the Euler equation (in state formation) to approximate $g'(k)$ around a steady state $\bar{k}$. Recall

$$H\left(g'(k)\, g(k), g(k), k\right) = 0, \ \forall k. \tag{8.8}$$

The steps are the same as before.

(i) Differentiate the Euler equation with respect to x. This yields a quadratic equation in $g'(x)$.

(ii) Evaluate the resulting quadratic equation at the steady state, x^*.

Differentiating (8.8) with respect to k gives

$$\left(g''(k)\, g(k) + \left(g'(k)\right)^2\right) H_{\ddot{k}} + g'(k)\, H_{\dot{k}} + H_k = 0.$$

Evaluating this expression at the steady state $k = \bar{k}$, where $g\left(\bar{k}\right) = \dot{k} = 0$ so that $g'\left(\bar{k}\right) g\left(\bar{k}\right) = \ddot{k} = 0$, the expression above simplifies to

$$H_{\ddot{k}}\left(0,0,\bar{k}\right)\left(g'\left(\bar{k}\right)\right)^2 + H_{\dot{k}}\left(0,0,\bar{k}\right) g'\left(\bar{k}\right) + H_k\left(0,0,\bar{k}\right) = 0. \tag{8.9}$$

Notice that this is a quadratic equation in $g'\left(\bar{k}\right)$.

We now derive the expressions for the coefficients of the quadratic equation. Recall that

$$H\left(\ddot{k},\dot{k},k\right)=F_{k}\left(k,\dot{k}\right)+\rho F_{\dot{k}}\left(k,\dot{k}\right)-F_{\dot{k}k}\left(k,\dot{k}\right)\dot{k}-F_{\dot{k}\dot{k}}\left(k,\dot{k}\right)\ddot{k}=0.$$

Differentiating this with respect to the three arguments gives

$$\begin{aligned}
H_{\ddot{k}}&=-F_{\dot{k}\dot{k}}\left(k,\dot{k}\right),\\
H_{\dot{k}}&=F_{k\dot{k}}\left(k,\dot{k}\right)+\rho F_{\dot{k}\dot{k}}\left(k,\dot{k}\right)-F_{\dot{k}k\dot{k}}\left(k,\dot{k}\right)\dot{k}-F_{\dot{k}k}\left(k,\dot{k}\right)-F_{\dot{k}\dot{k}\dot{k}}\left(k,\dot{k}\right)\ddot{k},\\
H_{k}&=F_{kk}\left(k,\dot{k}\right)+\rho F_{\dot{k}k}\left(k,\dot{k}\right)-F_{\dot{k}kk}\left(k,\dot{k}\right)\dot{k}-F_{\dot{k}\dot{k}k}\left(k,\dot{k}\right)\ddot{k}.
\end{aligned}$$

Evaluating the derivatives at the steady state $k=\bar{k}$ where $\dot{k}=\ddot{k}=0$, they simplify to

$$\begin{aligned}
H_{\dot{k}\dot{k}}&=-F_{\dot{k}\dot{k}}\left(\bar{k},0\right),\\
H_{\dot{k}}&=F_{k\dot{k}}\left(\bar{k},0\right)+\rho F_{\dot{k}\dot{k}}\left(\bar{k},0\right)-F_{\dot{k}k}\left(\bar{k},0\right)=\rho F_{\dot{k}\dot{k}}\left(\bar{k},0\right),\\
H_{k}&=F_{kk}\left(\bar{k},0\right)+\rho F_{\dot{k}k}\left(\bar{k},0\right).
\end{aligned}$$

Hence, we can now write (8.9) as

$$-F_{\dot{k}\dot{k}}\left(\bar{k},0\right)\left(g'\left(\bar{k}\right)\right)^{2}+\rho F_{\dot{k}\dot{k}}\left(\bar{k},0\right)g'\left(\bar{k}\right)+\left(F_{kk}\left(\bar{k},0\right)+\rho F_{\dot{k}k}\left(\bar{k},0\right)\right)=0.$$

Since the expression is quadratic, there may potentially be two candidate values for $g'\left(k\right)$.

Let $\lambda=g'\left(x^{*}\right)$, we can define the quadratic equation:

$$Q\left(\lambda\right)\coloneqq\left(-F_{\dot{k}\dot{k}}\right)\lambda^{2}+\left(\rho F_{\dot{k}\dot{k}}\right)\lambda+\left(F_{kk}+\rho F_{\dot{k}k}\right). \tag{8.10}$$

As in the discrete time case, the roots to $Q\left(\lambda\right)$ come in *almost reciprocal pairs.*[26]

Proposition 8.3. *If λ_1 solves $Q\left(\lambda_1\right)=0$, then so does $\lambda_2=-\lambda_1+\rho$.*

Proof. Suppose that λ_1 solves $Q\left(\lambda_1\right)=0$; i.e.

$$Q\left(\lambda_1\right)=\left(-F_{\dot{k}\dot{k}}\right)\lambda_1^{2}+\left(\rho F_{\dot{k}\dot{k}}\right)\lambda_1+\left(F_{kk}+\rho F_{\dot{k}k}\right)=0.$$

Consider $Q\left(-\lambda_1+\rho\right)$:

$$\begin{aligned}
Q\left(-\lambda_1+\rho\right)&=\left(-F_{\dot{k}\dot{k}}\right)\left(-\lambda_1+\rho\right)^{2}+\left(\rho F_{\dot{k}\dot{k}}\right)\left(-\lambda_1+\rho\right)+\left(F_{kk}+\rho F_{\dot{k}k}\right)\\
&=\left(-F_{\dot{k}\dot{k}}\right)\lambda_1^{2}+\left(-F_{\dot{k}\dot{k}}\right)\rho^{2}-\left(-F_{\dot{k}\dot{k}}\right)2\lambda_1\rho-\left(\rho F_{\dot{k}\dot{k}}\right)\lambda_1+\rho^{2}F_{\dot{k}\dot{k}}+\left(F_{kk}+\rho F_{\dot{k}k}\right)\\
&=\left(-F_{\dot{k}\dot{k}}\right)\lambda_1^{2}+2\rho F_{\dot{k}\dot{k}}\lambda_1-\rho F_{\dot{k}\dot{k}}\lambda_1+\left(F_{kk}+\rho F_{\dot{k}k}\right)\\
&=\left(-F_{\dot{k}\dot{k}}\right)\lambda_1^{2}+\left(\rho F_{\dot{k}\dot{k}}\right)\lambda_1+\left(F_{kk}+\rho F_{\dot{k}k}\right)\\
&=Q\left(\lambda_1\right)=0.
\end{aligned}$$

■

The theorem means that, if $\lambda_1<0$, then $\lambda_2>0$. Hence, if we find a solution

$$Q\left(g'\left(\bar{k}\right)\right)=0 \text{ with } g'\left(\bar{k}\right)<0,$$

[26] Reciprocal here means that one root is positive and the other is negative (ignoring the ρ term).

then this is *the* solution since the other root would give that $g'\left(\bar{k}\right) > 0$ which we know is not stable (recall (8.7)).

The solution satisfies the EE and TC since it converges to the steady state. The fact that it converges to the steady state also justifies the use of the approximation (i.e. linearisation) of the law of motion for capital since k stays in the neighbourhood of $\bar{k}$. It also means that there is *at most* one stable solution, which is reassuring since a convex problem should have at least one solution.

However, if $\lambda_1 > 0$ and $\lambda_2 > 0$, then the system is *not* locally stable. In this case, local arguments alone do not suffice to identify which one of the roots of Q is the solution for $g'\left(\bar{k}\right)$, but we know that one of them gives the value of $g'\left(\bar{k}\right)$.

We now analyse under what circumstances local stability (i.e. $g'\left(\bar{k}\right) < 0$) can be obtained in a one-dimensional continuous-time setting. Recall from (8.10) that

$$Q(\lambda) := \left(-F_{\dot{k}\dot{k}}\right)\lambda^2 + \left(\rho F_{\dot{k}\dot{k}}\right)\lambda + \left(F_{kk} + \rho F_{k\dot{k}}\right).$$

which can be written as

$$Q(\lambda) = \left(-F_{\dot{k}\dot{k}}\right)\left[\lambda^2 - \rho\lambda - \frac{F_{kk} + \rho F_{\dot{k}k}}{F_{\dot{k}\dot{k}}}\right]$$

so that

$$\begin{aligned}
Q(0) &= F_{\dot{k}\dot{k}}\left[\frac{F_{kk} + \rho F_{\dot{k}k}}{F_{\dot{k}\dot{k}}}\right], \\
Q(\rho) &= Q(0), \\
\frac{\partial Q(\lambda)}{\partial \lambda} &= \left(-F_{\dot{k}\dot{k}}\right)(2\lambda - \rho) < 0, \\
\frac{\partial^2 Q(\lambda)}{\partial \lambda^2} &= -2F_{\dot{k}\dot{k}} > 0,
\end{aligned}$$

where we assume that F is strictly concave so that $F_{\dot{k}\dot{k}} < 0$. These imply that $Q(\lambda)$ is strictly convex, U-shaped that has a negative slope as it crosses the y-axis. Since $Q(\rho) = Q(0)$, it attains its minimum beween $[0, \rho]$.

The solutions are:

$$\lambda = g'(k) = \frac{\rho \pm \sqrt{\rho^2 + 4\frac{-\left(F_{kk} + \rho F_{\dot{k}k}\right)}{\left(-F_{\dot{k}\dot{k}}\right)}}}{2}. \tag{8.11}$$

From (8.11), we have the following cases:

▷ $-(F_{kk} + \rho F_{\dot{k}k})/(-F_{\dot{k}\dot{k}}) > 0$:

$$\lambda_1 < 0 < \rho < \lambda_2$$

so that $g'\left(\bar{k}\right) = \lambda - 1$ is the locally stable steady state.

▷ $-(F_{kk} + \rho F_{\dot{k}k})/(-F_{\dot{k}\dot{k}}) < 0$:

$$0 < \lambda_1 < \lambda_2 < \rho,$$

which means that steady states are not locally stable.

Remark 8.2. The previous proposition also holds for higher dimensions. If the state is of dimension m, then there will be $2m$ roots satisfying

$$\lambda_{m+i} = -\lambda_i + \rho,\ \forall i = 1, 2, \ldots, m.$$

In the m dimensional case, the roots are *not* simply $\partial g_i/\partial k_i$. The roots are eigenvalues of the matrix of the derivatives of g. In the m dimensional case, we have

$$\dot{k}_i = g_i(k),\ \forall i = 1, 2, \ldots, m,$$

or, in vector notation,

$$\dot{\mathbf{k}} = \left(\dot{k}_1, \dot{k}_2, \ldots, \dot{k}_m\right) = \mathbf{g}\left(\mathbf{k}\right) = \mathbf{g}\left(k_1, k_2, \ldots, k_m\right)$$

and

$$\mathbf{g}\left(k_1, k_2, \ldots, k_m\right) = \left(g_1\left(\mathbf{k}\right), g_2\left(\mathbf{k}\right), \ldots, g_m\left(\mathbf{k}\right)\right).$$

Then, $\mathbf{g}'\left(\mathbf{k}\right)$ is the matrix

$$\frac{\partial \mathbf{g}\left(\bar{\mathbf{k}}\right)}{\partial \mathbf{k}} = \begin{bmatrix} \frac{\partial g_1}{\partial k_1} & \frac{\partial g_1}{\partial k_2} & \cdots & \frac{\partial g_1}{\partial k_m} \\ \frac{\partial g_2}{\partial k_1} & \frac{\partial g_2}{\partial k_2} & \cdots & \frac{\partial g_2}{\partial k_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial k_1} & \frac{\partial g_m}{\partial k_2} & \cdots & \frac{\partial g_m}{\partial k_m} \end{bmatrix}.$$

The eigenvalues of the matrix above control the (local) behaviour of k around $\bar{k}$.

8.3.5 Discount factor and local stability

Notice that as $\rho \to 0$,

$$\lambda = \frac{0 \pm 2\sqrt{-F_{kk}/\left(-F_{\dot{k}\dot{k}}\right)}}{2} = \pm\sqrt{F_{kk}/F_{\dot{k}\dot{k}}}.$$

Choosing the negative one:

$$\lim_{\rho \to 0} g'\left(\bar{k}\right) = -\sqrt{F_{kk}/F_{\dot{k}\dot{k}}} < 0.$$

Thus, as long as F is strictly concave, for small ρ, the steady state is globally stable.

We extend this result by finding an upper bound on ρ for which $g'\left(\bar{k}\right) < 0$. Assume that F is strictly concave so that $F_{kk}, F_{\dot{k}\dot{k}} < 0$. The smallest root λ_1 is negative if either:

- ▷ $F_{\dot{k}k} \leq 0$ (since we would have $\rho \pm \sqrt{\rho^2 + \text{positive}}$ in (8.11)); or
- ▷ if $F_{\dot{k}k} > 0$ and

$$4\frac{-\left(F_{kk} + \rho F_{\dot{k}k}\right)}{\left(-F_{\dot{k}\dot{k}}\right)} > 0 \Leftrightarrow F_{kk} + \rho F_{\dot{k}k} < 0 \tag{8.12}$$

$$\Leftrightarrow \rho < \frac{-F_{kk}}{F_{\dot{k}k}}.$$

Note that, since $F_{\dot{k}k} > 0$,

$$\rho < \frac{-F_{kk}}{F_{\dot{k}k}} = \frac{-F_{kk}}{\left|F_{\dot{k}k}\right|}.$$

Since F is constant returns to scale,[27]

$$F_{kk}F_{\dot{k}\dot{k}} = \left(F_{\dot{k}k}\right)^2.$$

To see this, recall the Euler's formula:

$$F\left(x, y\right) = xF_x + yF_y.$$

[27] This results means that the determinant of the Hessian matrix of a two-variable constant returns to scale function equals zero.

Derivative of F with respect to x and y are thus:

$$F_x = F_x + xF_{xx} + yF_{yx} \Leftrightarrow F_{xx} = -\frac{y}{x}F_{xy}$$
$$F_y = xF_{xy} + F_y + yF_{yy} \Leftrightarrow yF_{yy} = -\frac{x}{y}F_{xy}.$$

Then,

$$F_{xx}F_{yy} - F_{xy}^2 = \left(-\frac{y}{x}F_{yx}\right)\left(-\frac{x}{y}F_{xy}\right) - F_{xy}^2 = 0.$$

We can then write

$$\rho < \frac{-F_{kk}}{|F_{\dot{k}k}|} = \frac{-F_{kk}}{\sqrt{F_{kk}F_{\dot{k}\dot{k}}}} = \frac{\sqrt{-F_{kk}}\sqrt{-F_{kk}}}{\sqrt{-F_{kk}}\sqrt{-F_{\dot{k}\dot{k}}}} = \sqrt{\frac{F_{kk}}{F_{\dot{k}\dot{k}}}}.$$

Euler Equation at the steady state gives that (recall (8.4)).

$$F_k + \rho F_{\dot{k}} = 0 \Leftrightarrow F_k = -\rho F_{\dot{k}}. \tag{8.13}$$

Then,

$$\rho < \sqrt{\frac{F_{kk}}{F_{\dot{k}\dot{k}}}\frac{F_k}{F_k}} = \sqrt{\frac{F_{kk}}{F_{\dot{k}\dot{k}}}\frac{(-F_{\dot{k}})}{F_k}\rho} = \sqrt{\frac{F_{kk}/F_k}{F_{\dot{k}\dot{k}}/(-F_{\dot{k}})}\rho}$$
$$\Rightarrow \sqrt{\rho} < \sqrt{\frac{F_{kk}/F_k}{F_{\dot{k}\dot{k}}/(-F_{\dot{k}})}}$$
$$\Rightarrow \rho < \frac{-F_{kk}/F_k}{-F_{\dot{k}\dot{k}}/(-F_{\dot{k}})}.$$

Notice that this condition compares the curvature of F_k to $F_{\dot{k}}$ with the discount rate, and we see that the condition is satisfied if the marginal cost of accumulating capital, $-F_{\dot{k}}$, has a high curvature.

8.3.6 Steady state "capital stock" and the discount rate

Recall that, if $F_{\dot{k}k} > 0$, for local stability, we need (see (8.12)),

$$F_{kk} + \rho F_{\dot{k}k} < 0.$$

This is related to the sign of the comparative static of steady-state capital with respect to ρ. Define $\bar{k}(\rho)$ as solution to (8.13) so that

$$F_k\left(\bar{k}(\rho), 0\right) + \rho F_{\dot{k}}\left(\bar{k}(\rho), 0\right) = 0.$$

Differentiating with respect to ρ gives

$$F_{kk}\frac{d\bar{k}(\rho)}{d\rho} + F_{\dot{k}} + \rho F_{\dot{k}k}\frac{d\bar{k}(\rho)}{d\rho} = 0,$$

which gives us that

$$\frac{d\bar{k}(\rho)}{d\rho} = \bar{k}_\rho = \frac{(-F_{\dot{k}})}{F_{kk} + \rho F_{\dot{k}k}}.$$

Since $F_{\dot{k}} \leq 0$ (recall the proof of the sufficiency of EE and TC in continuous time), the numerator is positive. Thus, if $F_{\dot{k}k} < 0$ then $\bar{k}_\rho < 0$. Similarly, if $F_{\dot{k}k} > 0$ and $F_{kk} + \rho F_{\dot{k}k} < 0$ (the other case where the steady state is locally stable), we also have that $\bar{k}_\rho < 0$. That is, in both cases, we obtain the intuitive result that a higher discount rate (i.e. more impatience) leads to lower capital accumulation in the steady state.

8.3.7 Speed of convergence

The magnitude of $\left|g'\left(\bar{k}\right)\right|$ describes the speed of convergence. To see this, recall the linearised law of motion:

$$\dot{k} = g'\left(\bar{k}\right)\left(k - \bar{k}\right).$$

Rewriting this as

$$\frac{dk}{dt} = -g'\left(\bar{k}\right)\bar{k} + g'\left(\bar{k}\right)k,$$

which is differential equation in k.

Solving the differential equation To solve this, guess that the solution is of the form

$$k(t) = \exp\left[g'\left(\bar{k}\right)t\right]c(t).$$

Then,

$$\begin{aligned}\frac{dk(t)}{dt} &= g'\left(\bar{k}\right)\exp\left[g'\left(\bar{k}\right)t\right]c(t) + c'(t)\exp\left[g'\left(\bar{k}\right)t\right] \\ &= g'\left(\bar{k}\right)k(t) + c'(t)\exp\left[g'\left(\bar{k}\right)t\right].\end{aligned}$$

Substituting this into the expression for dk/dt,

$$\begin{aligned}-g'\left(\bar{k}\right)\bar{k} + g'\left(\bar{k}\right)k(t) &= g'\left(\bar{k}\right)k(t) + c'(t)\exp\left[g'\left(\bar{k}\right)t\right] \\ \Leftrightarrow c'(t) &= -g'\left(\bar{k}\right)\bar{k}\exp\left[-g'\left(\bar{k}\right)t\right].\end{aligned}$$

Integrating both sides with respect to t

$$\begin{aligned}\int c'(t)\,dt &= \int -g'\left(\bar{k}\right)\bar{k}\exp\left[-g'\left(\bar{k}\right)t\right]dt \\ \Rightarrow c(t) &= \bar{k}\exp\left[-g'\left(\bar{k}\right)t\right] + C.\end{aligned}$$

Hence,

$$\begin{aligned}k(t) &= \exp\left[g'\left(\bar{k}\right)t\right]\left(\bar{k}\exp\left[-g'\left(\bar{k}\right)t\right] + C\right) \\ &= \bar{k} + \exp\left[g'\left(\bar{k}\right)t\right]C.\end{aligned}$$

To pin down C, we need a boundary condition, which is that $k(0) = k_0$.

$$k_0 = \bar{k} + C \Leftrightarrow C = \bar{k} - k_0.$$

Thus, we obtain that

$$k(t) = \bar{k} + \left(\bar{k} - k_0\right) \exp\left[g'\left(\bar{k}\right) t\right].$$

Using integrating factor We can solve this first-order differential equation using an integrating factor.[28] The integrating factor is then given by

$$I(t) = \exp\left[\int^t -g'\left(\bar{k}\right) dt\right] = \exp\left[-g'\left(\bar{k}\right) t\right].$$

The solution is then given by

$$\begin{aligned}\exp\left[-g'\left(\bar{k}\right) t\right] k(t) &= \int^t -g'\left(\bar{k}\right) \bar{k} \exp\left[-g'\left(\bar{k}\right) t\right] dt + c \\ &= \bar{k} \exp\left[-g'\left(\bar{k}\right) t\right] + c.\end{aligned}$$

Initial condition is that $k(0) = k_0$. Notice that if $t = 0$, then above expression simplifies to

$$k_0 = \bar{k} + c \Rightarrow c = k_0 - \bar{k}.$$

This gives us a linearised version of the law of motion for capital:[29]

$$\begin{aligned}\exp\left[-g'\left(\bar{k}\right) t\right] k(t) &= \bar{k} \exp\left[-g'\left(\bar{k}\right) t\right] + \left(k_0 - \bar{k}\right) \\ \Rightarrow k(t) &= \bar{k} + \left(k_0 - \bar{k}\right) \exp\left[g'\left(\bar{k}\right) t\right].\end{aligned}$$

Thus, the if $g'\left(\bar{k}\right) < 0$, then $k(t) \to \bar{k}$. Moreover, the more negative $g'\left(\bar{k}\right)$, the faster is the convergence.

We can also solve the first-order differential equation,

$$\dot{k} = g'\left(\bar{k}\right)\left(k - \bar{k}\right) = g'\left(\bar{k}\right) k - g'\left(\bar{k}\right) \bar{k}$$

by the usual "guess and verify" method. Guess that the solution is given by

$$k(t) = \exp\left[g'\left(\bar{k}\right) t\right] c(t),$$

where $c(t)$ is the constant of variation. We also need the boundary condition, which in this case is the initial capital level:

[28] Recall that if we have a differential equation of the form

$$\frac{dy}{dt} + P(t) y = Q(t),$$

then we can define the integrating factor to be $I(t)$ such that $I'(t) = P(t) I(t)$. Multiplying both sides by $I(t)$ gives

$$\begin{aligned}I(t) \frac{dy}{dt} + P(t) I(t) y = Q(t) I(t) \Rightarrow \frac{dI(t) y}{dt} &= Q(t) I(t) \\ \Rightarrow I(t) y &= \int^t Q(t) I(t) dt + c,\end{aligned}$$

where c depends on the initial condition. To find $I(t)$:

$$\frac{I'(t)}{I(t)} = P(t) \Rightarrow \frac{d \ln(I(t))}{dt} = P(t) \Rightarrow \ln(I(t)) = \int^t P(t) dt \Rightarrow I(t) = \exp\left[\int^t P(t) dt\right].$$

[29] To verify this, differentiating the expression with respect to t yields

$$\dot{k}(t) = g'\left(\bar{k}\right)\left[k(0) - \bar{k}\right] \exp\left[g'\left(\bar{k}\right) t\right] = g'\left(\bar{k}\right)\left(k(t) - \bar{k}\right).$$

$k\left(0\right)=k_{0}$. Then,

$$\begin{aligned}\dot{k}&=g'\left(\bar{k}\right)\exp\left[g'\left(\bar{k}\right)t\right]c\left(t\right)+\dot{c}\left(t\right)\exp\left[g'\left(\bar{k}\right)t\right]\\&=g'\left(\bar{k}\right)k\left(t\right)+\dot{c}\left(t\right)\exp\left[g'\left(\bar{k}\right)t\right].\end{aligned}$$

Thus, we need

$$\begin{aligned}\dot{c}\left(t\right)\exp\left[g'\left(\bar{k}\right)t\right]&=-g'\left(\bar{k}\right)\bar{k}\\\Rightarrow\dot{c}\left(t\right)&=-\exp\left[-g'\left(\bar{k}\right)t\right]g'\left(\bar{k}\right)\bar{k}\\\Rightarrow c\left(t\right)=\int\dot{c}\left(t\right)dt&=\int-\exp\left[-g'\left(\bar{k}\right)t\right]g'\left(\bar{k}\right)\bar{k}dt\\&=-g'\left(\bar{k}\right)\bar{k}\int\exp\left[-g'\left(\bar{k}\right)t\right]dt\\&=-g'\left(\bar{k}\right)\bar{k}\left(-\frac{1}{g'\left(\bar{k}\right)}\exp\left[-g'\left(\bar{k}\right)t\right]+C\right)\\&=\bar{k}\exp\left[-g'\left(\bar{k}\right)t\right]--g'\left(\bar{k}\right)\bar{k}C.\end{aligned}$$

We therefore have that

$$\begin{aligned}k\left(t\right)&=\exp\left[g'\left(\bar{k}\right)t\right]\left(\bar{k}\exp\left[-g'\left(\bar{k}\right)t\right]-g'\left(\bar{k}\right)\bar{k}C\right)\\&=\bar{k}-\exp\left[g'\left(\bar{k}\right)t\right]g'\left(\bar{k}\right)\bar{k}C.\end{aligned}$$

We solve for C using the boundary condition.

$$\begin{aligned}k_{0}&=\bar{k}-\exp\left[g'\left(\bar{k}\right)\cdot0\right]g'\left(\bar{k}\right)\bar{k}C\\&=\bar{k}-g'\left(\bar{k}\right)\bar{k}C\\\Rightarrow C&=\frac{\bar{k}-k_{0}}{g'\left(\bar{k}\right)\bar{k}}.\end{aligned}$$

Hence,

$$\begin{aligned}k\left(t\right)&=\bar{k}-\exp\left[g'\left(\bar{k}\right)t\right]g'\left(\bar{k}\right)\bar{k}\frac{\bar{k}-k_{0}}{g'\left(\bar{k}\right)\bar{k}}\\&=\bar{k}+\left(k_{0}-\bar{k}\right)\exp\left[g'\left(\bar{k}\right)t\right]\end{aligned}$$

as we had before.

Remark 8.3. (*Half life*) Let us define τ as the time τ that it takes so that the system reaches close to the “half” of the difference between the initial point and the steady state.

$$k\left(\tau\right)-\bar{k}=\frac{1}{2}\left(k_{0}-\bar{k}\right).$$

Substituting for $k(\tau)$ the linearised law of motion for capital yields

$$\begin{aligned}\bar{k} + \left(k_0 - \bar{k}\right)\exp\left[g'\left(\bar{k}\right)\tau\right] - \bar{k} &= \frac{1}{2}\left(k_0 - \bar{k}\right) \\ \Rightarrow \exp\left[g'\left(\bar{k}\right)\tau\right] &= \frac{1}{2} \\ \Rightarrow \tau &= -\frac{\ln 2}{g'\left(\bar{k}\right)}.\end{aligned}$$

8.3.8 Neoclassical growth model

Recall that

$$F\left(k, \dot{k}\right) = U\left(f(k) - \delta k - \dot{k}\right)$$

so that

$$\begin{aligned}F_k\left(k, \dot{k}\right) &= \left(f'(k) - \delta\right)U', \\ F_{\dot{k}}\left(k, \dot{k}\right) &= -U', \\ F_{\dot{k}k}\left(k, \dot{k}\right) &= -\left(f'(k) - \delta\right)U'', \\ F_{\dot{k}\dot{k}}\left(k, \dot{k}\right) &= U'',\end{aligned}$$

where we add one extra derivative

$$F_{kk} = f''(k)\,U' + \left(f'(k) - \delta\right)^2 U''.$$

Evaluating them at the steady state $k = \bar{k}$ using the fact that $\rho = f'\left(\bar{k}\right) - \delta$ gives

$$\begin{aligned}F_k\left(\bar{k}, 0\right) &= \left(f'\left(\bar{k}\right) - \delta\right)U' = \rho U', \\ F_{\dot{k}}\left(\bar{k}, 0\right) &= -U', \\ F_{\dot{k}k}\left(\bar{k}, 0\right) &= -\rho U'', \\ F_{\dot{k}\dot{k}}\left(\bar{k}, 0\right) &= U'', \\ F_{kk}\left(\bar{k}, 0\right) &= f''\left(\bar{k}\right)U' + \rho^2 U''.\end{aligned}$$

The quadratic equation, $Q(\lambda) := \left(-F_{\dot{k}\dot{k}}\right)\lambda^2 + \left(\rho F_{\dot{k}\dot{k}}\right)\lambda + \left(F_{kk} + \rho F_{\dot{k}k}\right)$, then becomes

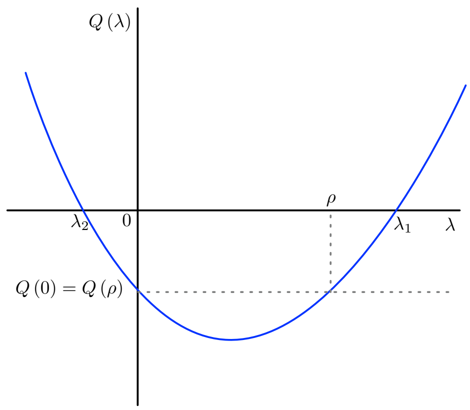
$$\begin{aligned}Q(\lambda) &= \left(-U''\right)\lambda^2 + \left(\rho U''\right)\lambda + \left(f''U' + \rho^2 U'' - \rho^2 U''\right) \\ &= \left(-U''\right)\lambda^2 + \left(\rho U''\right)\lambda + \left(f''U'\right) \\ &= \left(-U''\right)\left[\lambda^2 - \rho\lambda - \frac{U'}{U''}f''\right] \\ &= \left(-U''\right)\left[\lambda^2 - \rho\lambda - \frac{U'}{U''}f''\left(\frac{f'}{f'}\frac{c}{c}\frac{k}{k}\right)\right] \\ &= \left(-U''\right)\left[\lambda^2 - \rho\lambda - \left(\frac{cf'}{k}\right)\left(-\frac{U''}{U'}c\right)^{-1}\left(-\frac{f''}{f'}k\right)\right].\end{aligned}$$

Notice that $-\frac{U''}{U'}c$ is the elasticity of marginal utility (or, the reciprocal of the elasticity of intertemporal substitution) and $-\frac{f''}{f'}k$ is the elasticity of marginal productivity.

Now we check that one root is negative, say λ_2, and the other is then positive and larger than ρ. To do this, note that,

$$\begin{aligned}
Q(0) &= (-U'')\left[-\left(\frac{cf'}{k}\right)\left(-\frac{U''}{U'}c\right)^{-1}\left(-\frac{f''}{f'}k\right)\right] < 0 \because U'' < 0, \\
Q(\rho) &= Q(0), \\
\frac{\partial Q(\lambda)}{\partial \lambda} &= -U''(2\lambda - \rho), \\
\frac{\partial^2 Q(\lambda)}{\partial \lambda^2} &= -2U'' > 0.
\end{aligned}$$

Figure below plots $Q(\lambda)$, which visually shows that one root is negative and the other is positive and larger than ρ.



We can, in fact, solve for the roots using the quadratic formula:

$$\lambda = g'(k) = \frac{\rho \pm \sqrt{\rho^2 - 4\left(\frac{cf'}{k}\right)\left(-\frac{U''}{U'}c\right)^{-1}\left(-\frac{f''}{f'}k\right)}}{2},$$

and the negative is root is given by when $\pm$ is $-$.

