Macroeconomics I

Taught by Larry Christiano
Northwestern University, fall 2015

Ludvig Sinander
Northwestern University

This version: 28 October 2018

These notes are based on a macroeconomics course for first-year PhD students taught by Larry Christiano at Northwestern in fall 2015. The topics are (roughly speaking) dynamic programming, growth, and business cycles.

Contents

1   Mon 21 Sep 2015

1.1   Introduction

Macroeconomics   Macro is defined by its questions. Some of these are

(1) How and why do output, inflation, exchange rates, unemployment etc. fluctuate and co-move? Both over the long run (growth) and short run (business cycles). Includes 2007–08 crisis, slow vs. fast recoveries.

(2) Why do current account imbalances arise and persist, and what are their consequences?

(3) Does a market economy produce good outcomes, or should policymakers intervene actively? We answer normative questions of this sort using the Pareto criterion, taking as given preferences and technological constraints.

Books   We'll mainly be using Stokey, Lucas and Prescott (1989). Later on we'll also use Ljungqvist and Sargent (2012).

Course outline   The focus of this course is on technical tools rather than on economics per se. We first introduce a simple, deterministic environment called the neoclassical growth model,[1] with the aim of characterising efficient allocations. We next define markets and competitive equilibria, and prove welfare theorems that address how well decentralised markets work. We then move on to growth theory more broadly, introducing elements that break the link between Pareto optima and market equilibria. Finally we introduce stochastic shocks, yielding equilibria in the form of stochastic process rather than deterministic paths. We use this elaboration to study business cycles, and find that market equilibria may differ greatly from Pareto optima in this context.

Today   We first define the neoclassical growth model, and set up the social planner's problem as an optimisation problem in a sequence space. We aim to reformulate the social planner's program as a fixed-point problem in a functional space. To this end, we introduce the (far more general) Stokey–Lucas canonical dynamic program. We introduce some assumptions (4.3–4.5 and 4.7–4.9 in the book). We then establish the joint sufficiency of