Example 8.2. Let us impose some functional forms for the production function and the period utility:

$$\begin{aligned}
f(k) &= Ak^\alpha, \\
u(c) &= \frac{c^{1-\gamma} - 1}{1-\gamma}.
\end{aligned}$$

The steady steady equations are:

$$\begin{aligned}
f'(\bar{k}) = \rho + \delta \Rightarrow \alpha A\bar{k}^{\alpha-1} &= \rho + \delta. \\
\Rightarrow A\bar{k}^\alpha &= \frac{\rho + \delta}{\alpha}\bar{k}
\end{aligned}$$

We can obtain steady-state consumption from the law of motion for capital ($\dot{k}_t = G(k_t, 1) - \delta k_t - c_t$):

$$\bar{c} = A\bar{k}^\alpha - \delta\bar{k}.$$

Combining the two expressions, we can write

$$\begin{aligned}\bar{c} &= \frac{(\rho+\delta)}{\alpha}\bar{k} - \delta\bar{k} \\ &= \left(\frac{\rho + (1-\alpha)\,\delta}{\alpha}\right)\bar{k}.\end{aligned}$$

The elasticities are

$$\begin{aligned}-\frac{f''}{f'}k &= -\frac{(\alpha-1)\,\alpha A\bar{k}^{\alpha-2}}{\alpha A\bar{k}^{\alpha-1}}k = (1-\alpha)\,, \\ -\frac{y''}{u'}c &= -\frac{-\gamma c^{-\gamma-1}}{c^{-\gamma}}c = \gamma.\end{aligned}$$

Therefore,

$$\left(\frac{cf'}{k}\right)\left(-\frac{U''}{U'}c\right)^{-1}\left(-\frac{f''}{f'}k\right) = \left(\frac{\rho+(1-\alpha)\,\delta}{\alpha}\right)(\rho+\delta)\,\frac{1-\alpha}{\gamma} \geq 0.$$

Hence, indeed,

$$Q\,(0) = Q\,(\rho) = (-U'')\left[-\left(\frac{\rho+(1-\alpha)\,\delta}{\alpha}\right)(\rho+\delta)\,\frac{1-\alpha}{\gamma}\right] < 0$$

so that one root is negative.

8.4 Stability of continuous-time linear dynamic systems of higher dimensions

In continuous time and with higher dimensions, we would have

$$\dot{\mathbf{x}}\,(t) = m\,(\mathbf{x}\,(t))\,,\ \forall t \geq 0.$$

At the steady state,

$$0 = m\,(x^*)\,.$$

Linear approximation of $\dot{\mathbf{x}}\,(t)$ yields

$$\begin{aligned}\dot{\mathbf{x}}\,(t) = m\,(\mathbf{x}\,(t)) &\simeq m\,(\mathbf{x}^*) + m'\,(\mathbf{x}^*)\,(\mathbf{x}\,(t) - \mathbf{x}^*) \\ &\simeq m'\,(\mathbf{x}^*)\,(\mathbf{x}\,(t) - \mathbf{x}^*)\,.\end{aligned}$$

Define $\mathbf{A} \coloneqq m'\,(\mathbf{x}^*)$ and $\mathbf{y}\,(t) \coloneqq \mathbf{x}\,(t) - \mathbf{x}^*$ so that we may write

$$\dot{\mathbf{y}}\,(t) = \mathbf{A}\mathbf{y}\,(t)\,.$$

As before, we can diagonalise $\mathbf{A}$ as $\mathbf{A} = \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}$ and define $\mathbf{z}\,(t) \coloneqq \mathbf{P}\mathbf{y}\,(t)$ to obtain

$$\dot{\mathbf{z}}\,(t) = \mathbf{\Lambda}\mathbf{z}\,(t)\,,\ \forall t \geq 0,$$

which can be written as

$$\dot{z}_i\,(t) = \lambda_i z_i\,(t)\,,\ \forall i = 1, 2\ldots, n,\ \forall t \geq 0,.$$

This has the solution:[30]

$$z_i(t) = e^{\lambda_i t} z_i(0), \ \forall i = 1, 2 \ldots, n, \ \forall t \geq 0.$$

Proposition 8.4. *Let λ_i be such that for $i = 1, 2, \ldots, m$, we have $\lambda_i < 0$ and for $i = m+1, m+2, \ldots, n$, we have $\lambda_i \geq 0$. Thus, the eigenvalues of A are ordered so that the first m are negative. Consider the sequence*

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x}(t) - \mathbf{x}^*), \ \forall t \geq 0$$

for some initial condition $\mathbf{x}(0)$. Then

$$\lim_{t \to \infty} \mathbf{x}(t) = \mathbf{x}^*$$

if and only if the initial condition $\mathbf{x}(0)$ satisfies

$$\mathbf{x}(0) = \mathbf{P}^{-1} \hat{\mathbf{z}}(0) + \mathbf{x}^*,$$

where $\hat{\mathbf{z}}(0)$ is a vector with its $n - m$ last coordinates equal to zero; i.e.

$$\hat{z}_i(0) = 0, \ \forall i = m+1, m+2, \ldots, n$$

and where the remaining elements of $\hat{\mathbf{z}}(0)$ are arbitrary.

Proof. Recall that

$$z_i(t) = e^{\lambda_i t} z_i(0), \ \forall i = 1, 2 \ldots, n, \ \forall t \geq 0.$$

For any $\lambda_i > 0$ with positive initial value, the sequence is exploding. Hence, in order for the system to converge, it must be that, for all i with $\lambda_i > 0$, $z_i(0) = 0$. Recall that

$$\begin{aligned} \mathbf{z}(t) = \mathbf{P}\mathbf{y}(t) = \mathbf{P}(\mathbf{x}(t) - \mathbf{x}^*) &\Rightarrow \mathbf{P}^{-1}\mathbf{z}(t) = (\mathbf{x}(t) - \mathbf{x}^*) \\ &\Rightarrow \mathbf{x}(t) = \mathbf{P}^{-1}\mathbf{z}(t) + \mathbf{x}^* \\ &\Rightarrow \mathbf{x}(0) = \mathbf{P}^{-1}\mathbf{z}(0) + \mathbf{x}^*. \end{aligned}$$

Therefore, it must be the case that $\mathbf{x}(0) = \mathbf{P}^{-1}\hat{\mathbf{z}}(0) + \mathbf{x}^*$ for the system to converge. ■

Exercise 8.4. How should the previous statement be modified if λ_i can be complex?

Solution. We only consider the real part of the complex root and the conditions are the same with respect to the real part.

Exercise 8.5. Consider the following second-order linear differential equation:

$$\ddot{\mathbf{x}}(t) = \mathbf{A}_1 \dot{\mathbf{x}}(t) + \mathbf{A}_0 \mathbf{x}(t),$$

[30] To see this, rearrange

$$\begin{aligned} \lambda_i &= \frac{\dot{z}_i(t)}{z_i(t)} = \frac{d}{dt}[\ln z_i(t)] \\ \Rightarrow \int \lambda_i dt &= \lambda_i t = \ln z_i(t) + C \\ &\Rightarrow z_i(t) = e^{\lambda_i t} e^{-C}. \end{aligned}$$

Given the initial condition $z_i(0)$, $e^{-C} = z_i(0)$.

for $\mathbf{x}(t) \in \mathbb{R}^n$. Write this as a first-order differential equation in $\mathbf{X}(t) \in \mathbb{R}^{2n}$ with

$$\mathbf{X}(t) = \begin{pmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{x}(t) \end{pmatrix}$$

for $\dot{\mathbf{X}}(t) = \mathbf{J}\mathbf{X}(t)$. Given an expression for each of the four blocks of the $2n \times 2n$ matrix $\mathbf{J}$.

Solution. As before,

$$\dot{\mathbf{X}}(t) = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{I} & 0 \end{pmatrix} \mathbf{X}(t).$$

8.5 Saddle path for linearised dynamics

Consider a continuous-time problem with state x, controls u, objective function h, law of motion g and discount factor ρ as done in the notes before. The Hamiltonian is

$$H(x, u, \lambda) = h(x, u) + \lambda g(x, u),$$

where λ is the co-state variable. Let $x, u, \lambda \in \mathbb{R}$. The first-order conditions for the Hamiltonian are:

$$\begin{aligned} 0 &= H_u(x, u, \lambda), \\ \dot{\lambda} &= \rho\lambda - H_x(x, u, \lambda), \\ \dot{x} &= g(x, u). \end{aligned}$$

Using $H_u = 0$ to solve for $u = \mu(x, \lambda)$ as a function of x and λ, we get

$$H_u(x, \mu(x, \lambda), \lambda) = 0.$$

We then obtain the two dimensional dynamic system:

$$\begin{aligned} \dot{\lambda} &= \rho\lambda - H_x(x, \mu(x, \lambda), \lambda), \\ \dot{x} &= g(x, \mu(x, \lambda)). \end{aligned}$$

Linearising this system, we obtain

$$\begin{pmatrix} \frac{d\hat{\lambda}}{dt} \\ \frac{d\hat{x}}{dt} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \hat{\lambda}(t) \\ \hat{x}(t) \end{pmatrix},$$

where $\hat{\lambda}(t) = \lambda(t) - \lambda^*$ and $\hat{x}(t) = x(t) - x^*$.

The matrix $\mathbf{A} = [a_{ij}]$ has the derivatives of the two-dimensional system displayed above evaluated at the steady-state values. Using the previous result, we can write

$$\begin{pmatrix} \hat{\lambda}(t) \\ \hat{x}(t) \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} z_1(0)\, e^{\theta_1 t} \\ z_2(0)\, e^{\theta_2 t} \end{pmatrix},$$

where

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{\Theta}\mathbf{P}$$

and $\mathbf{\Theta}$ is a diagonal matrix with the eigenvalues of $\mathbf{A}$, denoted by θ_i in its diagonal.

Let us assume that $\theta_1 < 0 < \theta_2$. For the system to converge to the steady state, we must have $z_2(0) = z_2(t) = 0$ for all t, where z is given by

$$\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \mathbf{P} \begin{pmatrix} \hat{\lambda}(t) \\ \hat{x}(t) \end{pmatrix}.$$

Then,

$$\begin{pmatrix} z_1(0)\, e^{\theta_1 t} \\ 0 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} \hat{\lambda}(t) \\ \hat{x}(t) \end{pmatrix}, \forall t \geq 0$$

or,

$$p_{21}\hat{\lambda}(t) + p_{22}\hat{x}(t) = 0.$$

This defines the saddle path as

$$\begin{aligned} \hat{\lambda}(t) &= -\frac{p_{22}}{p_{21}}\hat{x}(t), \\ \Rightarrow \lambda(t) &= \lambda^* - \frac{p_{22}}{p_{21}}(x(t) - x^*). \end{aligned}$$

8.6 Slope of the saddle path and of the optimal decision rule (continuous-time, one-dimensional case)

Let the function ϕ satisfying $\lambda = \phi(x)$ be the saddle path. We want to compute

$$\frac{d\lambda}{dx} = \phi'(x^*),$$

where x^* is the steady state. Notice that

$$\phi' = \frac{d\lambda}{dx} = \frac{d\lambda/dt}{dx/dt} \equiv \frac{\dot{\lambda}(\lambda, x)}{\dot{x}(\lambda, x)} = \frac{\rho\lambda - H_x(x, \mu(x,\lambda), \lambda)}{g(x, \mu(x,\lambda))}$$

At the steady state, $\dot{\lambda} = \dot{x} = 0$ which implies that $\phi' =$ ”0/0”. Thus, we use L'Hôpital's rule to find this ratio.

$$\begin{aligned} \lim_{x \to x^*} \phi'(x) \equiv \left.\frac{d\lambda}{dx}\right|_{x=x^*} &= \left.\frac{\dot{\lambda}(\lambda, x)}{\dot{x}(\lambda, x)}\right|_{x=x^*} \\ &= \left.\frac{\frac{d}{dx}\dot{\lambda}(\lambda, x)}{\frac{d}{dx}\dot{x}(\lambda, x)}\right|_{x=x^*} = \left.\frac{\frac{d\dot{\lambda}}{d\lambda}\frac{d\lambda}{dx} + \frac{d\dot{\lambda}}{dx}}{\frac{d\dot{x}}{d\lambda}\frac{d\lambda}{dx} + \frac{d\dot{x}}{dx}}\right|_{x=x^*} \\ &= \left.\frac{\frac{d\dot{\lambda}}{d\lambda}\phi' + \frac{d\dot{\lambda}}{dx}}{\frac{d\dot{x}}{d\lambda}\phi' + \frac{d\dot{x}}{dx}}\right|_{x=x^*} \end{aligned}$$

where

$$\begin{aligned}
\frac{d\dot{\lambda}}{d\lambda} &= \rho - H_{xu}\mu_\lambda - H_{x\lambda}, \\
\frac{d\dot{\lambda}}{dx} &= -H_{xx} - H_{xu}\mu_x, \\
\frac{d\dot{x}}{d\lambda} &= g_u\mu_\lambda, \\
\frac{d\dot{x}}{dx} &= g_x + g_u\mu_x, \\
0 &= H_{ux} + H_{uu}\mu_x, \\
0 &= H_{u\lambda} + H_{uu}\mu_\lambda.
\end{aligned}$$

Note that the above comes from differentiating the first-order conditions:

$$\begin{aligned}
\dot{\lambda} &= \rho\lambda - H_x\left(x, \mu\left(x, \lambda\right), \lambda\right), \\
\dot{x} &= g\left(x, \mu\left(x, \lambda\right)\right), \\
0 &= H_u\left(x, \mu\left(x, \lambda\right), \lambda\right).
\end{aligned}$$

We therefore have the following quadratic form for ϕ':

$$\begin{aligned}
\phi'\left(\frac{d\dot{x}}{d\lambda}\phi' + \frac{d\dot{x}}{dx}\right) &= \frac{d\dot{\lambda}}{d\lambda}\phi' + \frac{d\dot{\lambda}}{dx} \\
\Rightarrow \left(\phi'\right)^2\frac{d\dot{x}}{d\lambda} + \phi\left(\frac{d\dot{x}}{dx} - \frac{d\dot{\lambda}}{d\lambda}\right) - \frac{d\dot{\lambda}}{dx} &= 0.
\end{aligned}$$

One of the two roots of this quadratic equation is the slope of the saddle path. Unlike when we were analysing convergence, there is no "rule" as to which root will represent the stable steady state. Thus, when using this method, we would usually plot the phase diagram to know the slope of the saddle path from the diagram, and pick the appropriate root using the equation above accordingly.

We can also find the slope of the optimal control rule setting the control as a function of the state. Let

$$u^*\left(x\right) = \mu\left(x, \phi\left(x\right)\right).$$

Then

$$\left.\frac{du^*\left(x\right)}{dx}\right|_{x=x^*} = \mu_x + \mu_\lambda\phi'.$$

Finally, we can find the rate of change of the state as a function of the state (i.e. the speed of convergence, $g'\left(x\right)$):

$$\begin{aligned}
\left.\frac{d\dot{x}}{dx}\right|_{x=x^*} &= g_x + g_u\left.\frac{du^*\left(x^*\right)}{dx}\right|_{x=x^*} \\
&= g_x + g_u\left(\mu_x + \mu_\lambda\phi'\right).
\end{aligned}$$

Example 8.3. (*Neoclassical growth model*) Recall that

$$\begin{aligned}
\dot{\lambda} &= \lambda\left(\rho - \left(f'\left(k\right) - \delta\right)\right) \\
\dot{k} &= f\left(k\right) - \delta k - c\left(k\right).
\end{aligned}$$

Since $U'(c) = \lambda$, we can write

$$\dot{k} = f(k) - \delta k - (U')^{-1}(\lambda).$$

Then

$$\phi'(k) \equiv \frac{d\lambda}{dk} = \frac{\frac{d\dot{\lambda}}{d\lambda}\phi' + \frac{d\dot{\lambda}}{dk}}{\frac{d\dot{k}}{d\lambda}\phi' + \frac{d\dot{k}}{dk}},$$

where

$$\begin{aligned}
\frac{d\dot{\lambda}}{d\lambda} &= \rho - (f'(k) - \delta) \\
\frac{d\dot{\lambda}}{dk} &= \frac{d\lambda}{dk}(\rho - f'(k)) - \lambda f''(k) = \phi'(k)(\rho - f'(k)) - \lambda f''(k) \\
\frac{d\dot{k}}{d\lambda} &= -\frac{d(U')^{-1}}{d\lambda} = -\frac{d(U')^{-1}(\lambda)}{d\lambda} \\
\frac{d\dot{k}}{dk} &= f''(k) - \delta - \frac{d(U')^{-1}(\lambda)}{d\lambda}\frac{d\lambda}{dk} = f''(k) - \delta - \frac{d(U')^{-1}(\lambda)}{d\lambda}\phi'(k).
\end{aligned}$$

Hence,

$$\phi'(\bar{k}) = \frac{(\rho - f'(\bar{k}) - \delta)\phi'(\bar{k}) + \phi'(\bar{k})(\rho - f'(\bar{k})) - \lambda f''(\bar{k})}{-\frac{d(U')^{-1}(\lambda)}{d\lambda}\phi'(\bar{k}) + f''(\bar{k}) - \delta - \frac{d(U')^{-1}(\lambda)}{d\lambda}\phi'(\bar{k})}.$$

In the steady state, $\rho = f'(\bar{k}) - \delta$, so we can simplify above to

$$\phi'(\bar{k}) = \frac{\delta\phi'(\bar{k}) + \lambda f''(\bar{k})}{2\frac{d(U')^{-1}(\lambda)}{d\lambda}\phi'(\bar{k}) - f''(\bar{k}) + \delta}$$

and so

$$2\frac{d(U')^{-1}(\lambda)}{d\lambda}(\phi'(\bar{k}))^2 - (f''(\bar{k}) + \delta)\phi'(\bar{k}) + (\delta - \lambda f''(\bar{k})) = 0$$

This is a quadratic equation in $\phi'(k)$ which we can solves to obtain the slope of the saddle path in (k, λ) space.

$$\phi'(\bar{k}) = \frac{f''(\bar{k}) + \delta \pm \sqrt{(f''(\bar{k}) + \delta)^2 - 8\frac{d(U')^{-1}(\lambda)}{d\lambda}(\delta - \lambda f''(\bar{k}))}}{4\frac{d(U')^{-1}(\lambda)}{d\lambda}}.$$

Recall from section 7.4.4 when we drew the phase diagram that the slope of the saddle path is negative in (k, λ) space. Since $U'' < 0$, it follows that $d(U')^{-1}(\lambda)/d\lambda < 0$ and so the negative root is given by when $\pm$ is $+$.[31] That is,

$$\phi'(\bar{k}) = \frac{f''(\bar{k}) + \delta + \sqrt{(f''(\bar{k}) + \delta)^2 + 8\frac{d(U')^{-1}(\lambda)}{d\lambda}\lambda f''(\bar{k})}}{4\frac{d(U')^{-1}(\lambda)}{d\lambda}}.$$

[31] For example, let $U(c) = \ln c$. Then $U'(c) = 1/c$ so that

$$1/c = \lambda \Leftrightarrow \lambda = \frac{1}{c};$$

i.e. $(U')^{-1}(\lambda) = 1/\lambda$. Then, $d(U')^{-1}(\lambda)/d\lambda = -1/\lambda^2 < 0$.

We can also obtain the slope of the saddle path in (k,c) space. Recall that

$$\begin{aligned}\frac{\dot{c}}{c} &= \frac{1}{-\frac{U''(c)c}{U'(c)}}\left(f'(k)-\delta-\rho\right),\\ \dot{k} &= f(k)-\delta k - c.\end{aligned}$$

So the slope of the saddle path is

$$\tilde{\phi}'\left(\bar{k}\right) = \left.\frac{dc}{dk}\right|_{k=\bar{k}} = \left.\frac{\dot{c}}{\dot{k}}\right|_{k=\bar{k}} = \frac{\frac{1}{-\frac{u''}{u'}}\left(f'(k)-\delta-\rho\right)}{f(k)-\delta k-\phi(k)}.$$

Using L'Hôpital's rule and evaluating at the steady state

$$\tilde{\phi}'\left(\bar{k}\right) = \frac{\frac{1}{-\frac{u''}{u'}}\left(f''\left(\bar{k}\right)\right) + \overbrace{\left(f'\left(\bar{k}\right)-\delta-\rho\right)}^{=0}\frac{d}{dk}\left(\frac{1}{-\frac{u''}{u'}}\right)}{\underbrace{f'(k)-\delta}_{=\rho}-\phi'\left(\bar{k}\right)}.$$

Hence,

$$\left(\rho-\tilde{\phi}'\left(\bar{k}\right)\right)\tilde{\phi}'\left(\bar{k}\right) = \frac{f''\left(\bar{k}\right)}{-\frac{u''}{u'}}\frac{-\left(f'\left(\bar{k}\right)-\delta\right)}{-\left(f'\left(\bar{k}\right)-\delta\right)} = -\frac{\frac{-f''}{f'-\delta}}{-\frac{u''}{u'}}\rho \equiv -b$$

where we used the fact that $f'=\rho$ in the steady state and $b>0$. The quadratic equation is then

$$\tilde{\phi}'\left(\bar{k}\right)^2 - \rho\tilde{\phi}\left(\bar{k}\right) - b = 0.$$

Letting $\gamma \coloneqq \tilde{\phi}'\left(\bar{k}\right)$,

$$Q(\gamma) \coloneqq \gamma^2 - \rho\gamma - b.$$

Observe that

$$\begin{aligned}Q(0) &= -b < 0,\\ Q(\rho) &= -b < 0\\ \left.\frac{dQ}{d\gamma}\right|_{\gamma=0} &= -\rho < 0,\\ \frac{d^2Q}{d\gamma^2} &= 2 > 0.\end{aligned}$$

Hence, $Q(\gamma)$ has a minimum between $\gamma \in (0,\rho)$, is strictly convex and is downward sloping when $\gamma = 0$. So, the figure looks like the one we drew in section 8.3.8. We realise that one root is positive while the other is negative. Recall from section 7.4.4 that the slope of the saddle path is positive in (k,c) space. So,

$$\tilde{\phi}'\left(\bar{k}\right) = \frac{\rho+\sqrt{\rho^2+4b}}{2}$$

Since $U'(c(k)) = \lambda(k)$, we can write

$$c^*(k) = \left(U'\right)^{-1}(\lambda(k))$$

so that

$$\tilde{\phi}'\left(\bar{k}\right) = \left.\frac{dc^*\left(k\right)}{dk}\right|_{k=\bar{k}} = \left.\frac{d\left(U'\right)^{-1}\left(\lambda\left(k\right)\right)}{dk}\frac{d\lambda}{dk}\right|_{k=\bar{k}} = \frac{d\left(U'\right)^{-1}\left(\lambda\left(\bar{k}\right)\right)}{dk}\phi'\left(\bar{k}\right).$$

9 Principle of Optimality and Dynamic Programming

Bellman's Principle of Optimality provides conditions under which a programming problem expressed in sequence form is equivalent (as defined below) to a two-period recursive programming problem (called the Bellman equation). The first part of the note discusses this relation and introduces some concepts and techniques used to solve dynamic programming problems in discrete time. The last part of the note focuses on continuous-time dynamic programming problems and shows how they relate to the Maximum Principle.

9.1 Principle of Optimality

9.1.1 Recursive problem

Recall that the sequence problem is given by:

$$\begin{aligned} V^*(x_0) \coloneqq \max_{\{x_{t+1}\}_{t=0}^{\infty}} & \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \\ s.t. \quad & x_{t+1} \in \Gamma(x_t), \ \forall t \geq 0, \\ & x_0 \text{ given.} \end{aligned}$$

Equivalently, we can write the problem as

$$V^*(x_0) = \max_{x^\infty \in \Pi^\infty(x_0)} u(\{x_{t+1}\}), \tag{9.1}$$

where

$$\begin{aligned} x^t &\coloneqq (x_0, x_1, \ldots, x_t), \\ \Pi^t(x_0) &\coloneqq \left\{x^t : x_{s+1} \in \Gamma(x_s), s = 0, 1, \ldots, t-1\right\}, \\ u(\{x_{t+1}\}) &= \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t F(x_t, x_{t+1}), \end{aligned}$$

x^t is the set of states from period 0 to t, $\Pi^t(x_0)$ is the set of all possible states from period 0 to period t (thus, $\Pi^t(x_0)$ is an increasing set as t increases, i.e. $\Pi^t(x_0) \subseteq \Pi^{t+1}(x_0)$ for all t), and $u(\{x_{t+1}\})$ is the discounted (infinite) sum of the utility.

9.1.2 Bellman equation

The Bellman equation, which is a recursive problem, is to find a function $V : X \to \mathbb{R}$ such that

$$V(x) = \max_{y\in\Gamma(x)} \left[F(x, y) + \beta V(y)\right], \ \forall x \in X. \tag{9.2}$$

This is a *functional* equation in which the solution is a function V that must satisfy the equation above for all $x \in X$. Let $g(x)$ denote the maximiser of the right-hand side of (9.2). Then, $g(x)$ satisfies

$$V(x) = F(x, g(x)) + \beta V(g(x)).$$

If the function V were known (or if we know properties of it), then (9.2) is a two-period problem.

9.1.3 Principle of Optimality

The principle of optimality states that

$$V^*(x) = V(x), \ \forall x \in X.$$

That is, the solution to the two-period problem in (9.2) is equivalent to the infinite-dimension problem in (9.1). Once we have V and g, then we have the solution of (9.1) for any initial condition $x_0 \in X$.

Here, we sketch the basic reasoning behind the Principle of Optimality (this is not a rigorous proof!). Take the case where F is bounded so $|F(x,y)| \le B < \infty$ for all $(y,x) \in \mathrm{Gr}(\Gamma)$. Notice that, for any $x^\infty \in \Pi^\infty(x_0)$, since $\beta \in (0,1)$,

$$\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t |F(x_t, x_{t+1})| \le \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t B = \frac{B}{1-\beta}.$$

That is,

$$|u\{x_{t+1}\}| \le \frac{B}{1-\beta}.$$

Splitting the infinite sum into the first $T-1$ periods and the remaining periods:

$$u(\{x_{t+1}\}) = \sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1}) + \beta^T \left[\sum_{t=T}^{\infty} \beta^{t-T} F(x_t, x_{t+1})\right].$$

Then, we can write

$$\left| u(\{x_{t+1}\}) - \sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1}) \right| = \beta^T \left| \sum_{t=T}^{\infty} \beta^{t-T} F(x_t, x_{t+1}) \right| \le \beta^T \frac{B}{1-\beta}.$$

Since β^T can be made arbitrarily small by choosing T sufficiently large, we can approximate $u(\{x_{t+1}\})$—i.e. the value of a plan—by $\sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1})$ for sufficiently large T.

By definition, (9.2) means that $V(x_0)$ is the maximum value such that

$$V(x_0) \ge F(x_0, x_1) + \beta V(x_1), \ \forall (x_0, x_1) \in \Pi^1(x_0) \tag{9.3}$$

and for some $(x_0, x_1) \in \Pi^1(x_0)$, the above holds with equality; i.e.

$$V(x_0) = F(x_0, x_1) + \beta V(x_1).$$

Since (9.2) holds for all $x \in X$, it must also be the case that

$$V(x_1) \ge F(x_1, x_2) + \beta V(x_2), \ \forall (x_0, x_1, x_2) \in \Pi^2(x_0).$$

Substituting into (9.3),

$$\begin{aligned} V(x_0) &\ge F(x_0, x_1) + \beta V(x_1) \\ &\ge F(x_0, x_1) + \beta [F(x_1, x_2) + \beta V(x_2)] \\ &= F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 V(x_2), \ \forall (x_0, x_1, x_2) \in \Pi^2(x_0). \end{aligned}$$

Again, there exists some $(x_0, x_1, x_2) \in \Pi^2(x_0)$ such that

$$V(x_0) = F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 V(x_2).$$

Continuing in a similar fashion, we get

$$V(x_0) \geq \sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1}) + \beta^T V(x_T), \ \forall x^T \in \Pi^T(x_0)$$

and, for some $x^T \in \Pi^T(x_0)$,

$$V(x_0) = \sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1}) + \beta^T V(x_T). \tag{9.4}$$

Suppose that $V(x)$ is bounded; i.e.

$$|V(x)| \leq \frac{B}{1-\beta}.$$

Since this must hold for all $x \in X$, in particular, it must hold in the limit; i.e.

$$\lim_{T\to\infty} \beta^T V(x_T) = 0.$$

Thus, we can approximate the value of $V(x_0)$ arbitrarily well by choosing T sufficiently large; i.e. we conclude that

$$V(x_0) = V^*(x_0).$$

9.2 Bounded Dynamic Programming

We now study the Bellman equation for the case where F is bounded.

Definition 9.1. *(Contraction)* Let (S, ρ) be a metric space and $T : S \to S$ an operator. T is a contraction with modulus $\beta \in (0,1)$ if

$$\rho(Tx, Ty) \leq \beta\rho(x, y), \ \forall x, y \in S.$$

Thus, T is a contraction if it "shrinks" the distance between any two points by more than the fraction β.

Remark 9.1. Let $S = [0,1]$ so that x and y are scalars between zero and one. Let $\rho(x,y) = |x-y|$ and T be a function. Then, for T to be a contraction with modulus β, it must be that

$$\begin{aligned} &|T(x) - T(y)| \leq \beta|x-y| \\ \Rightarrow\ &\frac{|T(y) - T(x)|}{|y-x|} \leq \beta. \end{aligned}$$

Hence, this says that the slope of the function T is less than β.

In our main application, S will be the set of continuous and bounded functions from X to $\mathbb{R}$:

$$S := \left\{ f : X \to \mathbb{R},\ f \text{ is continuous, and } \|f\| = \sup_{x\in X} |f(x)| < \infty \right\}$$

and the metric is the sup norm:

$$\rho(f, g) = \|f - g\| = \sup_{x\in X} |f(x) - g(x)|.$$

Definition 9.2. A sequence $\{x_n\} \subseteq S$ is a *Cauchy sequence* if, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\rho(x_n, x_m) < \varepsilon$ for all $n, m > N$.

A Cauchy sequence is a sequence whose elements become arbitrarily close to each other as the sequence progresses.

Definition 9.3. A linear space S is complete if any Cauchy sequence converges.

Theorem 9.1. *The set of bounded and continuous functions is complete.*

Theorem 9.2. *(Contraction Mapping Theorem) If T is a contraction in (S, ρ) with modulus β, then:*

(i) there is a unique fixed point $s^* \in S$,

$$s^* = Ts^*;$$

(ii) iterations of T converge to the fixed point

$$\rho(T^n s_0, s^*) \leq \beta^n \rho(s_0, s^*), \ \forall s_0 \in S,$$

where $T^{n+1}s = T(T^n s)$.

Corollary 9.1. *Let S be a complete metric space and $S' \subset S$ be closed. Let T be a contraction on S and $s^* = Ts^*$. If $TS' \subset S'$ (i.e. if $s' \in S'$, then $T(s') \in S'$), then $s^* \in S'$. Moreover, if $S'' \subset S'$ and $TS' \subset S''$ (i.e. if $s' \in S'$, then $T(s') \in S''$), then $s^* \in S''$.*

The first statement of Corollary 9.1 means that, if applying the operator T to any point in a closed set S' means that the "output" stays within the set, then the fixed point must also be in that set. The second means that, if applying the operator T to any point in S' means that the output is in the set S'', then the fixed point must also be in S''. Note that S'' need not be closed, whereas S' must be closed.

The following proposition allows us to determine if something is a contraction.

Proposition 9.1. *(Blackwell Sufficient Conditions) Let S be the space of bounded functions of X and $\|\cdot\|$ the* sup *norm. Let $T : S \to S$. Then, T is a contraction if*

(i) T is monotone; i.e. $Tf(x) \leq Tg(x)$ for any $x \in X$ and g, f such that $f(x) \leq g(x)$ for all $x \in X$;

(ii) T discounts; i.e. there exists $\beta \in (0, 1)$ such that, for any $a \in \mathbb{R}_+$,

$$T(f + a)(x) \leq Tf(x) + a\beta, \ \forall x \in X, f \in S.$$

Proof. First, we can always write

$$f - g = f - g.$$

By the definition of sup norm ($\|f(x)\| = \sup |f(x)|$):

$$\begin{aligned} f - g &\leq \|f - g\| \\ \Rightarrow f &\leq g + \|f - g\|. \end{aligned}$$

Since T is monotone, then

$$Tf \leq T(g + \|f - g\|).$$

Since T discounts, setting $a = \|f - g\|$, we can write

$$\begin{aligned} Tf &\leq T\left(g + \|f - g\|\right) \\ &\leq Tg + \beta \left\|f - g\right\| \\ \Rightarrow Tf - Tg &\leq \beta \left\|f - g\right\|. \end{aligned}$$

Of course, we can reverse the roles of f and g to obtain that

$$\begin{aligned} Tg &\leq Tf + \beta \left\|g - f\right\| \\ &= Tf + \beta \left\|f - g\right\| \\ \Rightarrow Tg - Tf &\leq \beta \left\|f - g\right\|. \end{aligned}$$

Combining both inequalities,

$$\left\|Tf - Tg\right\| \leq \beta \left\|f - g\right\|;$$

i.e. T is a contraction. ■

We define the Bellman operator T as

$$(Tv)(x) = \max_{y \in \Gamma(x)} \left[F(x, y) + \beta v(y)\right]. \tag{9.5}$$

Theorem 9.3. *Assume that F is bounded and continuous and that Γ is continuous and has range that is compact. Let T be the Bellman operator as defined in (9.5). Then, T maps the set of continuous and bounded functions S onto itself. Moreover, T is a contraction.*

Proof. That T maps the set of continuous and bounded follow from the Theorem of the Maximum. That T is a contraction follows since T satisfies the Blackwell conditions:

(i) Monotonicity. For $f \geq v$,

$$\begin{aligned} (Tv)(x) &= \max_{y \in \Gamma(x)} \left[F(x, y) + \beta v(y)\right] \\ &= F(x, g(x)) + \beta v(g(x)) \\ &\leq F(x, g(x)) + \beta f(g(x)) \because f \geq v \\ &\leq \max_{y \in \Gamma(x)} \left[F(x, y) + \beta f(y)\right] = (Tf)(x), \end{aligned}$$

where the last inequality comes from the fact that $Tf(x)$ is the maximum.

(ii) Discounts. For $a > 0$,

$$\begin{aligned} (T(v + a))(x) &= \max_{y \in \Gamma(x)} \left[F(x, y) + \beta(v(y) + a)\right] \\ &= \max_{y \in \Gamma(x)} \left[F(x, y) + \beta v(y)\right] + \beta a \\ &= (Tv)(x) + \beta a. \end{aligned}$$

■

Corollary 9.1 is useful in establishing the properties of the value function V and the optimal policy g.

Theorem 9.4. *(Monotonicity) Assume that $F(\cdot, y)$ is increasing and that Γ is increasing (i.e. $\Gamma(x) \subseteq \Gamma(x')$ for $x \leq x'$). Then, the fixed point v^* satisfying $v^* = Tv^*$ is increasing. If $F(\cdot, y)$ is strictly increasing, so is v^*.*

Theorem 9.5. *(Concavity) Assume that X is convex and Γ is convex; i.e. $y \in \Gamma(x)$ and $y' \in \Gamma(x')$ implies that*

$$y^\theta := \theta y' + (1-\theta) y \in \Gamma(\theta x' + (1-\theta) x) =: \Gamma\left(x^\theta\right), \forall x, x' \in X, \forall \theta \in (0,1).$$

Further, assume that F is concave in (x, y). Then, the fixed point v^* satisfying $v^* = Tv^*$ is concave in x. Moreover, if $F(\cdot, y)$ is strictly concave, so is v^*.

Proof. (*Theorem 9.4*) By Corollary 9.1, it suffices to show that Tf is increasing if f is increasing. Let $x \leq x'$:

$$\begin{aligned}
(Tf)(x) &= \max_{y \in \Gamma(x)} [F(x, y) + \beta f(y)] \\
&= F(x, g(x)) + \beta f(g(x)) \\
&\leq F(x', g(x)) + \beta f(g(x))
\end{aligned}$$

since $F(\cdot, y)$ is increasing (the inequality will hold strictly if $F(\cdot, y)$ is strictly increasing). But by definition,

$$\begin{aligned}
(Tf)(x) &\leq F(x', g(x)) + \beta f(g(x)) \\
&\leq F(x', g(x')) + \beta f(g(x')) \\
&= \max_{y \in \Gamma(x')} [F(x', y) + \beta f(y)] \\
&= (Tf)(x').
\end{aligned}$$

■

Proof. (*Theorem 9.5*) By Corollary 9.1, it suffices to show that Tf is concave if f is concave. That is, we wish to show that

$$(Tf)\left(x^\theta\right) \geq (1-\theta)(Tf)(x') + \theta (Tf)(x), \forall \theta \in (0,1).$$

Since F and f are concave then, for any $x, x' \in X$,

$$\begin{aligned}
F\left(x^\theta, y^\theta\right) &\geq (1-\theta) F(x', y') + \theta F(x, y), \\
f\left(y^\theta\right) &\geq (1-\theta) f(y') + \theta f(y),
\end{aligned}$$

for all $y \in \Gamma(x)$ and $y' \in \Gamma(x')$. Summing the two while multiplying the second by β:

$$\begin{aligned}
F\left(x^\theta, y^\theta\right) + \beta f\left(y^\theta\right) &\geq (1-\theta) F(x', y') + \theta F(x, y) + \beta [(1-\theta) f(y') + \theta f(y)] \\
&= (1-\theta)[F(x', y') + \beta f(y')] + \theta [F(x, y) + \beta f(y)]
\end{aligned}$$

for all $\theta \in (0,1)$. Above holds when $y = g(x)$ and $y' = g(x')$; i.e.

$$\begin{aligned}
F\left(x^\theta, y^\theta\right) + \beta f\left(y^\theta\right) &\geq (1-\theta)[F(x', g(x')) + \beta f(g(x'))] + \theta [F(x, g(x)) + \beta f(g(x))] \\
&= (1-\theta)(Tf)(x') + \theta (Tf)(x).
\end{aligned}$$

By definition,

$$\begin{aligned}(Tf)\left(x^{\theta}\right) &= \max_{y\in\Gamma(x^{\theta})}\left[F\left(x^{\theta},y\right)+\beta f\left(y\right)\right]\\ &= F\left(x^{\theta},g\left(x^{\theta}\right)\right)+\beta f\left(g\left(x^{\theta}\right)\right)\\ &\geq F\left(x^{\theta},y^{\theta}\right)+\beta f\left(y^{\theta}\right),\ \forall y^{\theta}\in\Gamma\left(x^{\theta}\right).\end{aligned}$$

Hence,

$$(Tf)\left(x^{\theta}\right)\geq(1-\theta)\left(Tf\right)\left(x'\right)+\theta\left(Tf\right)\left(x\right),\ \forall\theta\in(0,1).$$ ■

9.2.1 Envelope: Differentiability of the value function

Let us work out a heuristic argument to find an expression for the derivative of the value function. We will assume that V is differentiable and that the policy function g is also differentiable with respect to x. Assume that $(y,x)\in\operatorname{Int}(\operatorname{Gr}(\Gamma))$.

First, the first-order condition of the problem

$$\max_{y} F\left(x,y\right)+\beta V\left(y\right)$$

evaluated at the optimum, $y=g\left(x\right)$, is

$$F_y\left(x,g\left(x\right)\right)+\beta V'\left(g\left(x\right)\right)=0.$$

Now, differentiating both sides of

$$V\left(x\right)=F\left(x,g\left(x\right)\right)+\beta V\left(g\left(x\right)\right)$$

with respect to x gives

$$\begin{aligned}V'\left(x\right) &= F_x\left(x,g\left(x\right)\right)+F_y\left(x,g\left(x\right)\right)g'\left(x\right)+\beta V'\left(g\left(x\right)\right)g'\left(x\right)\\ &= F_x\left(x,g\left(x\right)\right)+\underbrace{\left(F_y\left(x,g\left(x\right)\right)+\beta V'\left(g\left(x\right)\right)\right)}_{=0\because\text{FOC}}g'\left(x\right).\end{aligned}\tag{9.6}$$

Hence,

$$V'\left(x\right)=F_x\left(x,g\left(x\right)\right).$$

This is called the *envelope* condition.

Remark 9.2. The formal proof (Benveniste and Scheinkman Theorem) requires that V is concave, $F\left(\cdot,y\right)\in C^1$ and that $\left(g\left(x\right),x\right)\in\operatorname{Int}(\operatorname{Gr}(\Gamma))$. Strictly speaking, the theorem does not require g to be differentiable.

9.2.2 First-order and the envelope conditions

The first-order and the envelope conditions are respectively given by

$$\begin{aligned}0 &= F_y\left(x,g\left(x\right)\right)+\beta V'\left(g\left(x\right)\right),\\ V'\left(x\right) &= F_x\left(x,g\left(x\right)\right)\end{aligned}$$

for all x such that $\left(g\left(x\right),x\right)\in\operatorname{Int}(\operatorname{Gr}(\Gamma))$.

Notice that combining the two gives the familiar Euler Equation:

$$0=F_y\left(x,g\left(x\right)\right)+\beta F_x\left(g\left(x\right),g\left(g\left(x\right)\right)\right).$$

9.2.3 Neoclassical growth model

Exercise 9.1. Show that the neoclassical growth model satisfy the assumptions of Theorem *9.4*.

Solution. Recall that the neoclassical growth model is given by

$$\begin{aligned} V^*(k_0) := \max_{\{k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t U\left(f\left(k_t\right) - k_{t+1}\right) \\ s.t. \quad & 0 \leq k_{t+1} \leq f\left(k_t\right), \\ & k_0 \text{ given,} \end{aligned}$$

where $f(k) = G(k,1) + (1-\delta)k$ and G is strictly increasing and strictly concave in k, satisfies Inada conditions, and U is strictly increasing and strictly concave. Then, the Bellman equation for this problem is given by

$$V(k) = \max_{k' \in [0, f(k)]} \left[U\left(f(k) - k'\right) + \beta V\left(k'\right)\right].$$

Since U is strictly increasing and $f(\cdot)$ is also strictly increasing in k; i.e. the condition that $F(\cdot, y)$ is increasing is satisfied. Let $\Gamma(x) = [0, f(x)]$. Since $f(\cdot)$ is strictly increasing, we also have that Γ is increasing.

Exercise 9.2. Show that the neoclassical growth model satisfy the assumptions of Theorem *9.5*.

Solution. Since $k \in \mathbb{R}$, X is convex, and since $\Gamma(x) = [0, f(k)]$ is an interval, Γ is also convex. Note that

$$\begin{aligned} F_x &= U' f' \\ F_y &= -U', \\ F_{xx} &= U''\left(f'\right)^2 + U' f'' < 0 \\ F_{yy} &= U'' < 0 \\ F_{yx} = F_{xy} &= -U'' f' > 0. \end{aligned}$$

Moreover,

$$\begin{aligned} F_{xx}F_{yy} - \left(F_{yx}\right)^2 &= \left(U''\left(f'\right)^2 + U' f''\right)U'' - \left(-U'' f'\right)^2 \\ &= U' f'' U'' > 0. \end{aligned}$$

Hence, F is strictly concave.

Example 9.1. *(Neoclassical growth model)* The Bellman equation for the Neoclassical problem is given by

$$V(k) = \max_{k' \in [0, f(k)]} \left[U\left(f(k) - k'\right) + \beta V\left(k'\right)\right].$$

Thus, the first-order condition evaluated at the optimal $k' = g(k)$ is

$$U'\left(f(k) - g(k)\right) = \beta V'\left(g(k)\right).$$

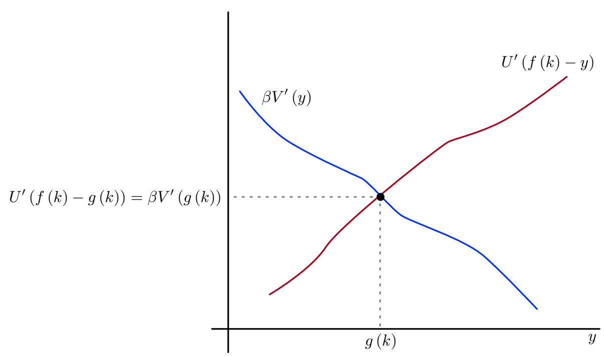
The envelop condition, evaluated at the optimal, is given by

$$V'\left((k)\right) = U'\left(f(k) - g(k)\right) f'(k).$$

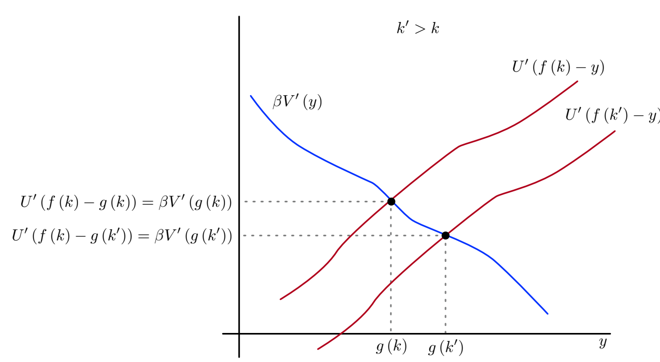
9.2.4 Exercises

Exercise 9.3. Show that g, the decision rule for the neoclassical growth model, is increasing in k. Graph $U'(f(k)-y)$ against $\beta V'(y)$ as functions of y, using the result that V is strictly concave. The intersection of the two functions gives $y=g(k)$. Increase the value of k to, say, $k'>k$, and argue that the new intersection occurs at a larger value for y; i.e. g is strictly increasing so that $g(k')>g(k)$. Use this graph to show that the slope of g is bounded by the slope of f. Show that the optimal consumption function $c(k):=f(k)-g(k)$ is increasing in k.

Solution. Recall that V is a strictly increasing and strictly concave function (Exercises 9.1 and 9.2). Similarly, U is a strictly increasing and strictly concave function (by assumption). Thus, $U'(f(k)-y)$ curve is strictly increasing in y, whereas $\beta V'(y)$ is a strictly decreasing function in y. Thus, there can be at most one intersection. That there in fact is an intersection follows if we impose Inada conditions on f and U.



Suppose now that k increases to $k'>k$, which affects the $U'(f(k)-y)$ curve. In particular, since $f(k)$ is strictly increasing and U' is strictly decreasing, $U'(f(k)-y)$ is strictly decreasing in k so that a higher k shifts the curve down. As can be seen from the figure below, this means that $g(k')>g(k)$.



We can also prove this formally.

Suppose, by way of contradiction, that $g(k')\leq g(k)$ with $k'>k$. Since U is strictly concave and $f(k')$ is strictly increasing in k,

$$U'(f(k')-g(k'))<U'(f(k)-g(k)).$$

Using the first-order condition, above implies that

$$\beta V'(g(k'))<\beta V'(g(k)),$$

which contradicts that V is strictly concave (since $g(k') \leq g(k) \Rightarrow \beta V'(g(k')) \geq \beta V'(g(k))$).

To show that the slope of g is bounded (above) by the slope of f. Notice from the figure that

$$U'(f(k') - g(k')) < U'(f(k) - g(k))$$

Since U is strictly concave, this implies that

$$\begin{aligned} & f(k') - g(k') > f(k) - g(k) \\ \Leftrightarrow\ & f(k') - f(k) > g(k') - g(k). \end{aligned}$$

Define Δ such that that $k' = k + \Delta$, then, from above, we can infer that

$$f'(k) := \lim_{\Delta \downarrow 0} \frac{f(k+\Delta) - f(k)}{\Delta} > \lim_{\Delta \downarrow 0} \frac{g(k+\Delta) - g(k)}{\Delta} =: g'(k).$$

That $f'(k) > g'(k)$ also means that (from market clearing):

$$c'(k) = f'(k) - g'(k) > 0.$$

That is, the optimal consumption function is increasing in k.

Exercise 9.4. Assume that $X = \mathbb{R}_+$, $\Gamma(x) = \mathbb{R}$ and $F \in C^2$, $F_x \geq 0$, $F_y \leq 0$, and F is strictly concave. Argue that g is increasing in x if and only if $F_{xy} \geq 0$.

Solution. Since $F(x, y)$ is increasing in x and Γ is (weakly) increasing, we know that V is increasing (Theorem 9.1). Moreover, Since X and Γ are convex and F is strictly concave, it follows that V is also strictly concave (Theorem 9.2). Differentiating the first-order condition with respect to x:

$$\begin{aligned} & 0 = F_{yx} + F_{yy} g' + \beta V'' g' \\ \Rightarrow\ & g' = -\frac{F_{yy} + \beta V''}{F_{xy}}. \end{aligned}$$

The numerator is strictly negative since F and V are strictly concave, thus, $g'(x) \geq 0 \Leftrightarrow F_{xy} \geq 0$.

Exercise 9.5. Consider the problem of an agent with wages w that saves with riskless gross rate of return $1 + r$. The budget constraint is

$$y + c = x(1 + r) + w,$$

where x is the beginning of period wealth and y are savings. Let $(1 + r)\beta = 1$, $w > 0$ and U strictly increasing, bounded, strictly concave and C^2. Write down the Bellman equation for this problem. Show that

$$\begin{aligned} c(x) &= w + rx, \\ g(x) &= x, \\ V(x) &= \frac{U(w + rx)}{1 - \beta} \end{aligned}$$

for all $x \in \mathbb{R}_+$.

Solution. The agent's sequence problem is

$$\max_{y_t \in \mathbb{R}} \quad \sum_{t=0}^{\infty} \beta^t U\left(x_t\left(1+r\right)+w_t-y_t\right).$$

where we used the budget constraint to replace c in the objective function. The Bellman equation is given by

$$V\left(x\right)=\max_{y\in\mathbb{R}}\left[U\left(x\left(1+r\right)+w-y\right)+\beta V\left(y\right)\right].$$

The first-order and the envelope conditions are:

$$U'\left(x\left(1+r\right)+w-g\left(x\right)\right)=\beta V'\left(g\left(x\right)\right), \tag{9.7}$$
$$\left(1+r\right)U'\left(x\left(1+r\right)+w-g\left(x\right)\right)=V'\left(x\right).$$

Since $\left(1+r\right)=1/\beta$, the envelope condition gives us hat

$$U'\left(x\left(1+r\right)+w-g\left(x\right)\right)=\beta V'\left(x\right). \tag{9.8}$$

Notice that the left-hand sides of (9.7) and (9.8) are the same so that

$$V'\left(g\left(x\right)\right)=V'\left(x\right).$$

Given the assumptions on U, we know that V is strictly concave so that it must be that

$$g\left(x\right)=x.$$

Substituting this into the budget constraint yields

$$\begin{aligned} c &= x\left(1+r\right)+w-g\left(x\right) \\ &= x\left(1+r\right)+w-x \\ &= w+rx. \end{aligned}$$

Substituting the expression for consumption and $g\left(x\right)=x$ into the Bellman equation while noting that $y=g\left(x\right)$ is the maximiser:

$$\begin{aligned} V\left(x\right) &= U\left(w+rx\right)+\beta V\left(x\right) \\ &= \frac{U\left(w+rx\right)}{1-\beta}. \end{aligned}$$

Exercise 9.6. *(Linear utility in the neoclassical growth model).* Let $U\left(c\right)=c$ and

$$f\left(k\right)=G\left(k,1\right)+\left(1-\delta\right)k,$$

where G is a neoclassical function: strictly increasing and strictly concave in k, satisfying Inada conditions. Assume that

$$0\leq k'\leq f\left(k\right).$$

Show that there is a $\bar{k}$ such that $k \leq f(k) \leq \bar{k}$ for all $0 \leq k \leq \bar{k}$. [Hint: graph $f(k)$ vs the 45 degree line.] Let $X = [0, \bar{k}]$. Write the Bellman equation. Show that V is increasing and concave in k. Show that, for k^* satisfying $\beta f'(k^*) = 1$, there is an $\varepsilon > 0$ such that for $k : |k - k^*| \leq \varepsilon$,

$$V(k) = f(k) - k^* + \frac{\beta}{1-\beta}(f(k^*) - k^*).$$

Characterise the optimal policy as much as possible.

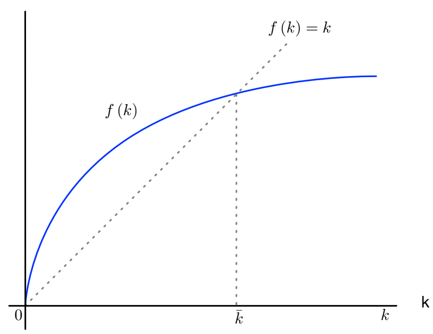
Solution. Since G satisfies Inada conditions, in particular,

$$G(0,1) = 0,$$
$$\lim_{k \to 0} G_1(k,1) = \infty.$$

Hence, $f(0) = 0$ and $\lim_{k\to 0} f'(k) = \infty$. That is, from the origin, $f(k)$ lies above the 45 degree line. Since $f(k)$ is concave and increasing;

$$f'(k) = G_1(k,1) + 1 - \delta > 0,$$

it must eventually cross the 45 degree line at $\bar{k}$. Then, for all $k \in [0, \bar{k}]$, $f(k) \geq k$.



The Bellman equation is given by

$$V(k) = \max_{y \in [0, \bar{k}]} [(f(k) - y) + \beta V(y)].$$

Since f is increasing in k, the period return function $F(x,y) := f(k) - y$ is strictly increasing in k. Moreover, $\Gamma(k) = [0, \bar{k}]$ is (weakly) increasing so that, from Theorem 9.1, we know that V is strictly increasing. That V is strictly concave follows from the fact that the state space $X = \mathbb{R}$ and $\Gamma(k)$ is convex, and that $f(k) - y$ is strictly concave in k (Theorem 9.2).

The corresponding sequence problem is given by

$$V^*(k_0) = \max_{\{k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t (f(k_t) - k_{t+1}).$$

The first-order condition is given by

$$-\beta^t + \beta^{t+1} f'(k^*_{t+1}) = 0 \Leftrightarrow f'(k^*_{t+1})\beta = 1.$$

Hence, k_{t+1}^* is constant for all t. Call this constant k^*. Then,

$$\begin{aligned}
V^*\left(k_0\right) &= \sum_{t=0}^{\infty} \beta^t \left(f\left(k_t^*\right) - k^*\right) \\
&= f\left(k_0\right) - k^* + \sum_{t=1}^{\infty} \beta^t \left(f\left(k^*\right) - k^*\right) \\
&= f\left(k_0\right) - k^* + \beta \sum_{t=0}^{\infty} \beta^t \left(f\left(k^*\right) - k^*\right) \\
&= f\left(k_0\right) - k^* + \frac{\beta}{1-\beta} \left(f\left(k^*\right) - k^*\right).
\end{aligned}$$

Exercise 9.7. Let the adjustment cost model be

$$\begin{aligned}
F\left(x, y\right) &= -\frac{a}{2} y^2 - \frac{b}{2} \left(y - x\right)^2, \\
\Gamma\left(x\right) &= \mathbb{R}.
\end{aligned}$$

Show that $V\left(x\right) = -\frac{c}{2} x^2$. Find an expression for the coefficient c as a function of a, b and β. Why is this specification called an adjustment cost model?

Solution. The Bellman equation is given by

$$V\left(x\right) = \max_{y \in \mathbb{R}} \left[-\frac{a}{2} y^2 - \frac{b}{2} \left(y - x\right)^2 + \beta V\left(y\right)\right].$$

Substituting the "guess",

$$V\left(x\right) = \max_{y \in \mathbb{R}} \left[-\frac{a}{2} y^2 - \frac{b}{2} \left(y - x\right)^2 - \beta \frac{c}{2} y^2\right].$$

The first-order condition is

$$\begin{aligned}
0 &= -ay - b\left(y - x\right) - \beta c y \\
\Rightarrow y &= \frac{b}{a + b + \beta c} x, \\
y^2 &= \frac{b^2}{\left(a + b + \beta c\right)^2} x^2
\end{aligned}$$

Let's rearrange the maximand first,

$$-\frac{a}{2} y^2 - \frac{b}{2} \left(y - x\right)^2 - \beta \frac{c}{2} y^2 = -\frac{y^2}{2} \left(a + b + \beta c\right) + byx - \frac{b}{2} x^2.$$

Substituting for y and y^2,

$$
\begin{aligned}
&-\frac{y^2}{2}\left(a+b+\beta c\right)+byx-\frac{b}{2}x^2 \\
=&-\frac{1}{2}\frac{b^2}{a+b+\beta c}x^2+\frac{b^2}{a+b+\beta c}x^2-\frac{b}{2}x^2 \\
=&\frac{b}{2}\left(\frac{b}{a+b+\beta c}-1\right)x^2 \\
=&-\frac{1}{2}b\left(\frac{a+\beta c}{a+b+\beta c}\right)x^2
\end{aligned}
$$

Hence,

$$
\begin{aligned}
c &= b\left(\frac{a+\beta c}{a+b+\beta c}\right) \\
\Leftrightarrow \left(a+b\right)c+\beta c^2 &= ab+\beta bc \\
\Leftrightarrow 0 &= \beta c^2+\left(a+\left(1-\beta\right)b\right)c-ab \\
\Leftrightarrow c &= \frac{-\left(a+\left(1-\beta\right)b\right)\pm\sqrt{\left(a+\left(1-\beta\right)b\right)^2-4\beta ab}}{2\beta}.
\end{aligned}
$$

The specification is called an adjustment cost model since the term $\left(y-x\right)^2$ represents the cost of adjusting y while y^2 represents the implicit target of zero for y.

Exercise 9.8. Consider the Neoclassical growth model with log utility, Cobb-Douglas production function and 100% depreciation:

$$
\begin{aligned}
F\left(x,y\right) &= \log\left(x^\alpha-y\right), \\
\Gamma\left(x\right) &= \left[0,x^\alpha\right].
\end{aligned}
$$

Show that

$$
\begin{aligned}
V\left(x\right) &= a+b\log x, \\
g\left(x\right) &= cx^\alpha.
\end{aligned}
$$

Find expressions for a, b and c.

Solution. The Bellman equation is given by

$$V\left(x\right)=\max_{y\in[0,x^\alpha]}\left[\log\left(x^\alpha-y\right)+\beta V\left(y\right)\right].$$

Substituting the "guess",

$$V\left(x\right)=\max_{y\in[0,x^\alpha]}\left[\log\left(x^\alpha-y\right)+\beta\left(a+b\log y\right)\right].$$

The first-order condition is

$$\begin{aligned} 0 &= -\frac{1}{x^\alpha - y} + \beta\frac{b}{y} \\ \Rightarrow y &= \beta b\left(x^\alpha - y\right) \\ &= \underbrace{\frac{\beta b}{1+\beta b}}_{=c} x^\alpha. \end{aligned}$$

Substituting into the maximand,

$$\begin{aligned} &\log\left(x^\alpha - y\right) + \beta\left(a + b\log y\right) \\ =& \log\left(x^\alpha - \frac{\beta b}{1+\beta b}x^\alpha\right) + \beta\left(a + b\log\frac{\beta b}{1+\beta b}x^\alpha\right) \\ =& \log\left(x^\alpha\frac{1+\beta b-\beta b}{1+\beta b}\right) + a\beta + \beta b\log\frac{\beta b}{1+\beta b} + \beta b\log x^\alpha \\ =& \log\left(x^\alpha\right) + \log\frac{1}{1+\beta b} + a\beta + \beta b\log\frac{\beta b}{1+\beta b} + \beta b\log x^\alpha \\ =& \left(1+\beta b\right)\alpha\log x + a\beta + \log\frac{1}{1+\beta b} + \beta b\log\frac{\beta b}{1+\beta b}. \end{aligned}$$

Hence,

$$\begin{aligned} b &= \left(1+\beta b\right)\alpha \\ &= \frac{\alpha}{1-\alpha\beta}, \\ c &= \frac{\beta b}{1+\beta b} = \frac{1}{\frac{1}{\beta b}+1} = \frac{1}{\frac{1}{\frac{\alpha\beta}{1-\alpha\beta}}+1} \\ &= \alpha\beta, \\ a &= a\beta + \log\frac{1}{1+\beta b} + \beta b\log\frac{\beta b}{1+\beta b} \\ &= \frac{1}{1-\beta}\left(\log\frac{1}{1+\frac{\alpha\beta}{1-\alpha\beta}} + \frac{\alpha\beta}{1-\alpha\beta}\log\alpha\beta\right) \\ &= \frac{1}{1-\beta}\left(\log\left(1-\alpha\beta\right) + \frac{\alpha\beta}{1-\alpha\beta}\log\alpha\beta\right). \end{aligned}$$

Exercise 9.9. Consider the Neoclassical growth model with 100% depreciation,

$$\begin{aligned} f\left(k\right) &= \left[\alpha k^{1-\frac{1}{\rho}} + \left(1-\alpha\right)^{1-\frac{1}{\rho}}\right]^{\frac{1}{1-\frac{1}{\rho}}}, \\ U\left(c\right) &= \frac{c^{1-\frac{1}{\sigma}}-1}{1-\frac{1}{\sigma}}. \end{aligned}$$

Look for the relationship between parameters ρ and σ such that a solution of the type

$$g\left(k\right) = af\left(k\right)$$

for some number $a \in (0,1)$ and c and b such that

$$V(k) = c + bU(af(k)).$$

[Hint: The previous exercise is a special case of this. In the previous case, the elasticity of substitution of capital is one, and the intertemporal elasticity of substitution σ is also one.]

Solution. The feasibility constraint is that

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

Letting $x = k_t$ and $y = k_{t+1}$ and $\delta = 1$, the Bellman equation is given by

$$V(x) = \max_{y \in [0, f(x)]} \{U(f(x) - y) + \beta V(y)\}.$$

The first-order condition is

$$\begin{aligned} U'(f(x) - y) &= \beta b U'(y) \\ \Rightarrow [f(x) - y]^{-\frac{1}{\sigma}} &= \beta b y^{-\frac{1}{\sigma}} \\ \Rightarrow f(x) - y &= (\beta b)^{-\sigma} y \\ \Rightarrow y &= \underbrace{\frac{1}{1 + (\beta b)^{-\sigma}}}_{=a} f(x). \end{aligned}$$

Substituting into the maximand,

$$\begin{aligned} c + bU(af(x)) =& U(f(x) - y) + \beta[c + bU(y)] \\ \Rightarrow c + b\frac{(af(x))^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} =& U(f(x) - af(x)) + \beta[c + bU(af(x))] \\ =& \beta c + \beta b U(af(x)) + U(f(x)(1-a)) \\ =& \beta c + \beta b \frac{[af(x)]^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} + \frac{[f(x)(1-a)]^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} \\ =& \beta c + \frac{1}{1 - \frac{1}{\sigma}} \left[\beta b [af(x)]^{1-\frac{1}{\sigma}} + [f(x)(1-a)]^{1-\frac{1}{\sigma}} - (1 + \beta b)\right] \\ =& \beta c + \frac{1}{1 - \frac{1}{\sigma}} \left[f(x)^{1-\frac{1}{\sigma}} \left(\beta b a^{1-\frac{1}{\sigma}} + (1-a)^{1-\frac{1}{\sigma}}\right) - (1 + \beta b)\right]. \end{aligned}$$

So we wish to verify if

$$f(x)^{1-\frac{1}{\sigma}} \left(\beta b a^{1-\frac{1}{\sigma}} + (1-a)^{1-\frac{1}{\sigma}}\right) - (1 + \beta b) = b\left((af(x))^{1-\frac{1}{\sigma}} - 1\right).$$

Equating coefficients,

$$1 + \beta b = b \Rightarrow b = \frac{1}{1 - \beta},$$

and

$$\begin{aligned}
\beta b a^{1-\frac{1}{\sigma}} + (1-a)^{1-\frac{1}{\sigma}} &= b a^{1-\frac{1}{\sigma}} \\
\Rightarrow a^{1-\frac{1}{\sigma}} b (1-\beta) &= (1-a)^{1-\frac{1}{\sigma}} \\
\Rightarrow a^{1-\frac{1}{\sigma}} &= (1-a)^{1-\frac{1}{\sigma}} \\
\Rightarrow a &= 1-a \\
\Rightarrow a &= \frac{1}{2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
a = \frac{1}{1+(\beta b)^{-\sigma}} &= \frac{1}{2} \\
\Rightarrow 1 &= \left(\frac{\beta}{1-\beta}\right)^{-\sigma} \\
\Rightarrow \sigma &= 0.
\end{aligned}$$

We also need that

$$c = \beta c \Rightarrow c = 0.$$

9.3 Continuous-time Bellman equation

Consider the following discrete-time Bellman equation,

$$V(x_t) = \max_{u_t \in U} \left[\Delta h(x_t, u_t) + \frac{1}{1+\Delta\rho} V(x_{t+\Delta})\right]$$

subject to

$$x_{t+\Delta} = x_t + \Delta g(x_t, u_t).$$

We will analyse the continuous-time Bellman equation as a limit of the discrete one.

Notice that, if we simply take the limit as Δ goes to zero, we are simply left with $V(x_t) = V(x_t)$, which is not very useful. Using Taylor expansion (around x_t), we can write

$$\begin{aligned}
V(x_{t+\Delta}) &= V(x_t + \Delta g(x_t, u_t)) \\
&= V(x_t) + V'(x_t)\Delta g(x_t, u_t) + o(\Delta g(x_t, u_t)),
\end{aligned}$$

where $d(z) = o(z)$ means that $\lim_{z\to 0} d(z)/z = 0$. Then the Bellman equation is

$$V(x_t) = \max_{u_t \in U} \left[\Delta h(x_t, u_t) + \frac{1}{1+\Delta\rho}\left(V(x_t) + V'(x_t)\Delta g(x_t, u_t) + o(\Delta g(x_t, u_t))\right)\right]$$

Multiplying both sides by the positive constant $1+\Delta\rho$ yields

$$(1+\Delta\rho)V(x_t) = \max_{u_t \in U}\left[(1+\Delta\rho)\Delta h(x_t, u_t) + V(x_t) + V'(x_t)\Delta g(x_t, u_t) + o(\Delta g(x_t, u_t))\right].$$

We can move $V(x_t)$ inside the max to the left-hand side since it does not depend on u_t,

$$\Delta\rho V(x_t) = \max_{u_t\in U}\left[(1+\Delta\rho)\Delta h(x_t,u_t) + V'(x_t)\Delta g(x_t,u_t) + o(\Delta g(x_t,u_t))\right].$$

Dividing both sides by Δ,

$$\rho V(x_t) = \max_{u_t\in U}\left[(1+\Delta\rho) h(x_t,u_t) + V'(x_t) g(x_t,u_t) + \frac{o(\Delta g(x_t,u_t))}{\Delta}\right].$$

Taking the limit as Δ goes to zero,

$$\rho V(x_t) = \max_{u_t\in U}\left[h(x_t,u_t) + V'(x_t) g(x_t,u_t)\right],$$

where we implicitly assume that the limit as $\Delta \to 0$ of the max with respect to u_t is the same as the max with respect to u_t of the limit as $\Delta \to 0$. Removing the time indices, we obtain the continuous-time version of the Bellman equation,

$$\rho V(x) = \max_{u\in U}\left[h(x,u) + V'(x) g(x,u)\right]. \tag{9.9}$$

Under the regularity conditions, the max of the RHS can be characterised using the following first-order condition for u:

$$0 = h_u(x,u^*(x)) + V'(x) g_u(x,u^*(x)),$$

which defines the optimal decision rule $u^*(x)$. Thus, the following two equations summarise the dynamic programming problem:

$$\begin{aligned}\rho V(x) &= h(x,u^*(x)) + V'(x) g(x,u^*(x)), \qquad (9.10)\\ 0 &= h_u(x,u^*(x)) + V'(x) g_u(x,u^*(x))\end{aligned}$$

for all $x \in X$. Notice that these are two functional equations (i.e. solutions are functions). The functions V and u^* are both functions of x.

9.4 Bellman equation and the Maximum Principle

We now show the sense in which the Bellman equation and the first-order conditions above imply the equations for the Maximum Principle (i.e. Hamiltonian) derived previously.

Recall that in the optimal-control approach, $u^*(t)$ maximises

$$H(x,u,\lambda) = h(x,u) + \lambda g(x,u).$$

From (9.9), in the dynamic programming approach, $u^*(t)$ maximises

$$h(x,u) + V'(x) g(x,u). \tag{9.11}$$

Hence, the two approaches are consistent only if

$$\lambda \equiv V'(x);$$

i.e. the co-state in the calculus of variations approach is the derivative of the value function. This is consistent with the interpretation for the discounted value of the co-state variables offered before: the marginal value of an extra unit of the state variable.

Second, using that $\lambda \equiv V'(x)$, we can see that the first-order condition from optimal-control approach

$$H_u(x,u,\lambda) = 0$$

is equivalent to

$$h_u(x,u) + \lambda g_u(x,u) = 0,$$

which is the derivative of (9.11) with respect to u set to zero.

Third, differentiating (9.10) with respect to time yields

$$\begin{aligned}
\rho V'\dot{x} &= h_x\dot{x} + h_u u^{*\prime}\dot{x} + \left(V''g + V'g_x + V'g_u u^{*\prime}\right)\dot{x} \\
&= h_x\dot{x} + \left(V''g + V'g_x\right)\dot{x} + \left(\underbrace{h_u + V'g_u}_{=0\because \text{FOC}}\right)u^{*\prime}\dot{x} \\
&= h_x\dot{x} + \left(V''g + V'g_x\right)\dot{x}.
\end{aligned}$$

Outside of the steady state, $\dot{x} \neq 0$, so we can divide through to obtain

$$\rho V' = h_x + V''g + V'g_x. \tag{9.12}$$

Differentiating $\lambda \equiv V'(x)$ with respect to time yields and using that $\dot{x} = g(x,u)$,

$$\dot{\lambda} = V''\dot{x} = V''g.$$

Substituting above and $\lambda \equiv V'(x)$ into (9.12) yields

$$\dot{\lambda} = \rho\lambda - \left(h_x + V'g_x\right),$$

which is equivalent to

$$\dot{\lambda} = \rho\lambda - H_x(x,u,\lambda),$$

which is the law of motion of the co-state variable obtained using the Maximum Principle.

Part V

Ramsey Taxation

The core of this section is the analysis of the steady-state effects of linear taxes to labour and (net) capital income. We adapt the arguments used before for existence and uniqueness of steady state with distortionary taxation. The main result is a comparative static results of the effect of labour and (net) capital income tax rates and of government purchases on the steady-state level of consumption, labour supply and investment.

We also discuss different interpretations and alternative formulations of the utility function.

10 Steady-states: Neoclassical growth model (deterministic case)

In studying the neoclassical model, we have so far assumed that labour is supplied inelastically by the household. In this section, we extend the model by endogenousing the household's labour supply decision and study the competitive equilibrium.

We discuss two extreme versions.

- ▷ version 1: firms own capital and household owns firms;
- ▷ version 2: households own capital and rents capital to firms.

We will see that the definition of the consumption possibility set, budget constraint and production possibility sets differ in the two interpretations of the competitive equilibrium. Nevertheless, in both cases, we will find that the equilibrium prices and aggregate quantities are identical.

We then analyse the steady state, paying particular attention to the fact that labour supply is endogenous. We show that, if leisure is normal (i.e. higher income implies greater time spent on leisure), then there is a unique steady state. We also introduce government purchases, financed with lump-sum taxes. Using that goods are normal, we analyse the effect of government purchases and taxes on output. We also solve a parametric example calibrated roughly to match the US economy.

10.1 The setup

10.1.1 Households

There is a representative household who derives utility from sequences of a unique consumption good and leisure, $\{c_t, 1-n_t\}_{t=0}^{\infty}$ whose preferences are represented by the following utility function:

$$u\left(\{c_t, 1-n_t\}_{t=0}^{\infty}\right) = u\left(c_0, 1-n_0, c_1, 1-n_1, \ldots\right),$$

where c_t is consumption in period t, n_t is labour, and we have normalised total time endowment to one per period so that $1-n_t$ is leisure in period t.

10.1.2 Planning problem

We start with the following planning problem:

$$\begin{aligned} \max_{\{c_t,n_t,x_t,k_{t+1}\}_{t=0}^{\infty}} \quad & u\left(c_0, 1-n_0, c_1, 1-n_1, \ldots\right) \\ s.t. \quad & c_t + x_t = F\left(k_t, n_t\right),\ \forall t \geq 0, \\ & k_{t+1} = x_t + k_t\left(1-\delta\right),\ \forall t \geq 0, \\ & k_0 \text{ given}, \end{aligned}$$

where x_t is investment and k_t is capital. F is a neoclassical, constant returns to scale production (i.e. number of firms is irrelevant) and $\delta \in [0,1]$ is the depreciation rate. Since k_t is given in period t, the choice variance in period t is k_{t+1} with respect to capital.

We consider two versions of the competitive equilibrium corresponding to this economy. In both cases, we assume that all trade takes place at time $t = 0$ (i.e. complete markets/deterministic).

10.1.3 Version 1: Firms as owners of capital

- ▷ Commodity space: Let $\mathbb{R}^{2T}$ for $T = \infty$. More formally,

  $$L \coloneqq \left\{ \{c_t, n_t\}_{t=0}^{\infty} : (c_t, n_t) \in \mathbb{R}^2,\ \forall t \geq 0 \right\};$$

  i.e. the commodity space is the set of all pairs of real sequences.

- ▷ Production possibility set of the firm:

  $$Y \coloneqq \left\{ \{c_t, n_t\}_{t=1}^{\infty} : x_t + c_t \leq F\left(k_t, n_t\right),\ k_{t+1} = x_t + k_t\left(1-\delta\right),\ x_t \in \mathbb{R},\ \forall t \geq 0,\ k_0 \text{ given} \right\}.$$

- ▷ Consumption possibility set:

  $$X \coloneqq \left\{ \{c_t, n_t\}_{t=0}^{\infty} : c_t \geq 0,\ 0 \leq n_t \leq 1 \right\}.$$

- ▷ Household budget constraint:

  $$\sum_{t=0}^{\infty} p_t\left(c_t + w_t \ell_t\right) = \pi + \sum_{t=0}^{\infty} p_t w_t,$$

  where p_t is the price of consumption good at time t, and $p_t w_t$ is the price of a unit of labour at time t. Both prices are in units of the numeraire good. Thus, w_t is the time-t real wage in terms of time-t units of the consumption good. We assume that the profits made by firms π are paid to the household (i.e. households own shares in firms). Note also $\ell_t = 1 - n_t$.

Firm's problem:

$$\pi = \max_{\{\{c_t,n_t\}\in Y\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t\left(c_t - w_t n_t\right).$$

Since π is strictly increasing in c_t, firms will set c_t such that $x_t + c_t = F\left(k_t, n_t\right)$ (i.e. holds with equality). Using the law of motion for capital ($k_{t+1} = x_t + k_t\left(1-\delta\right)$),

$$c_t = F\left(k_t, n_t\right) - k_{t+1} + k_t\left(1-\delta\right).$$

Hence, we can write the firm's problem as

$$\pi = \max_{\{k_{t+1}, n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t \left(F\left(k_t, n_t\right) - w_t n_t - \left(k_{t+1} - k_t\left(1-\delta\right)\right)\right).$$

First-order conditions with respect to n_t and k_{t+1} are

$$\begin{aligned} w_t &= F_n\left(k_t, n_t\right), \\ p_t &= p_{t+1}\left[F_k\left(k_{t+1}, n_{t+1}\right) + \left(1-\delta\right)\right] \\ \Rightarrow -\left(1-\delta\right) + \underbrace{\frac{p_t}{p_{t+1}}}_{:=1+r_t} &= F_k\left(k_{t+1}, n_{t+1}\right) \\ \Rightarrow \delta + r_t &= F_k\left(k_{t+1}, n_{t+1}\right), \end{aligned}$$

where r_t is the real interest rate.[32]

We want to simplify the expression for the maximised profits π. Since F is CRS, by the Euler's Theorem,

$$F\left(k_t, n_t\right) = F_n\left(k_t, n_t\right) n_t + F_k\left(k_t, n_t\right) k_t$$

and using the first-order condition with respect to n_t, $F_n\left(k_t, n_t\right) = w_t$,

$$F_k\left(k_t, n_t\right) k_t = F\left(k_t, n_t\right) - w_t n_t.$$

Since F is CRS, its partial derivatives are homogeneous of degree zero. Define $\kappa_t := k_t / n_t$, then

$$F_k\left(\kappa_t, 1\right) k_t = F\left(k_t, n_t\right) - w_t n_t.$$

We can therefore write the maximised profits as:

$$\pi = \sum_{t=0}^{\infty} p_t \left(F_k\left(\kappa_t, 1\right) k_t - \left(k_{t+1} - k_t\left(1-\delta\right)\right)\right).$$

[32]To see why, observe that, in the budget constraint, we have

$$\begin{aligned} &\cdots + p_t c_t + p_{t+1} c_{t+1} + \cdots \\ \Rightarrow &\cdots + c_t + \frac{p_{t+1}}{p_t} c_{t+1} + \cdots \end{aligned}$$

and so p_{t+1}/p_t can be thought of as the discount factor for consumption good; i.e. real interest rate.

Expanding the summation:

$$
\begin{aligned}
&\sum_{t=0}^{\infty} p_t \left(F_k \left(\kappa_t, 1 \right) k_t - \left(k_{t+1} - k_t \left(1 - \delta \right) \right) \right) \\
&= p_0 \left[F_k \left(\kappa_0, 1 \right) k_0 - k_1 + k_0 \left(1 - \delta \right) \right] + p_1 \left[F_k \left(\kappa_1, 1 \right) k_1 - k_2 + k_1 \left(1 - \delta \right) \right] + \cdots \\
&= p_0 \left[F_k \left(\kappa_0, 1 \right) k_0 + k_0 \left(1 - \delta \right) \right] + \left(-p_0 k_1 + p_1 F_k \left(\kappa_1, 1 \right) k_1 + p_1 k_1 \left(1 - \delta \right) \right) - p_1 k_2 \cdots \\
&= p_0 \left[F_k \left(\kappa_0, 1 \right) k_0 + k_0 \left(1 - \delta \right) \right] + p_0 k_1 \left(-1 + \frac{p_1}{p_0} F_k \left(\kappa_1, 1 \right) + \frac{p_1}{p_0} \left(1 - \delta \right) \right) \\
&\qquad + p_1 k_2 \left(-1 + \frac{p_2}{p_1} F_k \left(\kappa_2, 1 \right) + \frac{p_2}{p_1} \left(1 - \delta \right) \right) + \cdots \\
&= p_0 \left[F_k \left(\kappa_0, 1 \right) k_0 + k_0 \left(1 - \delta \right) \right] + \sum_{t=0}^{\infty} p_t k_{t+1} \left(-1 + \frac{F_k \left(\kappa_{t+1}, 1 \right) + \left(1 - \delta \right)}{1 + r_t} \right).
\end{aligned}
$$

Normalising $p_0 = 1$ and noting that k_0 is given,

$$
\begin{aligned}
\pi &= k_0 \left(F_k \left(\kappa_0, 1 \right) + \left(1 - \delta \right) \right) \\
&\qquad + \sum_{t=0}^{\infty} p_t k_{t+1} \left(-1 + \frac{F_k \left(\kappa_{t+1}, 1 \right) + \left(1 - \delta \right)}{1 + r_t} \right).
\end{aligned}
$$

The first-order condition with respect to k_{t+1} can be written as

$$
\delta + r_t = F_k \left(\kappa_{t+1}, 1 \right)
$$

so that

$$
\begin{aligned}
\pi &= k_0 \left(F_k \left(\kappa_0, 1 \right) + \left(1 - \delta \right) \right) \\
&\qquad + \sum_{t=0}^{\infty} p_t k_{t+1} \left(-1 + \frac{\delta + r_t + \left(1 - \delta \right)}{1 + r_t} \right) \\
&= k_0 \left(F_k \left(\kappa_0, 1 \right) + \left(1 - \delta \right) \right).
\end{aligned}
$$

Hence, firms earn a profit in period 0 only (which accrues to the household who are the owners of the firm).

Consumer's problem:

$$
\begin{aligned}
\max_{\{c_t, n_t\}_{t=0}^{\infty} \in X} \quad & u \left(c_0, 1 - n_0, c_1, 1 - n_1, \ldots \right) \\
s.t. \quad & \sum_{t=0}^{\infty} p_t \left(c_t + w_t \ell_t \right) = \pi + \sum_{t=0}^{\infty} p_t w_t.
\end{aligned}
$$

The first-order condition with respect to c_t is

$$
u_{c_t} \left(\cdot \right) = \lambda p_t,
$$

where λ is the Lagrange multiplier on the constraint. Taking the ratio of first-order condition with respect to c_{t+1} and c_t

gives

$$\frac{u_{c_{t+1}}(\cdot)}{u_{c_t}(\cdot)} = \frac{p_{t+1}}{p_t} = \frac{1}{1+r_t}$$
$$\Rightarrow 1 = \frac{u_{c_{t+1}}(\cdot)}{u_{c_t}(\cdot)}(1+r_t).$$

The first-order condition with respect to n_t is

$$u_{\ell_t}(\cdot) = \lambda p_t w_t,$$

where $\ell_t \coloneqq 1 - n_t$. Taking the ratio between the first-order condition with respect to ℓ_t and c_t,

$$\frac{u_{\ell_t}(\cdot)}{u_{c_t}(\cdot)} = w_t.$$

10.1.4 Version 2: Households who own capital and rent to the firm

- ▷ Commodity space: Let $\mathbb{R}^{3T}$ for $T = \infty$. More formally,

$$L \coloneqq \left\{\{c_t, n_t, k_t\}_{t=0}^{\infty} : (c_t, n_t, k_t) \in \mathbb{R}^3,\ \forall t \geq 0\right\}.$$

- ▷ Production possibility set of the firm:

$$Y_t \coloneqq \{\{c_t, n_t, k_t\} : c_t \leq F(k_t, n_t)\}.$$

- ▷ Consumption possibility set:

$$X \coloneqq \{\{c_t, n_t, k_t\}_{t=0}^{\infty} : 0 \leq n_t \leq 1,\ x_t + c_t \leq F(k_t, n_t),\ k_{t+1} = x_t + k_t(1-\delta),\ \forall t \geq 0,\ k_0 \text{ given}\}.$$

- ▷ Household budget constraint:

$$\sum_{t=0}^{\infty} p_t (x_t + c_t + w_t \ell_t) = \sum_{t=0}^{\infty} p_t (w_t + \upsilon_t k_t),$$

  where p_t is the price of consumption good at time t, and $p_t w_t$ is the price of a unit of labour at time t, $p_t \upsilon_t$ is the rental price of capital at time t. Thus, υ_t is the time t rental price of capital in terms of time-t consumption goods.

Firm's problem:

$$\pi = \max_{\{\{c_t, n_t, k_t\} \in Y_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t (c_t - w_t n_t - \upsilon_t k_t).$$

Since π is strictly increasing in c_t, $c_t = F(k_t, n_t)$ at the optimal, so

$$\pi = \max_{\{\{n_t, k_t\} \in Y_t\}_{t=0}^{\infty} \in Y} \sum_{t=0}^{\infty} p_t (F(k_t, n_t) - w_t n_t - \upsilon_t k_t).$$

The first-order condition with respect to k_t and n_t are therefore

$$F_k(k_t, n_t) = \upsilon_t,$$
$$F_n(k_t, n_t) = w_t.$$

for all $t \geq 0$. Note that since F is CRS,

$$\begin{aligned} F(k_t, n_t) &= F_k(k_t, n_t) k_t + F_n(k_t, n_t) n_t \\ &= \upsilon_t k_t + w_t n_t. \end{aligned}$$

Hence, at the optimal,

$$\pi_t = F(k_t, n_t) - w_t n_t - \upsilon_t k_t = p_t(\upsilon_t k_t + w_t n_t - w_t n_t - \upsilon_t k_t) = 0, \ \forall t \geq 0.$$

Substituting $x_t = k_{t+1} - k_t(1-\delta)$ into the budget constraint yields

$$\begin{aligned} \sum_{t=0}^{\infty} p_t(w_t + \upsilon_t k_t) &= \sum_{t=0}^{\infty} p_t(k_{t+1} - k_t(1-\delta) + c_t + w_t \ell_t) \\ \sum_{t=0}^{\infty} p_t(w_t n_t - c_t) &= \sum_{t=0}^{\infty} p_t(k_{t+1} - k_t(1-\delta) - \upsilon_t k_t) \\ &= p_0(k_1 - k_0(1-\delta) - \upsilon_0 k_0) + p_1(k_2 - k_1(1-\delta) - \upsilon_1 k_1) + \cdots \\ &= -p_0 k_0(\upsilon_0 + (1-\delta)) + k_1(p_0 - p_1((1-\delta) + \upsilon_1)) + \cdots \\ &= -p_0 k_0(\upsilon_0 + (1-\delta)) + \sum_{t=0}^{\infty} k_{t+1} \underbrace{(p_t - p_{t+1}(1-\delta+\upsilon_{t+1}))}_{=0}. \end{aligned}$$

We will see from the consumer's problem that the second term on the right-hand side is zero so that budget constraint simplifies to (while normalising $p_0 = 1$)

$$\sum_{t=0}^{\infty} p_t c_t = k_0(\upsilon_0 + (1-\delta)) + \sum_{t=0}^{\infty} p_t w_t n_t.$$

We can think of the first term on the right-hand side as the consumer's initial asset endowment. We can define

$$a_0 = (\upsilon_0 + (1-\delta)).$$

Consumer's problem:

$$\begin{aligned} \max_{\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty} \in X} \quad & u(c_0, 1-n_0, c_1, 1-n_1, \ldots) \\ s.t. \quad & \sum_{t=0}^{\infty} p_t(k_{t+1} - k_t(1-\delta) + c_t + w_t \ell_t) = \sum_{t=0}^{\infty} p_t(w_t + \upsilon_t k_t). \end{aligned}$$

The first-order conditions with respect to c_t and n_t are as before:

$$\begin{aligned} \frac{u_{c_{t+1}}(\cdot)}{u_t(\cdot)}(1+r_t) &= 1, \\ \frac{u_{\ell_t}(\cdot)}{u_{c_t}(\cdot)} &= w_t. \end{aligned}$$

The first-order condition with respect to k_{t+1} is given by

$$\begin{aligned}p_t - p_{t+1}(1-\delta) - p_{t+1}\upsilon_{t+1} = 0,\\ \Leftrightarrow \frac{p_t}{p_{t+1}} = 1-\delta+\upsilon_{t+1}\\ \Leftrightarrow \upsilon_{t+1} = r_t + \delta.\end{aligned}$$

10.1.5 Comparing versions 1 and 2

Recall the first-order conditions from version 1 in which we assumed firms own capital:

$$\delta + r_t = F_k(k_{t+1}, n_{t+1}), \tag{10.1}$$
$$\begin{aligned}w_t &= F_n(k_t, n_t),\\ \pi &= k_0(F_k(\kappa_0, 1) + (1-\delta)),\\ 1 &= \frac{u_{c_{t+1}}(\cdot)}{u_{c_t}(\cdot)}(1+r_t),\\ w_t &= \frac{u_{\ell_t}(\cdot)}{u_{c_t}(\cdot)}.\end{aligned}$$

From version 2 in which firms rent capital from households:

$$\upsilon_t = F_k(k_t, n_t), \tag{10.2}$$
$$\upsilon_{t+1} = r_t + \delta, \tag{10.3}$$
$$\begin{aligned}w_t &= F_n(k_t, n_t),\\ a_0 &= p_0(k_0(1-\delta) + \upsilon_0),\\ 1 &= \frac{u_{c_{t+1}}(\cdot)}{u_{c_t}(\cdot)}(1+r_t),\\ w_t &= \frac{u_{\ell_t}(\cdot)}{u_{c_t}(\cdot)}.\end{aligned}$$

To see that they are equivalent, notice that rolling (10.2) one period forward and substituting (10.3) yields (10.1). We can also write $\upsilon_0 = F_k(k_0, n_0) = F_k(\kappa_0, 1)$ so that we realise that $\pi = a_0$. That is, when the firms own capital, the initial endowment accrues to the firms; when the households own the capital, the same endowment accrues to the households. Since the first-order conditions are equivalent between the two versions, we realise that equilibrium prices and aggregate quantities are identical between the two.

10.2 Steady states

Definition 10.1. An equilibrium is in an steady state if

$$\begin{aligned}r_t &= r, & \upsilon_t &= \upsilon, & x_t &= x.\\ w_t &= w, & \ell_t &= \ell,\\ c_t &= c, & k_t &= k\end{aligned}$$

for all $t \geq 0$ with $k_0 = k$. Alternatively, we can define a steady state as an initial condition k_0 such that, if the economy starts at that value, then it continues at that same value.

Suppose that u is additively separable with discount factor β and utility v; i.e.

$$u\left(c_0, 1-n_0, c_1, 1-n_1, \ldots\right) = \sum_{t=0}^{\infty} \beta^t v\left(c_t, 1-n_t\right). \tag{10.4}$$

In case preferences are time separable, consumer's problem yields

$$\begin{aligned} \beta \frac{v_c\left(c_{t+1}, \ell_{t+1}\right)}{v_c\left(c_t, \ell_t\right)}\left(1+r_t\right) &= 1, \\ \frac{v_\ell\left(c_t, \ell_t\right)}{v_c\left(c_t, \ell_t\right)} &= w_t. \end{aligned}$$

Hence, in the steady state where $c_t = c$, $\ell_t = \ell = 1-n$, and $w_t = w$,

$$\beta(1+r) = 1, \tag{10.5}$$

$$\frac{v_\ell(c, \ell)}{v_c(c, \ell)} = w. \tag{10.6}$$

From the firm's problem, in the steady state,

$$F_k(k, n) = \upsilon, \tag{10.7}$$

$$F_n(k, n) = w. \tag{10.8}$$

Finally, we have that

$$\upsilon = r + \delta, \tag{10.9}$$

$$x = \delta k, \tag{10.10}$$

$$c + x = F(k, n). \tag{10.11}$$

From now, we make the following assumption.

Assumption. *Assume that the utility function is additively separable as in (10.4), and that both consumption and leisure are normal goods in $v(c, 1-n)$.*

Proposition 10.1. *Assume that utility is additively separable as in (10.4), and that consumption is a normal good. Then, if $v(c, 1-n)$ is such that leisure is a normal good, then there is a unique steady state.*

Proof. Define ρ as

$$\frac{1}{1+\rho} \equiv \beta.$$

Then, using (10.5),

$$\frac{1+r}{1+\rho} = 1 \Rightarrow r = \rho.$$

Thus, (10.9) becomes

$$\upsilon = \rho + \delta. \tag{10.12}$$

Using (10.7) and the fact that F_k is homogeneous of degree zero,

$$F_k(\kappa, 1) = F_k(k, n) = \upsilon = \rho + \delta. \tag{10.13}$$

Thus, κ is pinned down by the equation above. We can then pin down w using (10.8) and the fact that F_n is homogenous of degree zero:

$$w = F_n(k, n) = F_n(\kappa, 1).$$

To pin down consumption, we use the market clearing condition (obtained by combining (10.10) and (10.11)), and the Euler's Theorem,

$$\begin{aligned} c + \delta k &= F(k, n) \\ &= F_k(k, n)\, k + F_n(k, n)\, n \\ &= \upsilon k + wn \\ &= (\rho + \delta)\, k + wn \\ \Leftrightarrow c &= \rho k + wn. \end{aligned} \tag{10.14}$$

where we also used (10.7), (10.8) and (10.12). If we know k and n, then we can pin down c. Since we know $\kappa = k/n$, we only need to pin down k.

To pin down k, consider the following problem:

$$\begin{aligned} \max_{c,\ell} \quad & v(c, \ell) \\ s.t. \quad & c + \ell w = \rho k + w, \end{aligned} \tag{10.15}$$

where the constraint is the market clearing condition, (10.14), with $\ell = 1 - n$ substituted in. This is the intratemporal problem that determines the relationship between consumption and leisure (and the constraint is the equilibrium budget constraint for that period). The first-order conditions are

$$\begin{aligned} v_c(c, \ell) = \lambda,\ & v_\ell(c, \ell) = \lambda w. \\ \Rightarrow \frac{v_\ell(c, \ell)}{v_c(c, \ell)} &= w, \end{aligned}$$

where λ is the Lagrange multiplier on the constraint. Observe that the optimality conditions coincides with the intratemporal condition (10.6). We can express this as

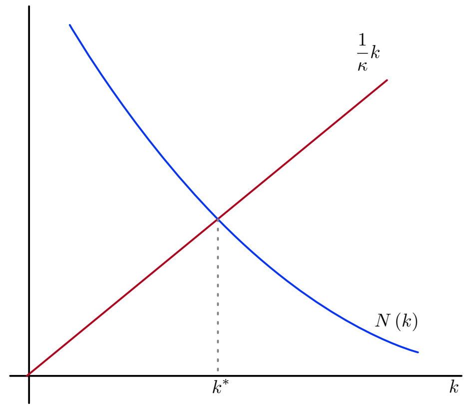
$$\frac{v_\ell(\rho k + wn, 1 - n)}{v_c(\rho k + wn, 1 - n)} = w.$$

Hence, we can use this equation, together with the fact that $\kappa = k/n = k/(1 - \ell)$ to pin down k.

We now find the value of k for which this is true. Denote the solution of the maximisation problem (10.15) above by $c = C(k)$ and $\ell = L(k)$. Since leisure is normal, L is strictly increasing, and hence $n = N(k) \coloneqq 1 - L(k)$ is strictly decreasing. In this notation, k is the solution to

$$k/n = \kappa \Rightarrow N(k) = \frac{1}{\kappa} k.$$

Since $N(0) > 0$ and N is strictly decreasing, while the right-hand side is strictly increasing in k, there is a unique solution k^*.



■

10.3 Government purchases

Suppose we add government purchases to the model. The feasibility constraint becomes

$$c_t + x_t + g_t = F(k_t, n_t).$$

We also suppose that government purchases are funded by lump-sum taxes levied on households, denoted τ_t per period. The household budget constraint is therefore

$$\sum_{t=0}^{\infty} p_t (x_t + c_t + w_t \ell_t) = \sum_{t=0}^{\infty} p_t (w_t + \upsilon_t k_t - \tau_t)$$

and the government budget constraint is

$$\sum_{t=0}^{\infty} p_t g_t = \sum_{t=0}^{\infty} p_t \tau_t.$$

The first-order conditions for firm's problem does not change, so that $\kappa = k/n$ is still pinned down by

$$F_k(\kappa, 1) = \rho + \delta$$

and w is pinned down by

$$w = F_n(\kappa, 1)$$

as before.

Suppose government purchases are constant; i.e. $g_t = g$. In the steady state, we must now have that

$$\begin{aligned} c + x + g &= F(k, n) \\ \Leftrightarrow c + \delta k + g &= (\rho + \delta) k + wn \\ \Leftrightarrow c &= wn + \rho k - g. \end{aligned}$$

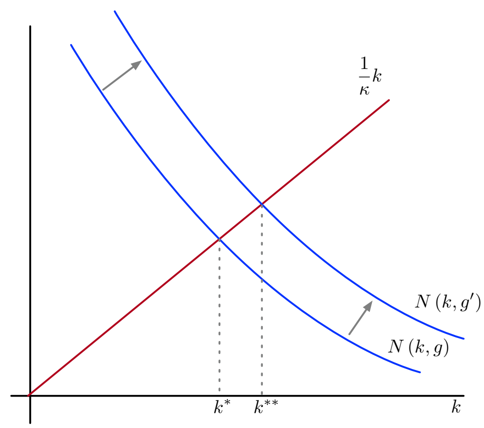
So if we know n and k, we can compute c. It remains to pin down k (or n) and we require an additional equation as before.

So consider again the household's intratemporal problem:

$$\begin{aligned}\max_{c,\ell} \quad & v(c,\ell) \\ s.t. \quad & c + \ell w = \rho k + w - g.\end{aligned} \tag{10.16}$$

The first-order condition for this problem, which is the same as before, gives us the additional equation necessary to pin down k.

To see how changes in g affects k, n and w, denote the solution of the maximisation problem (10.16) above by $c = C(k,g)$ and $\ell = L(k,g)$. Note that an increase in g has the effect of reducing wealth, so that ℓ is strictly decreasing in g. Hence, $N(k,g) = 1 - L(k,g)$ is strictly increasing in g. Thus, an increase in g to g' increases labour supply. As can be seen from the figure, this also implies a higher k in the steady state. Since κ does not change, this implies that k and ℓ change proportionately so that w does not change.



10.4 Neutral productivity shock

Suppose we replace the production function with $AF(k,n)$ with $A > 1$. Note that A increases productivity of both labour and capital equal—i.e. an increase in A represents a neutral technology shock. The feasibility constraint is now

$$c_t + x_t = AF(k,n).$$

Firm's first-order conditions are now given by

$$\begin{aligned}AF_k(\kappa,1) &= \upsilon, \\ AF_n(\kappa,1) &= w.\end{aligned} \tag{10.17}$$

Hence,

$$AF_k(\kappa,1) = \upsilon = \rho + \delta. \tag{10.18}$$

Notice that κ is now affected by A, as well as by ρ and δ. Using above, we can rewrite feasibility as

$$\begin{aligned} c + x &= AF_k(k,n)\,k + AF_n(k,n)\,n \\ &= (\rho+\delta)\,k + w(A)\,n \\ \Leftrightarrow c + \ell w(A) &= w(A) + \rho k. \end{aligned}$$

where $w(A)$ indicates that w's dependence on the technology parameter A.

Consider the following problem:

$$\begin{aligned} \max_{c,\ell} \quad & v(c,\ell) \\ s.t. \quad & c + lw(A) = \rho k + w(A). \end{aligned} \tag{10.19}$$

The first-order conditions are as before and the solution satisfies $c + \ell w(A) = \rho k + w(A)$. with k such that $k/n = k/(1-\ell) = \kappa(A)$.

Suppose A increases from A to A'. Consider what happens to the steady state values of r, w, $\kappa = k/n$ and k, as well as n:

- ▷ $r = \rho$—i.e. no change;
- ▷ from (10.18), we see that a higher A must be offset by a lower F_k. Since F is concave, this means that $\kappa(A)$ mst be higher;
- ▷ the effect on w is ambiguous—increase in κ leads to a fall in marginal productivity of labour; however, this is offset by the increase in the A parameter (see (10.17));
- ▷ since we do not know the effect of A on w, we do not know what happens to C or L (or N).

The ambiguity arises because a permanent change in A has both substitution effect and an income effect. In the case where we have preferences that are consistent with a balance growth path, leisure stays constant, since the substitution and the income effects exactly offset each other.

10.5 Labour-augmenting productivity shock

Now suppose that the production function is given by $F(k, An)$ with $A > 1$. A now represents labour-augmenting technology. The feasibility constraint is now

$$c_t + x_t = F(k, An).$$

Firm's first-order conditions are now given by

$$\begin{aligned} F_k\left(\frac{\kappa}{A}, 1\right) &= \upsilon, \\ AF_n\left(\frac{\kappa}{A}, 1\right) &= w \end{aligned} \tag{10.20}$$

so that

$$F_k\left(\frac{\kappa}{A}, 1\right) = \upsilon = \rho + \delta;$$

i.e. κ depends on A. However, we can see from above that if A increases, then κ must increase by the same proportion. That is, the ratio κ/A must remain constant. Hence, an increase in A leads to a higher w.

Consider the following problem:

$$\begin{aligned} \max_{c,\ell} \quad & v(c,\ell) \\ s.t. \quad & c + \ell w = \rho k + w. \end{aligned} \tag{10.21}$$

Denote the solution as $L(k,A)$ and define $N(k,A) = 1 - L(k,A)$. An increase in A leads to a higher w, which has both income and substitution effects, the effect on L (or N) is ambiguous in this case. However, since there is no substitution effect for consumption, we realise that C increases.

10.6 Infinitely elastic savings in the long run

Suppose we compute the steady state for economies with different:

- ▷ utility function u;
- ▷ production function F;
- ▷ depreciation rate δ;
- ▷ discount factor β (or discount rate ρ).

When is that the steady state interest rate change? Recall that the interest rate is given by (10.5). Hence, the interest rate change only when β or ρ changes. This means that the long-run (i.e. steady-state) elasticity of supply of capital (i.e. savings) is infinite.

10.7 Parametric example

Let the utility function and production functions be

$$\begin{aligned} v(c,\ell) &= \alpha \log c + (1-\alpha)\log \ell, \\ F(k,n) &= k^{\theta} n^{1-\theta}. \end{aligned}$$

We will solve for the steady state of this model assuming that government purchases, as a fraction of GNP, are equal to $\eta \coloneqq g/y$.

In this case, we have that

$$\begin{aligned} v_\ell(c,\ell) &= \frac{1-\alpha}{\ell}, \\ v_c(c,\ell) &= \frac{\alpha}{c}, \\ F_k(k,n) &= \theta \frac{y}{k} = \theta \left(\frac{k}{n}\right)^{\theta-1}, \\ F_n(k,n) &= (1-\theta)\frac{y}{n} = (1-\theta)\left(\frac{k}{n}\right)^{\theta}. \end{aligned}$$

In this case, the steady-state equations become

$$
\begin{aligned}
\rho &= r, \\
\frac{1-\alpha}{\alpha}\frac{c}{\ell} &= w, && (10.22) \\
\theta \frac{y}{k} &= \upsilon, \\
(1-\theta)\frac{y}{n} &= w, && (10.23) \\
\upsilon &= r + \delta, \\
x &= \delta k, \\
c + x + g &= y.
\end{aligned}
$$

The following results are immediate:

$$
\begin{aligned}
&\frac{k}{y} = \frac{\theta}{\upsilon} = \frac{\theta}{\rho+\delta}, \\
&\frac{x}{k} = \delta, \\
&\frac{x}{y} = \frac{x}{k}\frac{k}{y} = \frac{\delta\theta}{\rho+\delta}, \\
&\frac{c}{y} = 1 - \frac{x}{y} - \frac{g}{y} = 1 - \frac{\delta\theta}{\rho+\delta} - \eta.
\end{aligned}
$$

For labour, equating (10.22) and (10.23) gives

$$
\begin{aligned}
(1-\theta)\frac{y}{n} &= \frac{1-\alpha}{\alpha}\frac{c}{\ell} \\
\Rightarrow (1-\theta)(1-n) &= n\frac{1-\alpha}{\alpha}\frac{c}{y} \\
\Rightarrow (1-\theta) &= n\left(\frac{1-\alpha}{\alpha}\frac{c}{y} + (1-\theta)\right) \\
\Rightarrow n &= \frac{\alpha(1-\theta)}{(1-\alpha)\left(1 - \frac{\delta\theta}{\rho+\delta} - \eta\right) + \alpha(1-\theta)}.
\end{aligned}
$$

Let

$$
\begin{aligned}
&\alpha = \frac{1}{3}, && \theta = 0.3, \\
&\delta = 0.075, && \rho = 0.075. \\
&\eta = 0.15.
\end{aligned}
$$

($\eta = 0.15$ means that government purchases represent 15% of GNP in the steady state). Then,

$$\begin{aligned} r &= 0.075, & v &= 0.15 \\ \frac{k}{y} &= 2, & \frac{x}{k} &= 0.075, \\ \frac{x}{y} &= 0.15, & \frac{c}{y} &= 0.7. \\ n &= \frac{1}{3}, \end{aligned}$$

11 Linear taxes in Neoclassical model

We continue to assume that utility is additively separable; however, suppose that linear taxes are levied on labour and (net of depreciation) capital income, denoted $\tau_{\ell t}$ and τ_{kt} respectively. We will analyse the steady state as a function of $g = g_t$, $\tau_k = \tau_{kt}$ and $\tau_\ell = \tau_{\ell t}$ for all $t \geq 0$. We assume that the lump-sum taxes τ adjust to satisfy the government's budget constraint. We further assume that v is twice continuously differentiable, strictly concave and strictly increasing in (c, ℓ), and that it satisfies the standard Inada conditions so that we may focus on interior solutions.

The consumer's budget constraint is now

$$\sum_{t=0}^{\infty} p_t \left(c_t + x_t + \tau_t\right) = \sum_{t=0}^{\infty} p_t \left[\left(1 - \tau_{\ell t}\right) w_t n_t + k_t \upsilon_t - \tau_{kt} \left(\upsilon_t - \delta\right) k_t\right], \tag{11.1}$$

where $\tau_{\ell t}$ is the tax rate on labour income and τ_{kt} is the tax rate on net (of depreciation) capital.[33] We also have a lump-sum tax τ_t. These are all denominated in time-t units of the consumption good. As before p_t is the Arrow-Debreu price of time-t consumption goods in terms of time-0 consumption goods, and w_t and υ_t are the (pre-tax) real wage and rental rate of capital in terms of time-t consumption goods.

The firm's problem is given by

$$\max_{\{k_t, n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t \left(F\left(k_t, n_t\right) - w_t n_t - \upsilon_t k_t\right).$$

The government's budget constraint is

$$\sum_{t=0}^{\infty} p_t \left(\tau_t + \tau_{\ell t} w_t n_t + \tau_{kt} \left(\upsilon_t - \delta\right) k_t\right) = \sum_{t=0}^{\infty} p_t g_t,$$

where g_t denotes government purchases in period t (in units of time-t consumption good).

The agent's problem is given by

$$\max_{\{c_t, \ell_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t v\left(c_t, \ell_t\right)$$

subject to the budget constraint (11.1). Letting λ be the Lagrange multiplier on the budget constraint, the first-order conditions with respect to c_t and $n_t = 1 - \ell_t$ are given by

$$\begin{aligned} v_c\left(c_t, \ell_t\right) &= \lambda p_t, \\ v_\ell\left(c_t, \ell_t\right) &= \lambda p_t \left(1 - \tau_{\ell t}\right) w_t. \end{aligned}$$

Hence,

$$\beta \frac{v_c\left(c_{t+1}, \ell_{t+1}\right)}{v_c\left(c_t, \ell_t\right)} = \frac{p_{t+1}}{p_t} = \frac{1}{1 + r_t}, \tag{11.2}$$

$$\frac{v_\ell\left(c_t, \ell_t\right)}{v_c\left(c_t, \ell_t\right)} = \left(1 - \tau_{\ell t}\right) w_t. \tag{11.3}$$

[33] In this set up, time-t income from renting capital equals

$$\upsilon_t k_t - \tau_{kt} \left(\upsilon_t - \delta\right) k_t = \left(1 - \tau_{kt}\right) \upsilon_t k_t + \tau_{kt} \delta k_t.$$

Hence, capital tax is not levied on depreciation.

For the firm's problem, the first-order conditions with respect to k_t and n_t are

$$F_k\left(k_t, n_t\right) = \upsilon_t, \tag{11.4}$$
$$F_n\left(k_t, n_t\right) = w_t; \tag{11.5}$$

i.e. they are the same as before.

Using the household's budget constraint, we can find an expression for υ_{t+1} in terms of δ, r_t and τ_{kt+1} that must hold if agents find it optimal to choose $k_{t+1} \in (0, \infty)$. Substituting the law of motion for capital to eliminate x_t from the household budget constraint:

$$\begin{aligned}
\sum_{t=0}^{\infty} p_t\left(c_t + x_t + \tau_t\right) &= \sum_{t=0}^{\infty} p_t\left[\left(1-\tau_{\ell t}\right) w_t n_t + k_t \upsilon_t - \tau_{kt} k_t\left(\upsilon_t - \delta\right)\right] \\
\Leftrightarrow \sum_{t=0}^{\infty} p_t\left(c_t + \tau_t - \left(1-\tau_{\ell t}\right) w_t n_t\right) &= \sum_{t=0}^{\infty} p_t\left[\left(1-\tau_{kt}\right) \upsilon_t k_t + \tau_{kt}\delta k_t - x_t\right] \\
&= \sum_{t=0}^{\infty} p_t\left[\left(1-\tau_{kt}\right) \upsilon_t k_t + \tau_{kt}\delta k_t - \left(k_{t+1} - \left(1-\delta\right) k_t\right)\right] \\
&= \sum_{t=0}^{\infty} p_t\left[\left(1-\tau_{kt}\right) \upsilon_t k_t + \tau_{kt}\delta k_t + \left(1-\delta\right) k_t\right] - \sum_{t=0}^{\infty} p_t k_{t+1} \\
&= \sum_{t=0}^{\infty} p_{t+1} k_{t+1}\left[\left(1-\tau_{kt+1}\right) \upsilon_{t+1} + \tau_{kt+1}\delta + \left(1-\delta\right)\right] - \sum_{t=0}^{\infty} p_t k_{t+1} \\
&\quad + p_0 k_0\left[\left(1-\tau_{k0}\right) \upsilon_0 + \tau_{k0}\delta + \left(1-\delta\right)\right] \\
&= \sum_{t=0}^{\infty} p_{t+1} k_{t+1} \underbrace{\left[\left(1-\tau_{kt+1}\right) \upsilon_{t+1} + \tau_{kt+1}\delta + \left(1-\delta\right) - \frac{p_t}{p_{t+1}}\right]}_{=0} \\
&\quad + p_0 k_0\left[\left(1-\tau_{k0}\right) \upsilon_0 + \tau_{k0}\delta + \left(1-\delta\right)\right].
\end{aligned}$$

Focusing on the term equals to zero,

$$\begin{aligned}
1 + r_t = \frac{p_t}{p_{t+1}} &= \left(1-\tau_{kt+1}\right) \upsilon_{t+1} + \tau_{kt+1}\delta + \left(1-\delta\right) \\
\Leftrightarrow r_t &= \left(1-\tau_{kt+1}\right)\left(\upsilon_{t+1} - \delta\right).
\end{aligned} \tag{11.6}$$

Now suppose that tax rates are constant over time, $\tau_{kt} = \tau_k$, $\tau_{\ell t} = \tau_\ell$ for all t. Let $\kappa \coloneqq k/n$ and β be the time discount factor such that $\beta = 1/\left(1+\rho\right)$. We can then characterise the steady state with the following set of equations. From (11.2), in the steady state

$$\begin{aligned}
\beta &= \frac{1}{1+\rho} = \frac{1}{1+r} \\
\Leftrightarrow r &= \rho
\end{aligned}$$

so that we still have an infinitely elastic supply of savings in the steady state. From (11.6) and using $r = \rho$,

$$\upsilon = \frac{\rho}{1-\tau_k} + \delta.$$

Hence, (11.4) becomes

$$F_k(\kappa, 1) = \frac{\rho}{1-\tau_k} + \delta.$$

(11.5) remains unchanged. (11.3) becomes

$$\frac{v_\ell(c,\ell)}{v_c(c,\ell)} = w(1-\tau_\ell).$$

From law of motion for capital,

$$x = \delta k = \delta\kappa(1-\ell).$$

Feasibility can then be written as

$$\begin{aligned} F(k,n) &= c + g + x \\ &= c + g + \delta\kappa(1-\ell) \\ \Leftrightarrow c &= (1-l)(F(\kappa,1) - \delta\kappa) - g. \end{aligned}$$

To summarise, we have the following equations:

$$\begin{aligned} r &= \rho, \\ F_k(\kappa,1) &= \frac{\rho}{1-\tau_k} + \delta, \\ w &= F_n(\kappa,1), \qquad (11.7) \\ \frac{v_\ell(c,\ell)}{v_c(c,\ell)} &= w(1-\tau_\ell), \qquad (11.8) \\ c &= (1-\ell)(F(\kappa,1) - \delta\kappa) - g, \\ x &= \delta\kappa(1-\ell) \\ k &= \kappa(1-\ell). \end{aligned}$$

Note, in particular, that the taxes have created some wedges:

- ▷ firms pay wages equal to marginal product of labour, $w = F_n(\kappa,1)$. However, consumers only get $(1-\tau_\ell)w$; i.e. there is now a wedge between what firms perceive and what agents perceive from the labour-income tax.
- ▷ with respect to capital, is also a wedge created by the capital-income tax: firms equate net marginal product of labour $F_k(\kappa,1) - \delta$ to $\rho/(1-\tau_k)$, whereas the return for consumers is given by ρ.

With taxes in place, we no longer have that the marginal rate of technical substitution equals the marginal rate of substitution due to the wedges.

11.1 Characterisation of normal goods

Consider the optimal choice for a static problem with utility v for two levels of income with constant relative prices.

Definition 11.1. If c and ℓ are normal goods, then

$$\frac{\partial}{\partial c}\left(\frac{v_\ell(c,\ell)}{v_c(c,\ell)}\right) > 0,$$
$$\frac{\partial}{\partial \ell}\left(\frac{v_\ell(c,\ell)}{v_c(c,\ell)}\right) < 0.$$

That is, the marginal rate of substitution between ℓ and c, v_ℓ/v_c, is strictly increasing in c and strictly decreasing in ℓ.

Proposition 11.1. *If c and ℓ are normal goods, then we can define a function $\phi(\ell,\omega)$ such that*

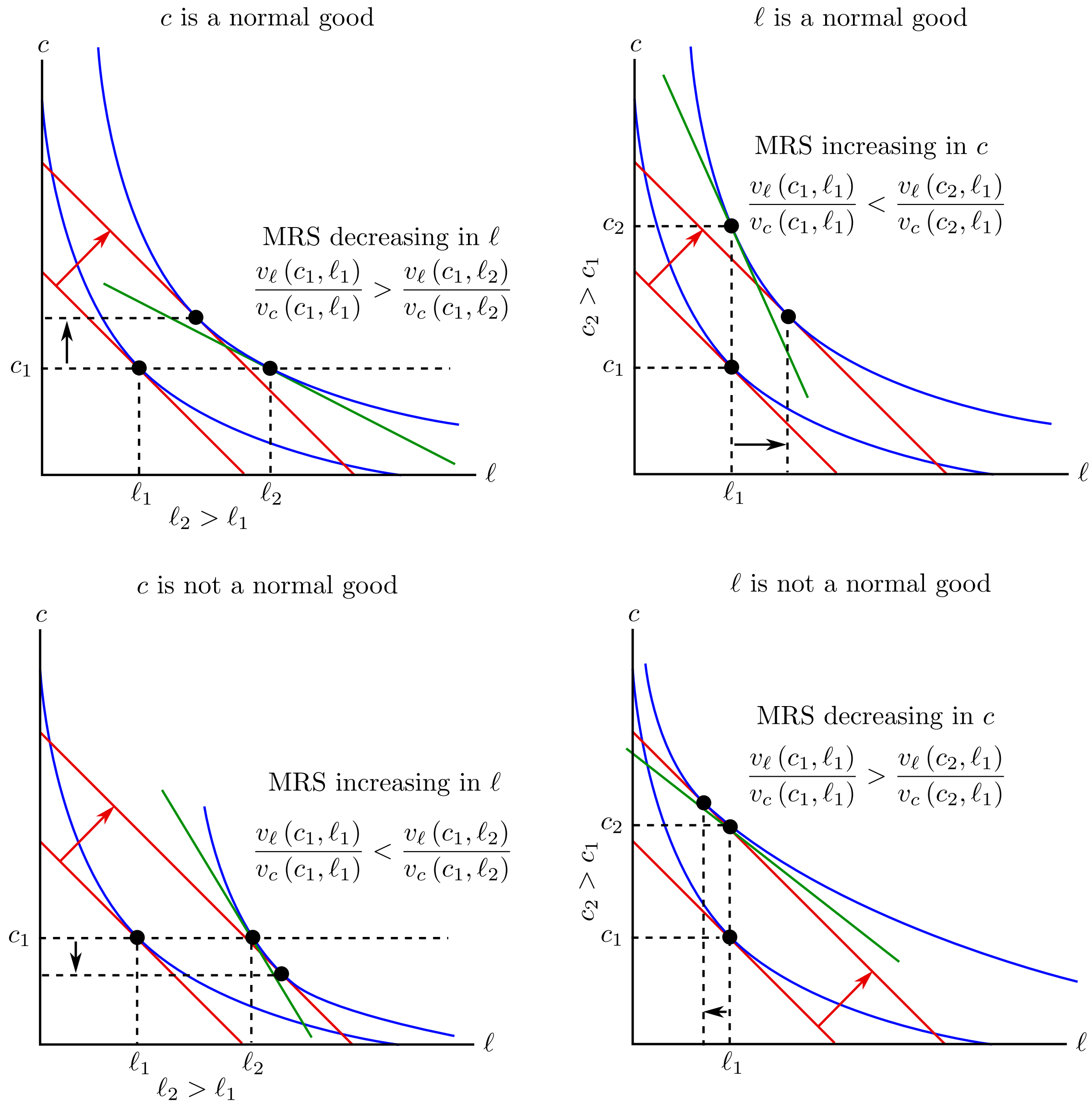
$$\frac{v_\ell(\phi(\ell,\omega),\ell)}{v_c(\phi(\ell,\omega),\ell)} = \omega,$$

and ϕ is increasing in ω and ℓ.

Why do these condition mean that c and ℓ are normal goods? Consider the figure below (we forget for now that we normalised time endowment to one) in which we increase income without changing the relative price of goods c_1 and c_2. It's useful to consider first the circumstances in which c and ℓ are not normal goods (i.e. inferior goods).

- ▷ Bottom left-hand panel shows the case in which an increase in income leads to lower consumption (but greater labour); i.e. c is an inferior good (while ℓ is a normal good). In this case, fixing c, we see that the marginal rate of substitution, v_ℓ/v_c is increasing in ℓ. In contrast, in the top left-hand panel, c is a normal good and we see that the MRS is decreasing in ℓ.
- ▷ The bottom right-hand panel shows the case in which ℓ is an inferior good. We can see that, fixing ℓ, MRS is decreasing in c. In contrast, in the top right-hand panel, ℓ is a normal good and we see that the MRS is increasing in c.

c and ℓ are normal goods



Proof. (*Proposition 11.1*) Since c and ℓ are normal goods

$$\sigma(x,\ell) \coloneqq \frac{v_\ell(x,\ell)}{v_c(x,\ell)}$$

is strictly increasing in x. Thus, ϕ is well-defined and there is a unique solution for ϕ:

$$\sigma(\phi,\ell) = \omega$$

for each (ω,ℓ). That there is a solution follows from the Inada conditions (since v_ℓ and v_c goes from 0 to infinity).

To show that ϕ is increasing, take the derivative of the expression with respect to ℓ,

$$\frac{\partial}{\partial c}\left(\frac{v_\ell(\phi(\ell,\omega),\ell)}{v_c(\phi(\ell,\omega),\ell)}\right)\frac{\partial\phi(\ell,\omega)}{\partial\ell} + \frac{\partial}{\partial\ell}\left(\frac{v_\ell(\phi(\ell,\omega),\ell)}{v_c(\phi(\ell,\omega),\ell)}\right) = 0.$$

Rearranging gives

$$\frac{\partial \phi\left(\ell, \omega\right)}{\partial \ell} = \frac{-\frac{\partial}{\partial \ell}\left(\frac{v_\ell}{v_c}\right)}{\frac{\partial}{\partial c}\left(\frac{v_\ell}{v_c}\right)} > 0.$$

Taking the derivative with respect to ω instead gives

$$\frac{\partial}{\partial c}\left(\frac{v_\ell\left(\phi\left(\ell, \omega\right), \ell\right)}{v_c\left(\phi\left(\ell, \omega\right), \ell\right)}\right)\frac{\partial \phi\left(\ell, \omega\right)}{\partial \omega} = 1.$$

Rearranging yields

$$\frac{\partial \phi\left(\ell, \omega\right)}{\partial \omega} = \frac{1}{\frac{\partial}{\partial c}\left(\frac{v_\ell}{v_c}\right)} > 0.$$ ■

Exercise 11.1. Suppose that

$$v\left(c, l\right) = \frac{\left(c^{1-\alpha}\ell^{\alpha} - 1\right)^{1-\gamma}}{1-\gamma}.$$

Find an expression for $\phi\left(\ell, \omega\right)$ in this case.

Solution. The partial derivatives are given by

$$\begin{aligned} v_c\left(c, \ell\right) &= \left(1-\alpha\right)\left(\frac{\ell}{c}\right)^{\alpha}\left(c^{1-\alpha}\ell^{\alpha} - 1\right)^{-\gamma}, \\ v_\ell\left(c, \ell\right) &= \alpha\left(\frac{c}{\ell}\right)^{1-\alpha}\left(c^{1-\alpha}\ell^{\alpha} - 1\right)^{-\gamma}. \end{aligned}$$

Then,

$$\frac{v_\ell\left(c, \ell\right)}{v_c\left(c, \ell\right)} = \frac{1-\alpha}{\alpha}\frac{c}{\ell}.$$

Hence,

$$\begin{aligned} \omega &= \frac{1-\alpha}{\alpha}\frac{c}{\ell} \\ \Rightarrow c = \phi\left(\ell, \omega\right) &= \frac{\alpha}{1-\alpha}\omega\ell. \end{aligned}$$

Clearly, $\phi\left(\ell, \omega\right)$ is increasing in both arguments.

11.2 Characterisation of the steady state

Consider the following system of two equations for the steady state (the first is a combination of (11.7) and (11.8)):

$$\begin{aligned} \frac{v_\ell\left(c, \ell\right)}{v_c\left(c, \ell\right)} &= F_n\left(\kappa, 1\right)\left(1-\tau_\ell\right), \\ c &= \left(1-\ell\right)\left(F\left(\kappa, 1\right) - \delta\kappa\right) - g. \end{aligned} \tag{11.9}$$

The first equates the marginal rate of substitution with post-tax wages, and the second is the market clearing condition.

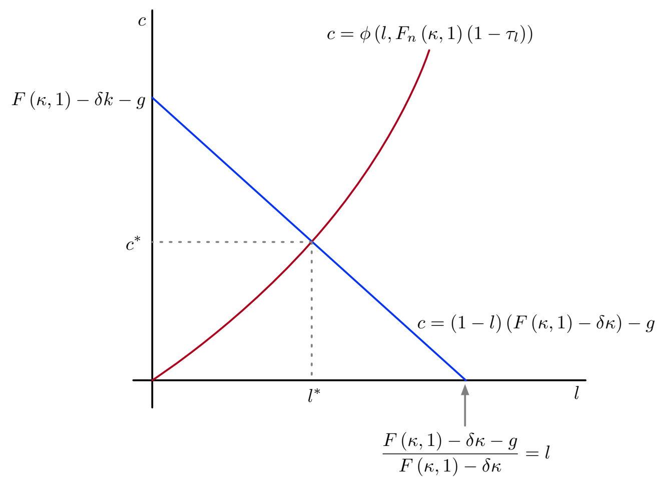
For given values of κ, g and τ_ℓ, we have two equations with two unknowns, c and ℓ. If $v\left(c, \ell\right)$ is such that c and ℓ are both normal goods, then, by Proposition (11.1), we can find $\phi\left(\ell, \omega\right)$ such that

$$\frac{v_\ell\left(\phi\left(\ell, \omega\right), \ell\right)}{v_c\left(\phi\left(\ell, \omega\right), \ell\right)} = \omega,$$

where $\omega = F_n(\kappa, 1)(1 - \tau_\ell)$. In other words,

$$c = \phi\left(\ell, F_n(\kappa, 1)(1 - \tau_\ell)\right). \tag{11.10}$$

Recall that ϕ is strictly increasing in both arguments, in particular, in ℓ. In contrast, (11.9) is strictly decreasing in ℓ (and also linear in ℓ). Hence, there is a unique solution for ℓ (and c) that solves the system of equations. We denote the solutions for ℓ and c respectively as $L(\kappa, g, \tau_\ell)$ and $C(\kappa, g, \tau_\ell)$. Note that (11.10) goes through the origin if it were to be consistent with a balanced growth path. check +



11.3 Comparative static on the steady state

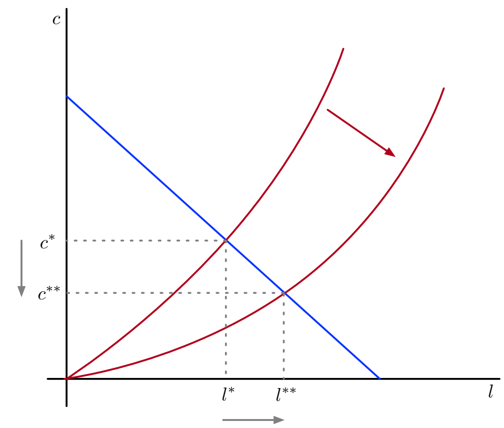
Recall that we have the following system of equations:

$$c = (1 - \ell)(F(\kappa, 1) - \delta\kappa) - g, \tag{11.11}$$

$$c = \phi\left(\ell, F_n(\kappa, 1)(1 - \tau_\ell)\right). \tag{11.12}$$

11.3.1 With respect to τ_ℓ

(11.11) is unaffected by τ_ℓ. Since ϕ is strictly increasing in its arguments, we realise that ϕ is strictly decreasing in τ_ℓ. Suppose τ_ℓ increases to τ_ℓ', then as can be seen from the figure below, (11.12) shifts down. Thus, a higher τ_ℓ leads to an increase in the steady-state leisure (ℓ) and a decrease in the steady-state consumption (c). That is, $L(\kappa, g, \tau_\ell)$ is strictly increasing in τ_ℓ while $C(\kappa, g, \tau_\ell)$ is strictly decreasing in τ_ℓ.



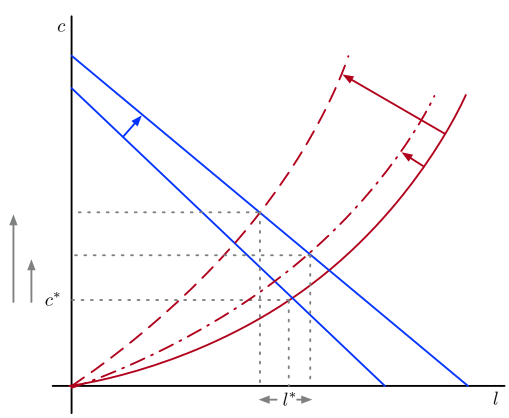
Note that we have no income effect here since we did not change g. Also, a higher steady-state ℓ (leisure) implies lower output.

11.3.2 With respect to κ

An increase in κ shifts (11.12) up since ϕ is strictly increasing in its arguments. What about (11.11)? Taking the derivative with respect to κ,

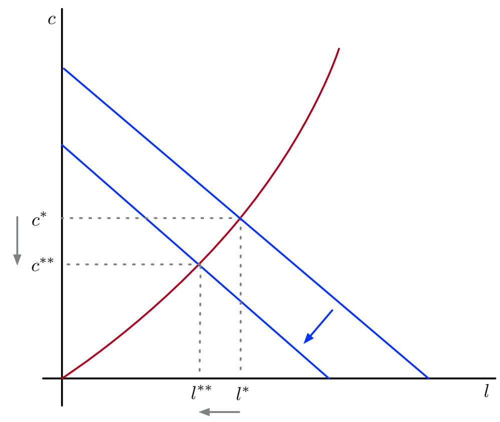
$$\begin{aligned}\frac{\partial c}{\partial \kappa} &= (1-\ell)\left(F_k(\kappa,1)-\delta\right) \\ &= (1-\ell)\left(\frac{\rho}{1-\tau_k}+\delta-\delta\right) = (1-\ell)\frac{\rho}{1-\tau_k} > 0.\end{aligned}$$

Hence, (11.11) also shifts up as κ increases. As can be seen from the figure, we can see that steady-state level of consumption will increase; however, whether ℓ will increase or decrease is ambiguous. Hence, although we can conclude that $C(\kappa, g, \tau_\ell)$ is strictly increasing in κ, we cannot be conclusive about how $L(\kappa, g, \tau_\ell)$ changes with κ.



11.3.3 With respect to g

(11.12) is unaffected by g while (11.11) shifts down when g increases. Hence, we can immediately conclude that $L(\kappa, g, \tau_\ell)$ and $C(\kappa, g, \tau_\ell)$ are strictly decreasing in g.



11.3.4 With respect to τ_k

Define the following functions,

$$C^*\left(\tau_k, \tau_\ell, g\right) \coloneqq C\left(\kappa^*\left(\tau_k\right), g, \tau_\ell\right),$$
$$L^*\left(\tau_k, \tau_\ell, g\right) \coloneqq L\left(\kappa^*\left(\tau_k\right), g, \tau_\ell\right),$$

where $\kappa^*\left(\tau_k\right)$ solves

$$F_k\left(\kappa^*\left(\tau_k\right), 1\right) = \delta + \frac{\rho}{1-\tau_k}.$$

Then, if τ_k increases, the right-hand side increases so that $\kappa^*\left(\tau_k\right)$ must fall. From before, we know that C would increase, but we cannot be certain about how L changes.

11.3.5 Summary

Maintain the assumption that c and ℓ are normal goods, we have the following comparative static results.

Steady-state quantities \ Policy parameter	τ_ℓ	τ_k	g
Interest rate (r)	$=$	$=$	$=$
Pre-tax wages (w) (F_n is increasing in κ)	$=$	$-$	$=$
Post-tax wages $(w\left(1-\tau_\ell\right))$	$-$	$-$	$=$
Capital-labour ratio (κ)	$=$	$-$	$=$
Pre-tax rental rate of capital $(\upsilon = \rho/(1-\tau_k)+\delta)$	$=$	$+$	$=$
Post-tax rental rate of capital ($(\upsilon-\delta)\left(1-\tau_k\right)$, υ given above)	$=$	$=$	$=$
Consumption (c)	$-$	$-$	$-$
Labour supply $(n = 1-\ell)$	$-$	?	$+$
Capital stock $(k = \kappa n = \kappa\left(1-\ell\right))$	$-$	?	$+$
Investment $(x = \delta k = \delta\kappa\left(1-\ell\right))$	$-$	?	$+$

12 Uzawa Model

The model we analyse in this subsection is an example of utility function with an endogenous discount rate.

12.1 Preliminaries

Recall that we assume

$$u\left(c_0, \ell_0, c_1, \ell_1, \ldots\right) = \sum_{t=0}^{\infty} \beta^t v\left(c_t, \ell_t\right).$$

How do we interpret this?

Infinite horizon Suppose that a period corresponds to a generation, and that parents are altruistic towards their children. We can capture this by assuming that parents care about their own consumption, as well as their children's well-being. In this case, we can write the utility of generation t as

$$U_t = v\left(c_t, \ell_t\right) + \beta U_{t+1}.$$

Solving forward,

$$\begin{aligned}
U_t &= v\left(c_t, \ell_t\right) + \beta v\left(c_{t+1}, \ell_{t+1}\right) + \beta U_{t+2} \\
&= v\left(c_t, \ell_t\right) + \beta v\left(c_{t+1}, \ell_{t+1}\right) + \beta^2 v\left(c_{t+2}, \ell_{t+2}\right) + \beta^2 U_{t+3} \\
&\ \ \vdots \\
&= \sum_{s=0}^{\infty} \beta^s v\left(c_{t+s}, \ell_{t+s}\right).
\end{aligned}$$

This type of rationalisation of an infinite horizon is called Barro bequests.

Independence from past consumption (additive separability) A feature of the discounted utility above is that it makes the marginal rate of substitution between consumption at dates t' and t, for $t' > t$, independent of the consumption at dated before t, which helps to simplify analysis. This feature will be shared for any utility function that satisfies

$$U_t = A\left(c_t, \ell_t, U_{t+1}\right),$$

where, again, we obtain U_t as a function of $\left(c_t, \ell_t, c_{t+1}, \ell_{t+1}, \cdots\right)$ by repeated substitution

$$\begin{aligned}
U_t &= A\left(c_t, \ell_t, U_t\left(c_{t+1}, \ell_{t+1}\right)\right) \\
&= A\left(c_t, \ell_t, A\left(c_{t+1}, \ell_{t+1}, U_{t+1}\left(c_{t+2}, \ell_{t+2}\right)\right)\right) \\
&\quad \cdots.
\end{aligned}$$

In this previous case, we had

$$A\left(c_t, \ell_t, U_{t+1}\right) = v\left(c_t, \ell_t\right) + \beta U_{t+1}.$$

12.2 Sequential and present-value budget constraints

Consider the following sequence of budget constraints:

$$q_t a_{t+1} + c_t + \ell_t w_t = w_t + a_t, \ \forall t \geq 0 \tag{12.1}$$

and a_0 given. We can interpret a_t as the period-t starting asset, q_t as the time-t price of a good at time $t+1$. Purchasing a_{t+1} assets at time t gives the agent a_{t+1} consumption goods at time $t+1$.

The following proposition establishes the equivalence between the present-value budget constraint and the sequential budget constraint (together with a transversality condition).

Proposition 12.1. *Let $q_t = 1/(1+r_t)$ and $\pi = a_0$. If $\{c_t, \ell_t\}$ satisfies the sequential budget constraints (12.1),*

$$q_t a_{t+1} + c_t + \ell_t w_t = w_t + a_t, \ \forall t \geq 0,$$

and a transversality condition,

$$\lim_{t\to\infty} a_{t+1} \prod_{s=0}^{t} q_s = 0.$$

Then $\{c_t, \ell_t\}$ satisfies the present-value budget constraint:

$$\sum_{t=0}^{\infty} p_t \left(c_t + w_t \ell_t\right) = \pi + \sum_{t=0}^{\infty} p_t w_t.$$

Conversely, if $\{c_t, \ell_t\}$ satisfies the present-value budget constraint, then it satisfies the sequential budget constraints.

Proof. To show that the present-value budget constraint implies sequential budget constraint, we can define the asset in period t as the present value of future consumption and earnings; i.e.

$$a_t = \frac{1}{p_t} \sum_{s=t}^{\infty} p_s \left(c_s + w_s \ell_s - w_s\right), \ \forall t \geq 0.$$

At $t = 0$, the equation is given by

$$a_0 = \pi = \frac{1}{p_0} \sum_{s=0}^{\infty} p_s \left(c_s + w_s \ell_s - w_s\right),$$

which is equal to the present-value budget constraint if we normalise p_0 to one. In period $t+1$, we have

$$a_{t+1} = \frac{1}{p_{t+1}} \sum_{s=t+1}^{\infty} p_s \left(c_s + w_s \ell_s - w_s\right). \tag{12.2}$$

Dividing by p_t converts the equation above in terms of period-t consumption good:

$$\frac{a_{t+1}}{p_t} = \frac{1}{p_t p_{t+1}} \sum_{s=t+1}^{\infty} p_s \left(c_s + w_s \ell_s - w_s\right). \tag{12.3}$$

Recalling that

$$q_t = \frac{1}{1+r_t} = \frac{p_{t+1}}{p_t} \Rightarrow \frac{1}{p_t} = \frac{q_t}{p_{t+1}}.$$

we can rewrite (12.3) as

$$\begin{aligned}\frac{q_t}{p_{t+1}}a_{t+1} &= \frac{q_t}{p_{t+1}}\frac{1}{p_{t+1}}\sum_{s=t+1}^{\infty} p_s\left(c_s + w_s\ell_s - w_s\right)\\ \Leftrightarrow q_t a_{t+1} &= \frac{q_t}{p_{t+1}}\sum_{s=t+1}^{\infty} p_s\left(c_s + w_s\ell_s - w_s\right).\end{aligned}$$

This shows that multiplying (12.2) by q_t converts the equation in terms of period-t good.

Subtracting a_t from $q_t a_{t+1}$, which gives the one period gain in asset,

$$q_t a_{t+1} - a_t = \frac{q_t}{p_{t+1}}\sum_{s=t+1}^{\infty} p_s\left(c_s + w_s\ell_s - w_s\right) - \frac{1}{p_t}\sum_{s=t}^{\infty} p_s\left(c_s + w_s\ell_s - w_s\right),$$

since $q_t = 1/\left(1+r_t\right) = p_{t+1}/p_t$,

$$\begin{aligned}q_t a_{t+1} - a_t &= \frac{1}{p_t}\sum_{s=t+1}^{\infty} p_s\left(c_s + w_s\ell_s - w_s\right) - \frac{1}{p_t}\sum_{s=t}^{\infty} p_s\left(c_s + w_s\ell_s - w_s\right)\\ &= -\frac{1}{p_t}\left(p_t\left(c_t + w_t\ell_t - w_t\right)\right)\\ &= -\left(c_t + w_t\ell_t - w_t\right)\\ \Rightarrow a_t &= q_t a_{t+1} + c_t + w_t\ell_t - w_t,\ \forall t \geq 0.\end{aligned}$$

Now, to show that the sequence budget constraint implies the present-value budget constraint, consider the period $t = 0$ budget first,

$$\begin{aligned}q_0 a_1 + \underbrace{c_0 + w_0\ell_0 - w_0}_{=z_0} &= a_0\\ q_0 a_1 + z_0 &= a_0, \qquad (12.4)\end{aligned}$$

where $z_t \coloneqq c_t + w_t\ell_t - w_t$. Period-1 and period-2 budget constraints are

$$\begin{aligned}q_1 a_2 + z_1 &= a_1,\\ q_2 a_3 + z_2 &= a_2.\end{aligned}$$

To convert this into period-t terms, we multiply the period-1 budget constraint by q_0 and period-2 budget constraint by $q_0 q_1$:

$$q_0 q_1 a_2 + q_0 z_1 = q_0 a_1, \qquad (12.5)$$

$$q_0 q_1 q_2 a_3 + q_0 q_1 z_2 = q_0 q_1 a_2. \qquad (12.6)$$

Summing (12.4), (12.5) and (12.6),

$$\begin{aligned}\left(q_0 q_1 q_2 a_3 + q_0 q_1 a_2 + q_0 a_1\right) + \left(q_0 q_1 z_2 + q_0 z_1 + z_0\right) &= a_0 + q_0 a_1 + q_0 q_1 a_2\\ \Rightarrow q_0 q_1 q_2 a_3 + q_0 q_1 z_2 + q_0 z_1 + z_0 &= a_0. \qquad (12.7)\end{aligned}$$

Note that

$$q_t = \frac{p_{t+1}}{p_t} \Rightarrow \prod_{s=0}^{T} q_s = q_0 q_1 \cdots q_T$$
$$= \frac{p_1}{p_0}\frac{p_2}{p_1}\cdots\frac{p_{T+1}}{p_T} = \frac{p_{T+1}}{p_0}.$$

Hence, we can write (12.7) as

$$a_0 = z_0\frac{p_0}{p_0} + z_1\frac{p_1}{p_0} + z_2\frac{p_2}{p_0} + a_3\frac{p_3}{p_0}$$
$$= \frac{p_3}{p_0}a_3 + \sum_{t=0}^{2}\frac{p_t}{p_0}z_t.$$

More generally, we have

$$a_0 = \frac{p_{T+1}}{p_0}a_{T+1} + \sum_{t=0}^{T}\frac{p_t}{p_0}z_t$$
$$= p_{T+1}a_{T+1} + \sum_{t=0}^{T} p_t\left(c_t + w_t\ell_t - w_t\right)$$

where, in the last line, we normalised $p_0 = 1$ and the definition of z_t. Then, taking limits,

$$\lim_{T\to\infty} a_0 = \lim_{T\to\infty} p_{T+1}a_{T+1} + \lim_{T\to\infty}\sum_{t=0}^{T} p_t\left(c_t + w_t\ell_t - w_t\right)$$
$$\Rightarrow a_0 = \pi = \sum_{t=0}^{\infty} p_t\left(c_t + w_t\ell_t - w_t\right)$$

since

$$\lim_{T\to\infty} p_{T+1}a_{T+1} = \lim_{T\to\infty}\left(q_0 q_1 \cdots q_T\right)a_{T+1} = 0$$

by assumption. ■

12.3 Uzawa's model

Consider the following function A:

$$U_t = A\left(c_t, \ell_t, U_{t+1}\right) = v\left(c_t, \ell_t\right) + e^{-(\rho + v(c_t,\ell_t))}U_{t+1},$$

where the discount rate depends on the utility level today. In particular, observe that, as agents become "richer" (i.e. higher v in period t), the agents become less patient.

Solving forward gives

$$
\begin{aligned}
U_t &= v\left(c_t, \ell_t\right) + e^{-\left(\rho + v\left(c_t, \ell_t\right)\right)} \Big(v\left(c_{t+1}, \ell_{t+1}\right) + e^{-\left(\rho + v\left(c_{t+1}, \ell_{t+1}\right)\right)} U_{t+2} \Big) \\
&= v\left(c_t, \ell_t\right) + e^{-\left(\rho + v\left(c_t, \ell_t\right)\right)} v\left(c_{t+1}, \ell_{t+1}\right) + e^{-\left[\left(\rho + v\left(c_t, \ell_t\right)\right) + \left(\rho + v\left(c_{t+1}, \ell_{t+1}\right)\right)\right]} U_{t+2} \\
&= v\left(c_t, \ell_t\right) + \exp\left[-\left(\rho + v\left(c_t, \ell_t\right)\right)\right] v\left(c_{t+1}, \ell_{t+1}\right) \\
&\qquad + \exp\left[-\left(\left(\rho + v\left(c_t, \ell_t\right)\right) + \left(\rho + v\left(c_{t+1}, \ell_{t+1}\right)\right)\right)\right] U_{t+2} \\
&= \quad \vdots \\
&= \sum_{s=0}^{\infty} \exp\left[-\sum_{i=0}^{s-1} \left(\rho + v\left(c_{t+i}, \ell_{t+i}\right)\right)\right] v\left(c_{t+s}, \ell_{t+s}\right),
\end{aligned}
$$

where $\sum_{i=0}^{-1} \left(\rho + v\left(c_{t+i}, \ell_{t+i}\right)\right) \coloneqq 0$.

Consider the problem of maximising the utility

$$
U_0 = \sum_{t=0}^{\infty} v\left(c_t\right) \exp\left[-\sum_{s=0}^{t-1} \left(\rho + \theta v\left(c_s\right)\right)\right]
$$

subject to

$$
\frac{1}{1+r} a_{t+1} + c_t = w + a_t,
$$

where we have fixed r and w, and suppressed labour to simplify the problem. Note that setting $\theta = 0$ gives the standard problem.

We can rewrite the sequential budget constraint as the present-value budget constraint:

$$
a_0 = \sum_{t=0}^{\infty} p_t \left(c_t - w\right),
$$

where w is the wage rate (we implicitly assume inelastic unit supply of labour). Let us fix the interest rate, i.e. $r_t = r$ for all $t \geq 0$, so that

$$
\frac{p_{t+1}}{p_t} = \frac{1}{1+r} \Rightarrow \frac{p_t}{p_0} = \left(\frac{1}{1+r}\right)^t.
$$

Then, the budget constraint becomes

$$
a_0 = \sum_{t=0}^{\infty} \left(\frac{1}{1+r}\right)^t \left(c_t - w\right).
$$

The Lagrangian for this problem is

$$
\mathcal{L} = \sum_{t=0}^{\infty} v\left(c_t\right) \exp\left[-\sum_{s=0}^{t-1} \left(\rho + \theta v\left(c_s\right)\right)\right] + \lambda \left[\sum_{t=0}^{\infty} \left(\frac{1}{1+r}\right)^t \left(w - c_t\right)\right].
$$

Before writing the first-order condition, note that

$$
\begin{aligned}
\frac{\partial U_0}{\partial c_t} = \cdots &+ \frac{\partial}{\partial c_t}\left(v\left(c_t\right)\exp\left[-\underbrace{\left(\rho+\theta v\left(c_0\right)+\cdots+\rho+\theta v\left(c_{t-1}\right)\right)}_{=\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)}\right]\right) \\
&\cdots + \frac{\partial}{\partial c_t}\left(v\left(c_{t+1}\right)\exp\left[-\underbrace{\left(\rho+\theta v\left(c_0\right)+\cdots+\rho+\theta v\left(c_t\right)+\rho+\theta v\left(c_{t+1}\right)\right)}_{=\sum_{s=0}^{t}\left(\rho+\theta v\left(c_s\right)\right)}\right]\right) \\
&\cdots + \frac{\partial}{\partial c_t}\left(v\left(c_{t+2}\right)\exp\left[-\underbrace{\left(\rho+\theta v\left(c_0\right)+\cdots+\rho+\theta v\left(c_t\right)+\rho+\theta v\left(c_{t+1}\right)+\rho+\theta v\left(c_{t+2}\right)\right)}_{=\sum_{s=0}^{t+1}\left(\rho+\theta v\left(c_s\right)\right)}\right]\right) \\
&= v'\left(c_t\right)\exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right] - v\left(c_{t+1}\right)\theta v'\left(c_t\right)\exp\left[-\sum_{s=0}^{t}\left(\rho+\theta v\left(c_s\right)\right)\right] \\
&\qquad - v\left(c_{t+2}\right)\theta v'\left(c_t\right)\exp\left[-\sum_{s=0}^{t+1}\left(\rho+\theta v\left(c_s\right)\right)\right] + \cdots \\
&= \exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right] v'\left(c_t\right) \\
&\quad - \theta\exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right] v'\left(c_t\right)\left(v\left(c_{t+1}\right)\frac{\exp\left[-\sum_{s=0}^{t}\left(\rho+\theta v\left(c_s\right)\right)\right]}{\exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right]}\right. \\
&\qquad \left. + v\left(c_{t+2}\right)\frac{\exp\left[-\sum_{s=0}^{t+1}\left(\rho+\theta v\left(c_s\right)\right)\right]}{\exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right]} + \cdots\right) \\
&= \exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right] v'\left(c_t\right) \\
&\quad - \theta\exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right] v'\left(c_t\right)\left(v\left(c_{t+1}\right)\exp\left[-\left(\rho+\theta v\left(c_t\right)\right)\right]\right. \\
&\qquad \left. + v\left(c_{t+2}\right)\exp\left[-\left(\left(\rho+\theta v\left(c_t\right)\right)+\left(\rho+\theta v\left(c_{t+1}\right)\right)\right)\right] + \cdots\right) \\
&= \exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right] v'\left(c_t\right) \\
&\quad \times\left[1-\theta\left(v\left(c_{t+1}\right)\exp\left[-\sum_{s=t}^{t}\left(\rho+\theta v\left(c_s\right)\right)\right] + v\left(c_{t+2}\right)\exp\left[\sum_{s=t}^{t+1}\left(\rho+\theta v\left(c_s\right)\right)\right] + \cdots\right)\right] \\
&= \exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right]\left(1-\theta\sum_{r=t+1}^{\infty} v\left(c_r\right)\exp\left[-\sum_{s=t}^{r-1}\left(\rho+\theta\left(c_s\right)\right)\right]\right) v'\left(c_t\right).
\end{aligned}
$$

Then, the first-order condition is given by

$$
\frac{\partial U_0}{\partial c_t} = \lambda\left(\frac{1}{1+r}\right)^t .
$$

The intertemporal marginal rate of substitution is then given by

$$\begin{aligned}
\frac{1}{1+r} &= \frac{\partial U_0/\partial c_{t+1}}{\partial U_0/\partial c_t} \\
&= \frac{\exp\left[-\sum_{s=0}^{t}\left(\rho+\theta v\left(c_s\right)\right)\right]\left(1-\theta\sum_{r=t+2}^{\infty}v\left(c_r\right)\exp\left[-\sum_{s=t+1}^{r-1}\left(\rho+\theta\left(c_s\right)\right)\right]\right)}{\exp\left[-\sum_{s=0}^{t-1}\left(\rho+\theta v\left(c_s\right)\right)\right]\left(1-\theta\sum_{r=t+1}^{\infty}v\left(c_r\right)\exp\left[-\sum_{s=t}^{r-1}\left(\rho+\theta\left(c_s\right)\right)\right]\right)}\frac{v'\left(c_{t+1}\right)}{v'\left(c_t\right)} \\
&= \exp\left[-\left(\rho+\theta v\left(c_t\right)\right)\right]\frac{1-\theta\sum_{r=t+2}^{\infty}v\left(c_r\right)\exp\left[-\sum_{s=t+1}^{r-1}\left(\rho+\theta\left(c_s\right)\right)\right]}{1-\theta\sum_{r=t+1}^{\infty}v\left(c_r\right)\exp\left[-\sum_{s=t}^{r-1}\left(\rho+\theta\left(c_s\right)\right)\right]}\frac{v'\left(c_{t+1}\right)}{v'\left(c_t\right)}.
\end{aligned}$$

Let us solve for the value of c such that $c_t = c$ for all $t \geq 0$ (i.e. steady state), and define discount factor as

$$\beta\left(c\right) \coloneqq \exp\left[-\left(\rho+\theta v\left(c\right)\right)\right].$$

Notice that

$$\begin{aligned}
\sum_{r=t+1}^{\infty}v\left(c\right)\exp\left[-\sum_{s=t}^{r-1}\left(\rho+\theta\left(c\right)\right)\right] &= v\left(c\right)\sum_{r=t+1}^{\infty}\beta\left(c\right)^{r-1-t} = v\left(c\right)\sum_{s=0}^{\infty}\beta\left(c\right)^{s} = \frac{v\left(c\right)}{1-\beta\left(c\right)}, \\
\sum_{r=t+2}^{\infty}v\left(c\right)\exp\left[-\sum_{s=t+1}^{r-1}\left(\rho+\theta\left(c\right)\right)\right] &= v\left(c\right)\sum_{r=t+2}^{\infty}\beta\left(c\right)^{r-2-t} = v\left(c\right)\sum_{s=0}^{\infty}\beta\left(c\right)^{s} = \frac{v\left(c\right)}{1-\beta\left(c\right)}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\left.\frac{\partial U_0/\partial c_{t+1}}{\partial U_0/\partial c_t}\right|_{c_t=c} &= \exp\left[-\left(\rho+\theta v\left(c\right)\right)\right]\frac{1-\theta\sum_{r=t+2}^{\infty}v\left(c\right)\exp\left[-\sum_{s=t+1}^{r-1}\left(\rho+\theta\left(c\right)\right)\right]}{1-\theta\sum_{r=t+1}^{\infty}v\left(c\right)\exp\left[-\sum_{s=t}^{r-1}\left(\rho+\theta\left(c\right)\right)\right]}\frac{v'\left(c\right)}{v'\left(c\right)} \\
&= \exp\left[-\left(\rho+\theta v\left(c\right)\right)\right] \\
&= \beta\left(c\right)
\end{aligned}$$

so that the intertemporal marginal rate of substitution simplifies to

$$\frac{1}{1+r} = \beta\left(c\right) = \exp\left[-\left(\rho+\theta v\left(c\right)\right)\right].$$

This contrasts with the standard case in which $\theta = 0$, when above expression did not depend upon c.

Now let $\hat{r}$ be such that

$$\frac{1}{1+r} = e^{-\hat{r}} \Rightarrow \hat{r} = \rho+\theta v\left(c\right). \tag{12.8}$$

Suppose $\theta = 0$, then c is determined by the budget constraint; i.e. c such that

$$a_0 = \sum_{t=0}^{\infty}\left(\frac{1}{1+r}\right)^{t}\left(c-w\right).$$

If $\theta > 0$, however, then (12.8) uniquely determines c (since v is strictly increasing).

Exercise 12.1. What is the interpretation of this equation in terms of the long-run distribution of consumption if there is heterogeneity in w?

Solution. [...] ???

12.3.1 Adding production

Let us know add a neoclassical production function $AF(k,n)$ to the set-up so far. We continue to assume that unit labour is supplied inelastically. It is still the case that, in the steady state,

$$\begin{aligned} \upsilon &= r+\delta, \\ F_k(k,1) &= \upsilon, \\ c+\delta k &= F(k,1). \end{aligned}$$

However, we now have that

$$\beta(c)(1+r) = 1.$$

We can combine the equations to obtain that

$$\beta(F(k,1)-\delta k)(1+F_k(k,1)-\delta) = 1.$$

Substituting the definition of $\beta(\cdot)$,

$$\begin{aligned} &\log\left[e^{-(\rho+\theta v(F(k,1)-\delta k))}(1+F_k(k,1)-\delta)\right] = \log(1) \\ &\Rightarrow \rho+\theta v(F(k,1)-\delta k) = \log(1-\delta+F_k(k,1)), \end{aligned}$$

where the left-hand side of the equation equals $\hat{r}$.

Since the left-hand side is increasing in k and the right-hand side is decreasing in k,[34] we realise that there is at most one solution.

Exercise 12.2. (*Productivity shock in the Uzawa model*) Suppose the production function $F(k,1)$ changes to $AF(k,1)$ for $A>1$ in the previous set up. What is the effect on the steady-state value of k and on the steady state interest rate r? Compare this case with the standard one where $\theta = 0$. What is the economic intuition behind the difference?

Solution. [...] ???

[34]Differentiating $F(k,1)-\delta k$ with respect to k yields $F_k(k,1)-\delta = \upsilon-\delta = r > 0$.

13 Ramsey Taxation

13.1 Neoclassical Growth Model: Deterministic Case

In this section, we study optimal taxation of capital and labour in a deterministic version of the Neoclassical growth model. We assume that the government has to fund a sequence of government purchases using linear taxation, either on capital and/or labour. We will find conditions under which we obtain the following two results:

- ▷ tax rates on capital should be zero;
- ▷ tax rates on labour income (or consumption) should be constant.

In doing so, we develop tools to study Ramsey linear taxation, namely the representation of an equilibrium with taxes by the implementability constraint.

13.1.1 The setup

Household's problem Let preferences be given by

$$\sum_{t=0}^{\infty} \beta^t U(c_t, n_t),$$

where c_t is consumption, n_t is time spent working. We assume that U is strictly concave in its two arguments (c, n), strictly increasing in c, strictly decreasing in n, and satisfy the standard Inada conditions.

The budget constraint of the household is

$$\sum_{t=0}^{\infty} p_t(c_t + k_{t+1}) = \sum_{t=0}^{\infty} p_t\left[(1-\tau_{\ell t}) w_t n_t + k_t R_{kt}\right],$$

and

$$R_{kt} = 1 + (1-\tau_{kt})(\upsilon_t - \delta), \ \forall t \geq 0,$$

where p_t is the Arrow-Debreu price of a consumption good at time t, w_t is the real wage, $\tau_{\ell t}$ is the tax rate on wages, and $w_t(1-\tau_{\ell t})$ the post-tax wage for each unit of time spent working. We use k_t to denote physical capital, and R_{kt} for the gross post-tax return on capital, υ_t is the rental rate on capital, δ is the depreciation rate on capital, and τ_{kt} is the tax rate on the net rental rate of capital. The gross post-tax return on capital has the form as before:

$$\begin{aligned} R_{kt+1} &= \upsilon_{t+1} + (1-\delta) - \tau_{kt+1}(\upsilon_{t+1} - \delta) \\ &= 1 + (1-\tau_{kt+1})(\upsilon_{t+1} - \delta). \end{aligned} \tag{13.1}$$

This says that, if the households buy k_{t+1} units of capital in period t, at price p_t, and they rent it in period $t+1$, they will receive rental income, $k_{t+1}\upsilon_{t+1}$, and depreciated capital, $(1-\delta)k_{t+1}$, and pay the taxes on net (of depreciation) return, $\tau_{kt+1}(\upsilon_{t+1} - \delta)k_{t+1}$. We assume, as a stylisation of the US tax code, that capital income taxes are levied on net returns; i.e. depreciation is deducted from the capital income to calculate the tax base.

The household's problem is then

$$\begin{aligned} \max_{\{c_t,n_t,k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t U\left(c_t, n_t\right) \\ s.t. \quad & \sum_{t=0}^{\infty} p_t \left(c_t + k_{t+1}\right) = \sum_{t=0}^{\infty} p_t \left[\left(1-\tau_{\ell t}\right) w_t n_t + k_t R_{kt}\right], \\ & R_{kt} = 1 + \left(1-\tau_{kt}\right)\left(\upsilon_t - \delta\right),\ \forall t \geq 0, \\ & k_0 \text{ given.} \end{aligned} \tag{13.2}$$

Firm's problem The firm's problem in any period t is given by

$$\max_{k_t,n_t} \quad F\left(k_t, n_t, A_t\right) - k_t \upsilon_t - w_t n_t,\ \forall t \geq 0, \tag{13.3}$$

where n_t are the labour services demanded, and k_t is the capital services demanded. We assume that $F\left(\cdot,\cdot,A_t\right)$ has constant returns to scale and it is neoclassical (strictly quasiconcave and satisfies Inada conditions) and that A_t is a productivity shifter.

Government's budget constraint The government budget constraint is given by

$$\sum_{t=0}^{\infty} p_t g_t = \sum_{t=0}^{\infty} p_t \left[\tau_{\ell t} w_t n_t + \tau_{kt} k_t \left(\upsilon_t - \delta\right)\right], \tag{13.4}$$

where g_t is the government purchases at time t. We assume that g_t is exogenous and that the government chooses $\tau_{\ell t}$ and τ_{kt} to satisfy above.

Feasibility Feasibility constraint in this case is

$$g_t + c_t + k_{t+1} = F\left(k_t, n_t, A_t\right) + \left(1-\delta\right) k_t,\ \forall t \geq 0, \tag{13.5}$$

13.1.2 Equilibrium

A competitive equilibrium with taxes where the government finances the (exogenously given) purchases $\{g_t\}_{t=0}^{\infty}$, and the initial condition k_0, is a set of

- ▷ allocations, $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$,
- ▷ prices, $\{p_t, w_t, R_{kt}, \upsilon_t\}_{t=0}^{\infty}$,
- ▷ taxes, $\{\tau_{\ell t}, \tau_{kt}\}_{t=0}^{\infty}$,

such that

- ▷ allocations, $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$, are feasible given $\{g_t\}_{t=0}^{\infty}$ and k_0;
- ▷ allocations, $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$, maximises the household utility given prices, $\{p_t, w_t, R_{kt}, \upsilon_t\}_{t=0}^{\infty}$, taxes, $\{\tau_{\ell t}, \tau_{kt}\}_{t=0}^{\infty}$, and k_0;
- ▷ firms maximise profits given prices $\{w_t, \upsilon_t\}$;
- ▷ government budget constraint holds.

13.1.3 First-order conditions

Households Letting λ_h denote the Lagrange multiplier on the budget constraint, the first-order conditions for (13.2) are

$$\beta^t U_{c_t} = \lambda_h p_t, \qquad \{c_t\} \tag{13.6}$$
$$-\beta^t U_{n_t} = \lambda_h p_t w_t (1-\tau_{\ell t}), \qquad \{n_t\} \tag{13.7}$$
$$0 = (-p_t + p_{t+1} R_{kt+1}) \lambda_h, \qquad \{k_{t+1}\} \tag{13.8}$$

for all $t \geq 0$. Combining these in the usual manner, we obtain

$$\beta \frac{U_{c_{t+1}}}{U_{c_t}} = \frac{p_{t+1}}{p_t} = \frac{1}{1+r_t}, \qquad \frac{\{c_{t+1}\}}{\{c_t\}}$$
$$-\frac{U_{n_t}}{U_{c_t}} = w_t (1-\tau_{\ell t}), \qquad \frac{\{n_t\}}{\{c_t\}} \tag{13.9}$$
$$1 + r_t = R_{kt+1}, \qquad \{k_{t+1}\} \tag{13.10}$$

where r_t is the real rate of interest.

In equilibrium, agents take the sequences of prices and taxes as given so that, at time $t = 0$, they can plan with the knowledge of all future actions. This means that we assume that the government is *committed* to this path of taxes (i.e. it won't renege on its path in the future). We should think of this as a version of rational expectation, in the sense that agents correctly anticipate future prices and taxes.

Firms The firm's problem, (13.3), gives the following (usual) first-order conditions:

$$w_t = F_n (k_t, n_t, A_t), \qquad \{n_t\} \tag{13.11}$$
$$v_t = F_k (k_t, n_t, A_t), \qquad \{k_t\} \tag{13.12}$$

for all $t \geq 0$.

Notice that prices for labour and capital differ for firms and households due to the taxes:

$$-\frac{U_{n_t}}{U_{c_t}} = (1-\tau_{\ell t}) w_t \neq w_t = F_n (k_t, n_t, A_t),$$
$$1 + r_t + \delta \neq 1 + \delta + \frac{r_t}{1-\tau_{kt+1}} = v_t = F_k (k_{t+1}, n_{t+1}, A_{t+1}).$$

This is due to the wedges created by the taxes. The tax wedges mean that household's marginal rate of substitution is no longer equated the the firm's marginal rate of transformation (i.e. the marginal product of capital net of depreciation plus one).

Remark 13.1. Marginal rate of transformation represents how many units of today's consumption must be given up (i.e. invested) in order to obtain one unit of consumption tomorrow. Hence,

$$MRT_{t,t+1} = \frac{1}{1 + F_k (k_{t+1}, n_{t+1}) - \delta}.$$

13.1.4 Implementability

Our goals is to solve for equilibrium taxes $\{\tau_{\ell t}, \tau_{kt}\}$ that finances the exogenous stream of government purchases $\{g_t\}$. However, optimising $\{\tau_{\ell t}, \tau_{kt}\}$ is analytically difficult, we therefore follow an alternative "primal" approach, where we let government choose allocations, instead of taxes. This involves eliminating taxes and prices from the household's budget constraint using the first-order conditions we derived above.

Recall the household's budget constraint:

$$\sum_{t=0}^{\infty} p_t \left(c_t + k_{t+1}\right) = \sum_{t=0}^{\infty} p_t \left[\left(1 - \tau_{\ell t}\right) w_t n_t + k_t R_{kt}\right].$$

First, let us collect the terms on k_{t+1}:

$$\begin{aligned}\sum_{t=0}^{\infty} p_t \left(c_t - \left(1 - \tau_{\ell t}\right) w_t n_t\right) &= \sum_{t=0}^{\infty} p_t k_t R_{kt} - \sum_{t=0}^{\infty} p_t k_{t+1} \\ &= p_0 k_0 R_{k0} + \sum_{t=0}^{\infty} k_{t+1} \left(p_{t+1} R_{kt+1} - p_t\right).\end{aligned}$$

Using (13.8), we can simplify above to:

$$\sum_{t=0}^{\infty} p_t \left(c_t - \left(1 - \tau_{\ell t}\right) w_t n_t\right) = p_0 k_0 R_{k0}.$$

We can then use the first-order condition for the household with respect to c_t, (13.6), to eliminate p_t's:

$$\begin{aligned}\sum_{t=0}^{\infty} \frac{\beta^t U_{c_t}}{\lambda_h} \left(c_t - \left(1 - \tau_{\ell t}\right) w_t n_t\right) &= \frac{U_{c_0}}{\lambda_h} k_0 R_{k0} \\ \Rightarrow \sum_{t=0}^{\infty} \beta^t U_{c_t} \left(c_t - \left(1 - \tau_{\ell t}\right) w_t n_t\right) &= U_{c_0} k_0 R_{k0}.\end{aligned}$$

We then use the ratio of the first-order condition for the household with respect to n_t and c_t, (13.9), to eliminate $\left(1 - \tau_{\ell t}\right) w_t$,

$$\begin{aligned}\sum_{t=0}^{\infty} \beta^t U_{c_t} \left(c_t + \frac{U_{n_t}}{U_{c_t}} n_t\right) &= U_{c_0} k_0 R_{k0} \\ \Rightarrow \sum_{t=0}^{\infty} \beta^t \left(U_{c_t} c_t + U_{n_t} n_t\right) &= U_{c_0} k_0 R_{k0}.\end{aligned}$$

Substituting the expression for R_{k0} yields

$$\begin{aligned}\sum_{t=0}^{\infty} \beta^t \left(U_{c_t} c_t + U_{n_t} n_t\right) &= U_{c_0} k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(v_0 - \delta\right)\right) \\ &= U_{c_0} k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(F_k\left(k_0, n_0\right) - \delta\right)\right),\end{aligned}$$

where we used (13.12). We therefore have one equation, called the implementability constraint,

$$\sum_{t=0}^{\infty} \beta^t \left(U_{c_t} c_t + U_{n_t} n_t\right) = U_{c_0} k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(F_{k0} - \delta\right)\right). \tag{13.13}$$

This is, in essence, the budget constraint of the agent after replacing taxes and prices using the relevant marginal rates of substitutions (i.e. using the first-order conditions that must be satisfied in the equilibrium).

The allocation that constitute an equilibrium must therefore satisfy the implementability constraint, (13.13), and the feasibility constraints, (13.5), in each period.

We can alternatively interpret implementability constraint as the budget constraint of the government where we replace prices using the relevant marginal rates of substitution. (Recall that, by Walras' Law, if two of the three constraints (household's budget constraint, government's budget constraint and the feasibility constraints) are satisfied, then the last constraint is automatically satisfied.)

To see that implementability constraint and the feasibility constraints imply that the budget constraint for the government holds, replace c_t in (13.13) using the feasibility constraint:

$$\sum_{t=0}^{\infty} \beta^t \left(U_{c_t}\left(F\left(k_t, n_t, A_t\right) - k_{t+1} + k_t\left(1-\delta\right) - g_t\right) + U_{n_t} n_t\right) = U_{c_0} k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(F_{k0} - \delta\right)\right).$$

Recovering prices,

$$\begin{aligned}
&\sum_{t=0}^{\infty} \frac{\beta^t U_{c_t}}{\lambda_h}\left(\left(F\left(k_t, n_t, A_t\right) - k_{t+1} + k_t\left(1-\delta\right) - g_t\right) + \frac{U_{n_t}}{U_{c_t}\lambda_h} n_t\right) = \frac{U_{c_0}}{\lambda_h} k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(F_{k0} - \delta\right)\right) \\
\Rightarrow &\sum_{t=0}^{\infty} p_t\left(F\left(k_t, n_t, A_t\right) - k_{t+1} + k_t\left(1-\delta\right) - g_t - \left(1-\tau_{\ell t}\right) w_t n_t\right) = p_0 k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(F_{k0} - \delta\right)\right).
\end{aligned}$$

By the Euler's Theorem, and the first-order conditions of the firm,

$$F\left(k_t, n_t, A_t\right) = F_k k_t + F_n n_t = \upsilon_t k_t + w_t n_t.$$

Hence,

$$\begin{aligned}
p_0 k_0 \left(1 + \left(1 - \tau_{k0}\right)\left(F_{k0} - \delta\right)\right) &= \sum_{t=0}^{\infty} p_t\left(\upsilon_t k_t + w_t n_t - k_{t+1} + k_t\left(1-\delta\right) - g_t - \left(1-\tau_{\ell t}\right) w_t n_t\right) \\
&= \sum_{t=0}^{\infty} p_t\left(-k_{t+1} + k_t\left(\upsilon_t + 1 - \delta\right) - g_t + \tau_{\ell t} w_t n_t\right).
\end{aligned}$$

Recall (13.1) which gives that

$$\begin{aligned}
R_{kt+1} &= 1 + \left(1 - \tau_{kt+1}\right)\left(\upsilon_{t+1} - \delta\right) \\
\Rightarrow \upsilon_{t+1} + 1 - \delta &= R_{kt+1} + \tau_{kt+1}\left(\upsilon_{t+1} - \delta\right),
\end{aligned}$$

we have

$$
\begin{aligned}
p_0 k_0 R_0 &= \sum_{t=0}^{\infty} p_t \left(-k_{t+1} + k_t \left(R_{kt} + \tau_{kt} \left(\upsilon_t - \delta\right)\right) - g_t + \tau_{\ell t} w_t n_t\right) \\
\sum_{t=0}^{\infty} p_t \left(-g_t + \tau_{\ell t} w_t n_t + \tau_{kt} k_t \left(\upsilon_t - \delta\right)\right) &= p_0 k_0 R_0 + \sum_{t=0}^{\infty} p_t \left(k_{t+1} - k_t R_{kt}\right) \\
&= p_0 k_0 R_0 + \sum_{t=0}^{\infty} p_t k_{t+1} - \left(p_0 k_0 R_0 + \sum_{t=0}^{\infty} p_{t+1} k_{t+1} R_{kt+1}\right) \\
&= \sum_{t=0}^{\infty} k_{t+1} \left(p_t - p_{t+1} k_{t+1} R_{kt+1}\right) \\
&= 0,
\end{aligned}
$$

where we used the (13.8). That is,

$$
\sum_{t=0}^{\infty} p_t g_t = \sum_{t=0}^{\infty} p_t \left[\tau_{\ell t} w_t n_t + \tau_{kt} k_t \left(\upsilon_t - \delta\right)\right].
$$

13.1.5 Recovering taxes

With the exception of τ_{k0}, there are no explicit taxes in the implementability constraint. However, given an allocation that satisfies the implementability constraint, as well as feasibility constraints, we can recover *tax wedges* by using the relevant marginal rate of substitution and marginal rate of transformation.

For labour, from (13.9)

$$
\begin{aligned}
1 - \tau_{\ell t} &= -\frac{U_{n_t}/U_{c_t}}{w_t} \\
&= -\frac{U_{n_t}/U_{c_t}}{F_n\left(k_t, n_t, A_t\right)},
\end{aligned}
\tag{13.14}
$$

where we also used (13.11) to express the wedge purely in terms of allocations.

For capital, from (13.10),

$$
\begin{aligned}
1 + r_t = R_{kt+1} &= 1 + \left(1 - \tau_{kt+1}\right)\left(\upsilon_{t+1} - \delta\right) \\
\Rightarrow 1 - \tau_{kt+1} &= \frac{r_t}{\upsilon_{t+1} - \delta} \\
&= \frac{\frac{1}{\beta\left(U_{c_{t+1}}/U_{c_t}\right)} - 1}{F_k\left(k_{t+1}, n_{t+1}, A_{t+1}\right) - \delta},
\end{aligned}
\tag{13.15}
$$

where we used (13.6) and (13.12) to express the wedge purely in terms of allocations.

13.2 Ramsey Problem

The Ramsey Problem is to find the policy $\{\tau_{\ell t}, \tau_{kt}\}_{t=0}^{\infty}$ that support a competitive equilibrium with taxes given an exogenous path of government purchases, $\{g_t\}_{t=0}^{\infty}$ assuming linear taxation. The problem can be characterised as follows:

$$\begin{aligned}
\max_{\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t U(c_t, n_t) \\
s.t. \quad & g_t + c_t + k_{t+1} = F(k_t, n_t, A_t) + (1-\delta) k_t, \ \forall t \geq 0 \\
& \sum_{t=0}^{\infty} \beta^t \left(U_{c_t}(c_t, n_t) c_t + U_{n_t}(c_t, n_t) n_t \right) = U_{c_0} k_0 \left(1 + (1-\tau_{k0})(F_{k0} - \delta)\right), \\
& k_0 \text{ given}, \\
& \tau_{k0} \leq \tau_{k0}^m.
\end{aligned}$$

The implementability constraint suggests that τ_{k0} should be set at the highest possible value because taxing initial capital, k_0, does not change the agent's decision since capital is, in the very first period, fixed in supply. Thus, τ_{k0} acts as a lump-sum tax. Taxing capital in later periods is different because they will distort capital accumulation by the households. Therefore, to make the problem interesting (e.g. avoid the case in which initial period tax funds the entirety of government purchases), we assume that there is an upper bound on τ_{k0}, say τ_{k0}^m.

13.2.1 Solving the Ramsey Problem

Let us write the Ramsey Problem as follows, letting λ be the Lagrange multiplier of the implementability constraint:

$$\begin{aligned}
\max_{\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \left[U(c_t, n_t) + \lambda \left(U_{c_t}(c_t, n_t) c_t + U_{n_t}(c_t, n_t) n_t \right) \right] - \lambda U_{c_0} k_0 \left(1 + (1-\tau_{k0})(F_{k0} - \delta)\right) \\
s.t. \quad & g_t + c_t + k_{t+1} = F(k_t, n_t, A_t) + (1-\delta) k_t, \ \forall t \geq 0 \\
& k_0 \text{ given}, \\
& \tau_{k0} \leq \bar{\tau}_{k0}.
\end{aligned}$$

The solution to this maximisation problem solves the Ramsey problem if the implementability constraint is satisfied. Equivalently, the maximisation problem above solves the Ramsey problem if λ is indeed the Lagrange multiplier of the implementability constraint.

Define

$$W(c_t, n_t; \lambda) := U(c_t, n_t) + \lambda \left(U_{c_t}(c_t, n_t) c_t + U_{n_t}(c_t, n_t) n_t \right).$$

We can interpret W has the social value of consumption and labour since $U(c_t, n_t)$ represents private value, and $\lambda(U_{c_t} c_t + U_{n_t} n_t)$ are the taxes corresponding to period-t allocation in terms of period-t utility, which represents the penalty for violating implementability.

Using W, we can rewrite the problem as

$$
\begin{aligned}
\max_{\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t W(c_t, n_t; \lambda) - \lambda U_{c_0} k_0 \left(1 + (1-\tau_{k0})(F_{k0} - \delta)\right) \\
s.t. \quad & g_t + c_t + k_{t+1} = F(k_t, n_t, A_t) + (1-\delta) k_t, \ \forall t \geq 0 \\
& k_0 \text{ given}, \\
& \tau_{k0} \leq \tau_{k0}^m.
\end{aligned}
$$

In this form, the problem is similar to the planner's problem for the neoclassical growth model except that: (i) the period-utility function $U(c, n)$ has been replaced by the social value function, $W(c, n, \lambda)$; (ii) the first term in the discounted utility is treated differently from others.

The first-order conditions with respect to n_t and k_{t+1} are (substitute for c_t using the feasibility constraint):

$$
\begin{aligned}
0 &= \beta^t \left(W_{c_t} F_{n_t}(k_t, n_t, A_t) + W_{n_t}\right) \\
\Leftrightarrow -\frac{W_{n_t}}{W_{c_t}} &= F_{n_t}(k_t, n_t, A_t), \qquad (13.16) \\
\beta^t W_{c_t} &= \beta^{t+1} W_{c_{t+1}} \left(F_{kt+1} + (1-\delta)\right) \\
\Leftrightarrow 1 &= \beta \frac{W_{c_{t+1}}}{W_{c_t}} \left(F_{kt+1} + (1-\delta)\right), \qquad (13.17)
\end{aligned}
$$

for all $t \geq 1$.The first-order condition for time $t = 0$ differ from those of others due to the period-0 term in the objective function.

13.2.2 Chamley-Judd Result: No taxation of capital in the steady state

Proposition 13.1. *Assume that $g_t \to \bar{g}$ and $A_t \to \bar{A}$ as $t \to \infty$. In a steady state of the Ramsey Problem corresponding to $\bar{g}$ and $\bar{A}$, there is no taxation of capital; i.e. $\bar{\tau}_k = 0$.*[a]

[a]Note that the statement only concerns the steady state, not the path to the state state!

Proof. To see this, notice that, in a steady state, $c_{t+1} = c_t = c$ and $n_{t+1} = n_t = n$ so that $W_{c_{t+1}} = W_{c_t}$. Thus, from (13.17), it must be that

$$1 = \beta \left(F_k + (1-\delta)\right). \qquad (13.18)$$

Substituting above into (13.15) gives that $\bar{\tau}_k = 0$. However, to see this more easily, we can write (13.10) using (13.6), (13.1) and (13.12)

$$
\begin{aligned}
1 &= \frac{1}{1+r_t} R_{kt+1} \\
&= \beta \frac{U_{c_{t+1}}}{U_{c_t}} \left(1 + (1-\tau_{kt+1})(\upsilon_{t+1} - \delta)\right) \\
&= \beta \frac{U_{c_{t+1}}}{U_{c_t}} \left(1 + (1-\tau_{kt+1})(F_{kt+1} - \delta)\right), \qquad (13.19)
\end{aligned}
$$

which, in a steady state, must be

$$1 = \beta\left(1 + (1 - \bar{\tau}_k)(F_k - \delta)\right). \tag{13.20}$$

For (13.18) and (13.20) to be both true, it must be that $\bar{\tau}_k = 0$. ■

To understand the intuition behind this result, let us analyse how a steady state looks like for a competitive equilibrium with constant taxes, $\bar{\tau}_k > 0$ and $\bar{\tau}_l > 0$. We will compute the ratio of the marginal rate of substitution to the marginal rate of transformation, which gives the tax wedge.

The marginal rate of substitution between consumption in period t and period $t+T$ is given by

$$MRS_{t,t+T} \coloneqq \frac{\beta^{t+T} U_{c_{t+T}}}{\beta^t U_{c_t}}.$$

The marginal rate of transformation between consumption at t and $t+T$ s given by

$$MRT_{t,t+T} \coloneqq \frac{1}{1-\delta+F_{kt+1}} \frac{1}{1-\delta+F_{kt+2}} \cdots \frac{1}{1-\delta+F_{kt+T}}.$$

The ratio is given by

$$\begin{aligned}
\frac{MRS_{t,t+T}}{MRT_{t,t+T}} &= \frac{\beta^T U_{c_{t+T}}}{U_{c_t}} \prod_{i=1}^{T} (1-\delta+F_{kt+i}) \\
&= \frac{\beta^T U_{c_{t+T}}}{U_{c_t}} \prod_{i=1}^{T} \left(1 + (1-\tau_{kt+i})(F_{kt+i}-\delta)\right) \frac{1-\delta+F_{kt+i}}{1+(1-\tau_{kt+i})(F_{kt+i}-\delta)}.
\end{aligned}$$

Recall that, in equilibrium, $F_{kt} = \upsilon_t$ (see (13.12)) so that

$$1 + (1-\tau_{kt+i})(F_{kt+i}-\delta) = 1 + (1-\tau_{kt+i})(\upsilon_{t+i}-\delta) = R_{kt+i}$$

and (see (13.9) and (13.10)):

$$\frac{\beta^T U_{c_{t+T}}}{U_{c_t}} = \prod_{i=1}^{T} \frac{1}{R_{kt+i}}.$$

Hence, in equilibrium,

$$\frac{MRS_{t,t+T}}{MRT_{t,t+T}} = \prod_{i=1}^{T} \frac{1-\delta+F_{kt+i}}{1+(1-\tau_{kt+i})(F_{kt+i}-\delta)}.$$

At the steady state,

$$\begin{aligned}
F_{kt} &= \bar{\upsilon}, \\
\rho &= (1-\tau_{kt+i})(F_{kt+i}-\delta)
\end{aligned}$$

so that

$$\begin{aligned}
\frac{MRS_{t,t+T}}{MRT_{t,t+T}} &= \left(\frac{1+(\bar{v}-\delta)}{1+(1-\bar{\tau}_k)(\bar{v}-\delta)}\right)^T = \left(\frac{1+\frac{\rho}{1-\bar{\tau}_k}}{1+\rho}\right)^T \\
&= \left(\frac{1-\bar{\tau}_k+\rho}{(1+\rho)(1-\bar{\tau}_k)}\right)^T = \left(\frac{(1+\rho)(1-\bar{\tau}_k)+\rho\bar{\tau}_k}{(1+\rho)(1-\bar{\tau}_k)}\right)^T \\
&= \left(1+\frac{\rho\bar{\tau}_k}{(1+\rho)(1-\bar{\tau}_k)}\right)^T.
\end{aligned}$$

This means that, if the tax rates on capital are positive in the steady state (i.e. $\bar{\tau}_k > 0$), then the wedge between the marginal rate of substitution and marginal rate of transformation increases exponentially with T.

This suggests that tax rates on capital income should not only be zero in the long run, i.e. in a steady state, but even fairly soon. To see this, we consider the following.

Proposition 13.2. *Suppose utility is given by*

$$U(c,n) = \begin{cases} \frac{c^{1-\sigma}n^{\varphi}}{1-\sigma} & \sigma \neq 1, \\ \frac{c^{1-\sigma}}{1-\sigma} + h(n) & \sigma > 0, \end{cases}$$

where $\sigma > 0$. Then, the optimal tax rate for capital is zero from $t \geq 2$ (even if the capital is far away from its steady-state level, and even if A_t and g_t are varying over time). Moreover, in the steady state, labour tax should be constant, $\tau_{\ell t} = \bar{\tau}_l > 0$.

Proof. First, note that

$$W_{c_t} = U_{c_t} + \lambda\left(U_{c_t c_t}c_t + U_{c_t} + U_{n_t c_t}n_t\right),$$

where, using the using the expression for $U(c,n)$ with $s=1$,

$$\begin{aligned}
U_{c_t} &= c^{-\sigma}n^{\varphi}, \\
U_{n_t} &= \varphi\frac{c^{1-\sigma}n^{\varphi-1}}{1-\sigma}, \\
U_{c_t c_t} &= -\sigma c^{-\sigma-1}n^{\varphi}, \\
U_{n_t c_t} &= \varphi c^{-\sigma}n^{\varphi-1},
\end{aligned}$$

which imply that

$$\begin{aligned}
U_{c_t c_t}c_t &= -\sigma c^{-\sigma}n^{\varphi} = -\sigma U_{c_t}, \\
U_{n_t,c_t}n_t &= \varphi c^{-\sigma}n^{\varphi} = \varphi U_{c_t}.
\end{aligned}$$

Thus,

$$\begin{aligned}
W_{c_t} &= U_{c_t} + \lambda\left(-\sigma U_{c_t} + U_{c_t} + \varphi U_{c_t}\right) \\
&= U_{c_t}\left(1+\lambda-\sigma\lambda+\varphi\lambda\right).
\end{aligned}$$

In this case, (13.17) is given by

$$\begin{aligned} 1 &= \beta \frac{W_{c_{t+1}}}{W_{c_t}} \left(F_{kt+1} + (1-\delta)\right) \\ &= \beta \frac{U_{c_{t+1}}}{U_{c_t}} \left(F_{kt+1} + (1-\delta)\right), \ \forall t \geq 1. \end{aligned}$$

From the consumer's problem, recall that (13.19) must hold; i.e.

$$1 = \beta \frac{U_{c_{t+1}}}{U_{c_t}} \left(1 + \left(1 - \tau_{kt+1}\right)\left(F_{kt+1} - \delta\right)\right), \ \forall t \geq 0.$$

Hence, it must be that $\tau_{kt} = 0$ for all $t \geq 2$.

If the preferences are given by the second line, then

$$\begin{aligned} U_{c_t} &= c^{-\sigma}, \\ U_{n_t} &= h'(n), \\ U_{c_t c_t} &= -\sigma c^{-\sigma-1}, \\ U_{n_t c_t} &= 0, \end{aligned}$$

so that

$$\begin{aligned} U_{c_t c_t} c_t &= -\sigma U_{c_t}, \\ U_{n_t, c_t} n_t &= 0. \end{aligned}$$

Then,

$$\begin{aligned} W_{c_t} &= U_{c_t} + \lambda\left(-\sigma U_{c_t} + U_{c_t}\right) \\ &= U_{c_t}\left(1 + \lambda(1-\sigma)\right), \\ \Rightarrow \frac{W_{c_{t+1}}}{W_{c_t}} &= \frac{U_{c_{t+1}}\left(1 + \lambda(1-\sigma)\right)}{U_{c_t}\left(1 + \lambda(1-\sigma)\right)} = \frac{U_{c_{t+1}}}{U_{c_t}}. \end{aligned}$$

Hence, we again obtain the result that $\tau_{kt} = 0$ for all $t \geq 2$.

Recall that the labour tax wedge is given by, (13.14),

$$1 - \tau_{\ell t} = -\frac{U_{n_t}/U_{c_t}}{F_{n_t}}.$$

Hence, it is immediate that, in the steady state, labour tax should be constant, $\tau_{\ell t} = \bar{\tau}_l > 0$. It must also be positive since $\tau_{kt} = 0$ and we assumed that government requires revenue in addition to τ_{k0}. ■

13.2.3 Equivalent tax schemes

There are other ways to obtain essentially the same allocation that correspond to zero capital tax rates and positive labour rates. We explore two examples here. Specifically, we show that a constant labour income tax and a zero tax on capital income tax is equivalent to:

- ▷ a constant gross income tax rate, together with investment tax credit;

▷ tax on consumption.

Gross tax with investment tax credit Consider the budget constraint of a household subject to a net income tax and investment tax credit:

$$\sum_{t=0}^{\infty} p_t (x_t + c_t) = \sum_{t=0}^{\infty} p_t \left((w_t n_t + \upsilon_t k_t) - \tau_t (n_t w_t + \upsilon_t k_t - x_t) \right),$$

where tax is levied on the entire labour and capital income, $w_t n_t + \upsilon_t k_t$, but the household also receives a tax credit for the investment. We can rewrite above as

$$\sum_{t=0}^{\infty} p_t \left((1-\tau_t) x_t + c_t \right) = \sum_{t=0}^{\infty} p_t \left((1-\tau_t)(w_t n_t + \upsilon_t k_t) \right). \tag{13.21}$$

The law of motion for capital is given by

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

We can rewrite the budget constraint using the law of motion for capital as

$$\begin{aligned}
\sum_{t=0}^{\infty} p_t c_t &= \sum_{t=0}^{\infty} p_t (1-\tau_t)(w_t n_t + \upsilon_t k_t - x_t) \\
&= \sum_{t=0}^{\infty} p_t (1-\tau_t)(w_t n_t + \upsilon_t k_t - (k_{t+1} - (1-\delta) k_t)) \\
&= \sum_{t=0}^{\infty} p_t (1-\tau_t)(w_t n_t - k_{t+1} + (\upsilon_t + 1 - \delta) k_t) \\
\Rightarrow \sum_{t=0}^{\infty} p_t (c_t - (1-\tau_t) w_t n_t) &= \sum_{t=0}^{\infty} p_t (1-\tau_t)(\upsilon_t + 1 - \delta) k_t - \sum_{t=0}^{\infty} p_t (1-\tau_t) k_{t+1} \\
&= p_0 (1-\tau_0)(\upsilon_0 + 1 - \delta) k_0 + \sum_{t=0}^{\infty} p_{t+1} (1-\tau_{t+1})(\upsilon_{t+1} + 1 - \delta) k_{t+1} \\
&\quad - \sum_{t=0}^{\infty} p_t (1-\tau_t) k_{t+1} \\
&= \sum_{t=0}^{\infty} p_{t+1} k_{t+1} \left[(1-\tau_{t+1})(\upsilon_{t+1} + 1 - \delta) - \frac{p_t}{p_{t+1}} (1-\tau_t) \right] \\
&\quad + p_0 (1-\tau_0)(\upsilon_0 + 1 - \delta) k_0.
\end{aligned}$$

Setting the term with k_{t+1} to zero, we obtain

$$\begin{aligned}
(1-\tau_{t+1})(\upsilon_{t+1} + 1 - \delta) &= \frac{p_t}{p_{t+1}} (1-\tau_t) \\
\Rightarrow \frac{1-\tau_{t+1}}{1-\tau_t} (\upsilon_{t+1} + 1 - \delta) &= 1 + r_t.
\end{aligned}$$

Thus, if income tax rates are constant, i.e. $\tau_t = \tau_{t+1}$, then

$$r_t + \delta = \upsilon_{t+1},$$

which is the same expression that is obtained if net capital income is taxed at the rate $\tau_{kt+1} = 0$. In this sense, a constant income tax rate, together with investment tax credit, is equivalent to a constant labour income tax and a zero tax on capital income tax.

Example 13.1. Consider buying one investment good at time t, which costs (net of tax credit), $1 - \tau_t$. The purchase is financed by borrowing so that the cash flow in period t is zero. At the end of period $t+1$, you must return $(1 + r_t)(1 - \tau_t)$. The capital purchased at time t is rented in period $t+1$ at the rental rate υ_{t+1}. After-tax revenue from renting the capital is $\upsilon_{t+1}(1 - \tau_{t+1})$. After renting, the investment is sold, and the return is given by the value net of depreciation, net of taxes, i.e. $(1-\delta)(1-\tau_{t+1})$. The time-t profits of these operations are

$$(1-\delta)(1-\tau_{t+1}) + \upsilon_{t+1}(1-\tau_{t+1}) - (1+r_t)(1-\tau_t).$$

The no-arbitrage condition requires above to equal zero so that

$$\frac{1-\tau_{t+1}}{1-\tau_t}(\upsilon_{t+1} + 1 - \delta) = 1 + r_t,$$

which is the same expression as we obtained above.

Example 13.2. Consider the effect of an investment tax credit on the present value of stream of income created by investing one unit of consumption good at time t net of the cost of acquiring this unit. The stream of income is given by (note that capital depreciates at the “end” of the period)

$$\begin{aligned}
&p_{t+1}k_t\left(\underbrace{\upsilon_{t+1}}_{\text{rental income}} - \underbrace{\tau_{kt+1}(\upsilon_{t+1}-\delta)}_{\text{tax on rent}}\right)\\
&+ p_{t+2}k_t \underbrace{(1-\delta)}_{\text{capital depreciates}} (\upsilon_{t+2} - \tau_{kt+2}(\upsilon_{t+2}-\delta))\\
&+ p_{t+3}k_t(1-\delta)^2(\upsilon_{t+3} - \tau_{kt+3}(\upsilon_{t+3}-\delta))\\
&\quad\vdots
\end{aligned}$$

Expressing above as an infinite sum, and combining with the cost of acquiring capital yields

$$-p_t k_t + \sum_{j=1}^{\infty} p_{t+j}k_t(1-\delta)^{j-1}(\upsilon_{t+j} - \tau_{kt+j}(\upsilon_{t+j}-\delta)).$$

If $\tau_{kt+j} = 0$ for all j, then we obtain

$$-p_t k_t + \sum_{j=1}^{\infty} p_{t+j}k_t(1-\delta)^{j-1}\upsilon_{t+j}. \tag{13.22}$$

Now, consider the case of uniform taxation of income with an investment tax credit:

$$-p_t(1-\tau_t)k_t + \sum_{j=1}^{\infty} p_{t+j}k_t(1-\delta)^{j-1}\upsilon_{t+j}(1-\tau_{t+j}).$$

In this case, we see that the investment tax credit decreases the price of purchasing capital, hence the $1-\tau_t$ on the left-hand side. Suppose that $\tau_{t+j} = \tau$ for all j, then

$$(1-\tau)\left(-p_t k_t + \sum_{j=1}^{\infty} p_{t+j} k_t (1-\delta)^{j-1} \upsilon_{t+j}\right).$$

Since, in equilibrium, above expression must sum to zero, we see that investment tax credit has the same effect as the zero tax rate on capital income.

Remark 13.2. To see that (13.22) is the same as the one we've been dealing with: set $k_t = 1$, then, notice that

$$\begin{aligned} p_t &= \sum_{j=1}^{\infty} p_{t+j} (1-\delta)^{j-1} \upsilon_{t+j} \\ &= p_{t+1}\upsilon_{t+1} + (1-\delta)\sum_{j=1}^{\infty} p_{t+1+j}(1-\delta)^{j-1}\upsilon_{t+1+j} \\ &= p_{t+1}\upsilon_{t+1} + (1-\delta) p_{t+1} \\ \Rightarrow 1 + r_t &= \upsilon_{t+1} + 1 - \delta. \end{aligned}$$

Exercise 13.1. Compare two cases: (i) setting tax rates to capital income $\tau_{k0} = \bar{\tau}$, $\tau_{kt} = 0$ from period $t \geq 1$ on, and setting labour income taxes to a constant value, $\tau_{\ell t} = \bar{\tau}$; and (ii) setting a tax on gross income $\tau_t = \bar{\tau}$ at a constant level and including an investment tax credit at the same rate. Write the budget constraint for the government in each of the two cases.

Solution. In case of gross taxation with tax credit for investment:

$$\begin{aligned} \sum_{t=0}^{\infty} p_t g_t &= \sum_{t=0}^{\infty} p_t \tau_t (w_t n_t + \upsilon_t k_t - x_t) \\ &= \sum_{t=0}^{\infty} p_t \tau_t (w_t n_t + \upsilon_t k_t - (k_{t+1} - (1-\delta) k_t)) \\ &= \sum_{t=0}^{\infty} p_t \tau_t (w_t n_t - k_{t+1} + (\upsilon_t + 1 - \delta) k_t) \\ &= \sum_{t=0}^{\infty} p_t \tau_t w_t n_t + \sum_{t=0}^{\infty} p_t \tau_t ((\upsilon_t + 1 - \delta) k_t) - \sum_{t=0}^{\infty} p_t \tau_t k_{t+1} \\ &= \sum_{t=0}^{\infty} p_t \tau_t w_t n_t + \sum_{t=0}^{\infty} p_{t+1} \tau_{t+1} ((\upsilon_{t+1} + 1 - \delta) k_{t+1}) - \sum_{t=0}^{\infty} p_t \tau_t k_{t+1} \\ &\quad + p_0 \tau_0 (\upsilon_0 + 1 - \delta) k_0 \\ &= \sum_{t=0}^{\infty} p_t \tau_t w_t n_t + \sum_{t=0}^{\infty} p_{t+1} k_{t+1} \left(\tau_{t+1} (\upsilon_{t+1} + 1 - \delta) - \frac{p_t}{p_{t+1}} \tau_t\right) \\ &\quad + p_0 \tau_0 (\upsilon_0 + 1 - \delta) k_0 \end{aligned}$$

Setting $\tau_t = \bar{\tau}$, we obtain

$$\begin{aligned}
\sum_{t=0}^{\infty} p_t g_t &= \bar{\tau} \sum_{t=0}^{\infty} p_t w_t n_t + \bar{\tau} \sum_{t=0}^{\infty} p_{t+1} k_{t+1} \underbrace{\left(\left(\upsilon_{t+1} + 1 - \delta \right) - \left(1 + r_t \right) \right)}_{=0 \because r_t + \delta = \upsilon_{t+1}} \\
&\quad + p_0 \bar{\tau} \left(\upsilon \upsilon_0 + 1 - \delta \right) k_0 \\
&= \bar{\tau} \sum_{t=0}^{\infty} p_t w_t n_t + \bar{\tau} p_0 \left(\upsilon_0 - \delta \right) k_0 .
\end{aligned}$$

When tax on capital and labour are separate, we have

$$\begin{aligned}
\sum_{t=0}^{\infty} p_t g_t &= \sum_{t=0}^{\infty} p_t \left(\tau_\ell w_t n_t + \tau_{kt} \left(\upsilon_t - \delta \right) k_t \right) \\
&= \sum_{t=0}^{\infty} p_t \tau_\ell w_t n_t + \sum_{t=0}^{\infty} p_t \tau_{kt} \left(\upsilon_t - \delta \right) k_t .
\end{aligned}$$

Substituting the particular values of taxes gives

$$\sum_{t=0}^{\infty} p_t g_t = \bar{\tau} \sum_{t=0}^{\infty} p_t w_t n_t + \bar{\tau} p_0 \left(\upsilon_0 - \delta \right) k_0 . \tag{13.23}$$

Now recall the government 's budget constraint (13.4)

$$\sum_{t=0}^{\infty} p_t g_t = \sum_{t=0}^{\infty} p_t \left[\tau_{\ell t} w_t n_t + \tau_{kt} k_t \left(\upsilon_t - \delta \right) \right] .$$

Setting $\tau_{k0} = \bar{\tau}$, $\tau k_t = 0$ for all $t \geq 1$ and $\tau_{\ell_t} = \bar{\tau}$ for all $t \geq 0$ yields

$$\sum_{t=0}^{\infty} p_t g_t = \bar{\tau} \sum_{t=0}^{\infty} p_t w_t n_t + \bar{\tau} p_0 \left(\upsilon_0 - \delta \right) k_0 ,$$

which is exactly the same as (13.23) we derived above under gross income tax with investment tax credit.f

Consumption tax Suppose we now tax consumption so that the budget constraint of the household subject to consumption tax rate, τ_{ct}, is given by

$$\sum_{t=0}^{\infty} p_t \left[\left(1 + \tau_{ct} \right) c_t + x_t \right] = \sum_{t=0}^{\infty} p_t \left(w_t n_t + k_t \upsilon_t \right) . \tag{13.24}$$

Notice that this can be rewritten in terms of tax on gross income with tax credit on investment. Take any period t,

$$\begin{aligned}
&\left(1 + \tau_{ct} \right) c_t + x_t = w_t n_t + k_t \upsilon_t \\
\Rightarrow\ & c_t + \frac{1}{\left(1 + \tau_{ct} \right)} x_t = \frac{1}{\left(1 + \tau_{ct} \right)} \left(w_t n_t + k_t \upsilon_t \right) .
\end{aligned}$$

Define $1 - \tau_t = 1 / \left(1 + \tau_{ct} \right)$, then

$$c_t + \left(1 - \tau_t \right) x_t = \left(1 - \tau_t \right) \left(w_t n_t + k_t \upsilon_t \right) .$$

Thus, we can rewrite (13.24) exactly as (13.21).

13.2.4 Labour tax

The next proposition establishes conditions under which the tax rate on labour, $\tau_{\ell t}$, does not vary over time, even though capital is not at its steady-state value, and even though both g_t and A_t could vary over time. The main condition is that the utility function has constant elasticities of substitution.

Proposition 13.3. *Let preferences be as follows*

$$U(c,\ell) = \begin{cases} \frac{c^{1-\sigma}(1-\ell)^{-\varphi(1-\sigma)}}{1-\sigma} & \sigma > 1, \varphi \geq \frac{\sigma}{\sigma-1} \\ \frac{c^{1-\sigma}-1}{1-\sigma} - A\frac{(1-\ell)^{1+\varphi}}{1+\varphi} & \sigma > 0, A > 0, \varphi \geq 0 \end{cases}$$

where labour supply is $n = 1-\ell$ and ℓ is leisure. Then the optimal tax rate on labour, $\tau_{\ell t}$, is constant from $\tau \geq 1$.

Proof. Consider the case in which utility is given by the first line, then

$$\begin{aligned} U_{n_t} = U_{1-\ell_t} &= -\varphi c^{1-\sigma}(1-\ell)^{-\varphi(1-\sigma)-1}, \\ U_{c_t} &= c^{-\sigma}(1-\ell)^{-\varphi(1-\sigma)}, \\ U_{n_t n_t} &= \varphi(\varphi(1-\sigma)+1)c^{1-\sigma}(1-\ell)^{-\varphi(1-\sigma)-2}, \\ U_{c_t c_t} &= -\sigma c^{-\sigma-1}(1-\ell)^{-\varphi(1-\sigma)}, \\ U_{c_t n_t} &= -\varphi(1-\sigma)c^{-\sigma}(1-\ell)^{-\varphi(1-\sigma)-1}. \end{aligned}$$

Then,

$$\begin{aligned} c_t U_{c_t c_t} &= -\sigma U_{c_t}, \\ n_t U_{n_t n_t} &= -(\varphi(1-\sigma)+1)U_{n_t}, \\ c_t U_{c_t n_t} &= (1-\sigma)U_{n_t}, \\ n_t U_{c_t n_t} &= -\varphi(1-\sigma)U_{c_t}. \end{aligned}$$

so that

$$\begin{aligned} W_{c_t} &= U_{c_t} + \lambda(U_{c_t} + U_{c_t c_t}c_t + U_{n_t c_t}n_t) \\ &= U_{c_t} + \lambda(U_{c_t} - \sigma U_{c_t} - \varphi(1-\sigma)U_{c_t}) \\ &= U_{c_t}(1+\lambda(1-\sigma)(1-\varphi)), \\ W_{n_t} &= U_{n_t} + \lambda(U_{c_t n_t}c_t + U_{n_t} + U_{n_t n_t}n_t) \\ &= U_{n_t} + \lambda((1-\sigma)U_{n_t} + U_{n_t} - (\varphi(1-\sigma)+1)U_{n_t}) \\ &= U_{n_t}(1+\lambda(1-\sigma)(1-\varphi)). \end{aligned}$$

Thus, (13.16) becomes

$$F_{n_t}(k_t, n_t, A_t) = -\frac{U_{n_t}(1+\lambda(1-\sigma)(1-\varphi))}{U_{c_t}(1+\lambda(1-\sigma)(1-\varphi))} = -\frac{U_{n_t}}{U_{c_t}}.$$

Recall that labour tax rate is given by (13.14), which becomes,

$$1 - \tau_{\ell t} = 1 \Rightarrow \tau_{\ell t} = 0$$

for all $t \geq 1$.

Consider the case in which utility is additively separable as given by the second line, then

$$\begin{aligned} U_n = U_{1-\ell} &= -A\,(1-\ell)^{\varphi}\,, \\ U_c &= c^{-\sigma} \\ U_{nn} &= -\varphi A\,(1-\ell)^{\varphi-1} \\ U_{cc} &= -\sigma c^{-\sigma-1} \\ U_{cn} &= 0. \end{aligned}$$

Then,

$$\begin{aligned} cU_{cc} &= -\sigma U_{c_t}, \\ nU_{nn} &= \varphi U_n \end{aligned}$$

so that

$$\begin{aligned} W_{c_t} &= U_{c_t} + \lambda\left(U_{c_t} + U_{c_t c_t} c_t + U_{n_t c_t} n_t\right) \\ &= U_{c_t}\left(1 + \lambda\,(1-\sigma)\right), \\ W_{n_t} &= U_{n_t} + \lambda\left(U_{c_t n_t} c_t + U_{n_t} + U_{n_t n_t} n_t\right) \\ &= U_{n_t}\left(1 + \lambda\,(1+\varphi)\right). \end{aligned}$$

Thus, (13.16) becomes

$$F_{n_t}\left(k_t, n_t, A_t\right) = -\frac{U_{n_t}\left(1 + \lambda\,(1+\varphi)\right)}{U_{c_t}\left(1 + \lambda\,(1-\sigma)\right)}.$$

Recall that labour tax rate is given by (13.14), which becomes,

$$1 - \tau_{\ell t} = \frac{1 + \lambda\,(1-\sigma)}{1 + \lambda\,(1+\varphi)}$$

for all $t \geq 1$. ■

The general intuition is that constant taxes smooth out distortion from taxation across periods—public debt plays a key role here (notice that we did not assume that government has to satisfy period budget constraint, only the present-value budget constraint). When preferences have constant elasticities, we obtain the (extreme) result of constant taxes. In general. optimal taxes will vary, however, they vary much less so than expenditure and capital.

13.3 Ramsey Taxation with stochastic government purchases (labour only, no capital)

In this section, we show that the conclusion of (almost) constant taxes in the deterministic case carries through in the stochastic case with complete markets. Specifically, we analyse the optimal (linear) labour taxes in an economy with:

- ▷ an infinitely lived representative agent;
- ▷ exogenously given stochastic government purchases—only source of shock in the economy;
- ▷ variable labour supply, linear production, and no capital;
- ▷ complete markets with bonds contingent on realisation of "shock";
- ▷ linear labour (or consumption) taxes;
- ▷ Ramsey problem with commitment.

13.3.1 The set up

Let $g_t \in G \subset \mathbb{R}_+$ denote government purchases at time t, and $g^t = (g_0, g_1, \ldots, g_t) \in G^t$ denote the history of government purchases up to time t. The probability of the path are described by the CDF $F^t(g^t)$, where $F^t : G^t \to [0,1]$. Thus, allocations are functions of both t as well as the history of government purchases, g^t . We assume the marginal productivity of labour to be one. Thus, feasibility condition is that

$$c_t\left(g^t\right) + g_t + \ell_t\left(g^t\right) \leq 1,\ \forall t \geq 0,\ \forall g^t \in G^t. \tag{13.25}$$

The agent's (expected) utility function is given by

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty} \beta^t U\left(c_t, l_t\right)\right] = \sum_{t=0}^{\infty} \beta^t \int U\left(c_t\left(g^t\right), l_t\left(g^t\right)\right) dF^t\left(g^t\right).$$

We assume that U is strictly concave in (c, n) and strictly increasing in (c, ℓ), and that it satisfies the standard Inada conditions. The economy has no intertemporal possibilities for consumption goods.

Household budget constraint The commodities in this economy are consumption and leisure indexed by time t and g^t: $c_t(g^t)$ and $\ell_t(g^t)$. Let $p_t(g^t)$ denote the time-0 Arrow-Debreu price of consumption at time t in state g^t. In the deterministic case, p_t is just the present value. Income taxes are contingent on history so $\tau_t(g^t)$. We assume that the government has promised coupons, denoted $b_t(g^t)$, to each agent, which can be thought of as social security. We use that, given linear production function, pre-tax wages in terms of date-t consumption are equal to one. Thus, the Arrow-Debreu budget constraint is given by

$$\sum_{t=0}^{\infty} \int p_t\left(g^t\right)\left[c_t\left(g^t\right)\right] dg^t \leq \sum_{t=0}^{\infty} \int p_t\left(g^t\right)\left[\left(1-\tau_t\left(g^t\right)\right)\left(1-\ell_t\left(g^t\right)\right) + b_t\left(g^t\right)\right] dg^t,$$

where the left-hand side is the present value of consumption and the right-hand side is the present value of income, which consists of post-tax labour income and receipts from coupons. We can rearrange above to

$$\begin{aligned} 0 \geq\ & p_0\left[c_0 - \left(1-\tau_0\right)\left(1-\ell_0\right) - b_0\right] \\ & + \sum_{t=1}^{\infty} \int p_t\left(g^t\right)\left[c_t\left(g^t\right) - \left(1-\tau_t\left(g^t\right)\right)\left(1-\ell_t\left(g^t\right)\right) - b_t\left(g^t\right)\right] dg^t, \end{aligned} \tag{13.26}$$

where we abbreviate the g^0 since there is only one state with respect to government purchases in period zero.

Government budget constraint The government's budget constraint equates the present value of tax revenues with the present value of the sum of government purchases and coupon payments:

$$\sum_{t=0}^{\infty} \int p_t\left(g^t\right)\left[g_t\left(g^t\right)+b_t\left(g^t\right)\right] dg^t \leq \sum_{t=0}^{\infty} \int p_t\left(g^t\right)\left[\tau_t\left(g^t\right)\left(1-\ell_t\left(g^t\right)\right)\right] dg^t .$$

We can rewrite above as

$$0 \geq p_0\left[g_0+b_0-\tau_0\left(1-\ell_0\right)\right]+\sum_{t=1}^{\infty} \int p_t\left(g^t\right)\left[g_t\left(g^t\right)+b_t\left(g^t\right)-\tau_t\left(g^t\right)\left(1-\ell_t\left(g^t\right)\right)\right] dg^t .$$

13.3.2 Equilibrium

A competitive equilibrium with taxes where the government finances the stochastic process for purchases $\{g_t\}$ and transfers $\{b_t\}$ is given by stochastic processes for:

- ▷ an allocation $\{c_t\left(g^t\right), \ell_t\left(g^t\right)\}$;
- ▷ a price system $\{p_t\left(g^t\right)\}$;
- ▷ taxes $\{\tau_t\left(g^t\right)\}$;

such that

(i) $\{c_t\left(g^t\right), \ell_t\left(g^t\right)\}$ is feasible for $\{g_t\}$;

(ii) $\{c_t\left(g^t\right), \ell_t\left(g^t\right)\}$ maximises the household utility given prices, $\{p_t\left(g^t\right)\}$, and taxes, $\{\tau_t\left(g^t\right)\}$, and transfers $\{b_t\left(g^t\right)\}$;

(iii) the government budget constraint holds.

Note that the firm's maximisation problem is subsumed by letting wages equal to one.

In an equilibrium, agents take the stochastic process for prices, $\{p_t\}$, taxes, $\{\tau_t\}$, and transfers, $\{b_t\}$, as given so that, at time $t=0$, they can plan with the knowledge of all future actions (contingent on all possible realisations of g^t). Thus, we implicitly assume that the government can commit to the path of taxes they announce at time $t=0$. We can think of this as a version of rational expectation in the sense that agents correctly anticipate future prices and taxes.

13.3.3 Household's problem

The household problem is given by

$$\begin{aligned}
\max_{\{c_t(g^t),\ell_t(g^t)\}} \quad & \sum_{t=0}^{\infty} \beta^t \int U\left(c_t\left(g^t\right), \ell_t\left(g^t\right)\right) dF^t\left(g^t\right) \\
s.t. \quad & 0 \geq p_0\left[c_0-\left(1-\tau_0\right)\left(1-\ell_0\right)-b_0\right] \\
& \quad +\sum_{t=1}^{\infty} \int p_t\left(g^t\right)\left[c_t\left(g^t\right)-\left(1-\tau_t\left(g^t\right)\right)\left(1-\ell_t\left(g^t\right)\right)-b_t\left(g^t\right)\right] dg^t .
\end{aligned}$$

The first-order conditions for each $t=0,1,\ldots$ and $g^t \in G^t$ are

$$\begin{aligned}
\beta^t U_c\left(c_t\left(g^t\right), \ell_t\left(g^t\right)\right) f^t\left(g^t\right) &= \lambda p_t\left(g^t\right), \\
\beta^t U_\ell\left(c_t\left(g^t\right), \ell_t\left(g^t\right)\right) f^t\left(g^t\right) &= \lambda p_t\left(g^t\right)\left(1-\tau_t\left(g^t\right)\right),
\end{aligned}$$

where λ is the Lagrange multiplier for the budget constraint and f^t is the density ("probability") of a particular history of g^t given information available as at period zero. The intratemporal marginal rate of substitution is given by

$$\frac{U_\ell\left(c_t\left(g^t\right),\ell_t\left(g^t\right)\right)}{U_c\left(c_t\left(g^t\right),\ell_t\left(g^t\right)\right)}=1-\tau_t\left(g^t\right), \tag{13.27}$$

where wages do not appear on the right-hand side since we normalised labour productivity to one. The intertemporal marginal rate of substitution between period-t consumption and period-0 consumption given by

$$\beta^t\frac{U_c\left(c_t\left(g^t\right),\ell_t\left(g^t\right)\right)}{U_c\left(c_0,\ell_0\right)}f^t\left(g^t\right)=\frac{p_t\left(g^t\right)}{p_0}, \tag{13.28}$$

where we note that $f^0\left(g^0\right)=1$.

13.3.4 Implementability

We now summarise the type of allocation that can constitute an equilibrium with taxes that finances the exogenously given stochastic stream of government purchases.

As before, we obtain implementability constraint by eliminating prices and taxes from the household's budget constraint using the first-order conditions—i.e. intertemporal and intratemporal marginal rates of substitution—from the household's problem. As before, by Walras' Law, we can also interpret the implementability constraint as the budget constraint of the government, where prices have been replaced by the relevant marginal rates of substitution.

First, divide the household's budget constraint, (13.26), through by p_0 (while setting inequality to bind):

$$0=\left[c_0-\left(1-\tau_0\right)\left(1-\ell_0\right)-b_0\right]+\sum_{t=1}^{\infty}\int\frac{p_t}{p_0}\left[c_t-\left(1-\tau_t\right)\left(1-\ell_t\right)-b_t\right]dg^t,$$

where we simplified the notation by removing the explicit reference of time-t consumption and leisure being a function of g^t. We then use (13.27) and (13.28) to eliminate $\left(1-\tau_t\left(g^t\right)\right)$ and $p_t\left(g^t\right)/p_0$:

$$0=\left[c_0-\frac{U_\ell\left(c_0,\ell_0\right)}{U_c\left(c_0,\ell_0\right)}\left(1-\ell_0\right)-b_0\right]+\sum_{t=1}^{\infty}\int\beta^t\frac{U_c\left(c_t,\ell_t\right)}{U_c\left(c_0,\ell_0\right)}f^t\left(g^t\right)\left[c_t-\frac{U_\ell\left(c_t,\ell_t\right)}{U_c\left(c_t,\ell_t\right)}\left(1-\ell_t\right)-b_t\right]dg^t.$$

Multiplying both sides by $U_c\left(c_0,\ell_0\right)$ while noting that $f^t\left(g^t\right)dg^t=dF^t\left(g^t\right)$,

$$\begin{aligned}0=&\left(c_0-b_0\right)U_c\left(c_0,\ell_0\right)-\left(1-\ell_0\right)U_\ell\left(c_0,\ell_0\right)\\&+\sum_{t=1}^{\infty}\beta^t\int\left[\left(c_t-b_t\right)U_c\left(c_t,\ell_t\right)-\left(1-\ell_t\right)U_\ell\left(c_t,\ell_t\right)\right]dF^t\left(g^t\right).\end{aligned} \tag{13.29}$$

Above is the implementability constraint in the current set-up.

13.3.5 Ramsey Problem

The Ramsey Problem is to find the policy $\{\tau_t\}$ that support a competitive equilibrium with taxes; i.e. to find the "optimal" taxes to finance a given stream of government purchases, subject to the fact that agents behave competitively for given policy. With our characterisation of an equilibrium with taxes, the Ramsey problem can be solved by finding the allocation that maximises the utility of the agents subject to feasibility, given by (13.25), for all $t\geq 0$, and subject to the implementability constraint, (13.29); i.e.

$$\begin{aligned}\max_{\{c_t(g^t),\ell_t(g^t)\}} \quad & \sum_{t=0}^{\infty}\beta^t\int U\left(c_t,\ell_t\right)dF^t\left(g^t\right)\\ s.t.\quad & c_t+g_t+\ell_t\leq 1,\forall t\geq 0,\ \forall g^t\in G^t,\\ & 0=\left(c_0-b_0\right)U_c\left(c_0,\ell_0\right)-\left(1-\ell_0\right)U_l\left(c_0,\ell_0\right)\\ & \quad+\sum_{t=1}^{\infty}\beta^t\int\left[\left(c_t-b_t\right)U_c\left(c_t,\ell_t\right)-\left(1-\ell_t\right)U_l\left(c_t,\ell_t\right)\right]dF^t\left(g^t\right).\end{aligned}$$

It is useful to rewrite the Ramsey problem above as follows.

$$\begin{aligned}\min_{\lambda}\max_{\{c_t(g^t),\ell_t(g^t)\}} & \sum_{t=0}^{\infty}\beta^t\int U\left(c_t,\ell_t\right)+\lambda\left[\left(c_t-b_t\right)U_c\left(c_t,l_t\right)-\left(1-l_t\right)U_l\left(c_t,l_t\right)\right]dF^t\left(g^t\right) \\ s.t.\qquad & c_t+g_t+\ell_t=1,\ \forall t\geq 0,\ \forall g^t\in G^t,\end{aligned}\tag{13.30}$$

where λ is the Lagrange multiplier for the implementability constraint. Written above, this is a sequence of static problem for a given λ (except for b_t). Define the period-state utility as

$$W\left(c,\ell;b_t,\lambda\right)\coloneqq U\left(c,l\right)+\lambda\left[\left(c-b_t\right)U_c\left(c,\ell\right)-\left(1-\ell\right)U_\ell\left(c,\ell\right)\right].$$

As before, we can interpret W as the social value of (c,ℓ). The first term represents the private value for the agent, and the second term is the net tax revenue (in terms of utils), corresponding to period-t, state-g^t allocation, representing the penalty from violating the constraint.

The first-order conditions for the Ramsey problem are:

$$\begin{aligned}W_c\left(c,\ell;b_t,\lambda\right)f^t\left(g^t\right)&=\mu_t,\\ W_\ell\left(c,\ell;b_t,\lambda\right)f^t\left(g^t\right)&=\mu_t,\end{aligned}$$

for all $t\geq 0$ and $g^t\in G^t$, where μ is the Lagrange multiplier on the feasibility constraint. Rearranging gives,

$$\frac{W_\ell\left(c,\ell;b_t,\lambda\right)}{W_c\left(c,\ell;b_t,\lambda\right)}=1.$$

We also have the feasibility constraint,

$$c_t+g_t+\ell_t=1,\ \forall t\geq 0,\ \forall g^t\in G^t.$$

Proposition 13.4. *In any period in which b_t and g_t coincide, τ_t are the same.*

Proof. Recall that (13.30) is a sequence of static problem for a given λ and b_t; i.e. for any given λ, and t, t' in which $b_t=b_{t'}$, $c_t\left(g^t\right)=c_{t'}\left(g^t\right)=c\left(g^t\right)$ and $\ell_t\left(g^t\right)=\ell_{t'}\left(g^t\right)=\ell_t\left(g^t\right)$, for any $g^t\in G^t$. It follows then that, for any given λ, and t, t' in which $b_t=b_{t'}$,

$$U\left(c_t\left(g^t\right),\ell_t\left(g^t\right)\right)=U\left(c_{t'}\left(g^t\right),\ell_{t'}\left(g^t\right)\right),\ \forall g^t\in G^t.$$

Given that tax rates are given by (13.27), we therefore obtain that

$$1-\tau_{t'}\left(g^t\right)=1-\tau_{t'}\left(g^t\right),\ \forall g^t\in G^t. \qquad\blacksquare$$

The next proposition says that even in periods in which government expenditures are zero, the optimal policy involves taxing the agents in such periods. The intuition is that some tax is better than none since, starting from zero taxes, it has no first-order loss (there is a second-order loss associated with the "Harberger Triangle"), but there is a first-order increase in revenue).

Proposition 13.5. *In any period in which $b_t = g_t = 0$, then $\tau_\ell > 0$ if $\lambda > 0$.*

Proof. See problem set 8, question 2 (note that the problem set uses l to denote labour, not leisure as in the notes). ■

Recall that when preferences exhibit constant elasticities of substitution, then the optimal labour tax does not vary over time, even if not at the steady state. Lucas-Stokey present 9 other examples in which the the same result holds. In each example, they specify processes for $\{g_t, b_t\}$, focusing on the optimal tax (and management of debt). The existence of complete markets mean that taxes depend on current g only.

Example 13.3. *(Deterministic expenditures).* We can use one-period debt or multi-period debt.

- ▷ $b_t = g_t = 0$ for all $t \geq 0$ and $g^t \in G^t$. No taxes and we obtain the first best. Of course, no debt.
- ▷ $b_t + g_t = 0$ for all $t \geq 0$ and $g^t \in G^t$. Again, no taxes, and we obtain the first best and no debt. But note that $b_t < 0$ if $g_t > 0$ (i.e. coupon payments are negative).
- ▷ $b_t = \bar{B}$, $g_t = \bar{G}$ for all $t \geq 0$ and $g^t \in G^t$ with $\bar{G} + \bar{B} > 0$. With constant expenditures, there is no need to issue new debt.
- ▷ $b_t = 0$ for all $t \geq 0$ and $g^t \in G^t$, but $g_T > 0$ and $g_t = 0$ for all $t \neq T$; i.e. there is government purchases in just one period, period T. Then, allocations are constant for all $t \neq T$ with positive taxes. Government saves before T (rolling over one-period assets) and pays the debt after T (paying interest in each period). Note that the the present value of all tax income is equal to the government expenditures (in this case, the present value of g_T). This means that, at $t = T$, the government has not collected enough tax income to be able to pay g_t. Thus, the government borrows money in periods before so that, in period $t = T$, it can pay g_T. It then repays the debt with subsequent tax income.

Example 13.4. *(Stochastic expenditures).* We can restrict to period-contingent debt—a one-period contingent debt purchased at t pays one unit in $t+1$ if $g_{t+1} = g'$ and zero otherwise for each value of g' that is possible at $t+1$.

- ▷ $b_t = 0$ for all $t \geq 0$ and $g_t \in G^t$, but $g_T > 0$ with probability α and $g_T = 0$ otherwise, and $g_T = 0$ for all $t \neq T$. This is the stochastic version of the last example from the deterministic case. Then, again, allocations are constant for all $t \neq T$ and positive taxes in those dates. The government saves before T and pays the debt thereafter.
  - ▷ At each date $t = 0, 1, \ldots, T-2$, the government buys bonds and accumulates assets.
  - ▷ At $t = T-1$, the government buys bonds contingent on $g_T > 0$.
  - ▷ In period $t = T$, suppose that $g_T > 0$, then the contingent bond pays g_T, which it can use to purchase g_T amount of goods. The government debt left is then the borrowing prior to period T plus the purchase price of the contingent bond. If instead $g_T = 0$, the contingent bond pays nothing, but, at the same time, the government does not have to purchase any goods. Thus, the government again has debt equal to the borrowing prior to period T plus the purchase price of the contingent bond.
  - ▷ At $t = T+1$, regardless of whether $g_T > 0$, the government inherits the same debt and rolls it over.
- ▷ $b_t = 0$ for all $t \geq 0$ and $g^t \in G^t$, and $g_t = \bar{G}$ for $t = T, T+S, T+2S$, where $0 \leq T \leq S$ and$g_0 = 0$ otherwise.

- ▷ Two types of allocations, one for $t = T, T+S, T+2S$, and another for the other t.
- ▷ In period $t = 0, 1, \ldots, T-1$, the government runs a surplus, and, at $t = T$, uses the surplus to pay for the deficit.
- ▷ The government repeats the cycle every S periods.

▷ $b_t = 0$ for all $t \geq 0$ and $g^t \in G^t$, and $g_0 = \bar{G} > 0$. If $g_t = \bar{G}$, then $g_{t+1} = \bar{G}$ with probability α and $g_{t+1} = 0$ otherwise.

- ▷ Two types of allocations, one for each value of g.
- ▷ Financed with (the same portfolio of) contingent bonds while expenditure is positive.
- ▷ The government buys more bonds at t when $g_t > 0$ contingent on $g_{t+1} > 0$.
- ▷ It switches to rolling over debt after g reaches zero.

▷ $b_t = 0$ for all $t \geq 0$ and $g^t \in G^t$, and $\{g_t\}$ is a sequence of iid random variables.

- ▷ Allocation depends on the current value of g.
- ▷ In each period, think of first paying the excess of expenditure over surplus.
- ▷ In each period, buy a portfolio of one-period bond with zero coupon paying the amount equal to primary surplus/deficit.
- ▷ Value of the portfolio to be purchased at each t is always the same.

▷ $b_t = 0$ for all $t \geq 0$ and $g^t \in G^t$, and $\{g_t\}$ follows a Markov process.

- ▷ Allocation depends on the current value of g.
- ▷ In each period, buy a portfolio of one-period bond with zero coupon paying the amount equal to the primary surplus/deficit.
- ▷ Value of the one-period portfolio at t depends on the current value of g.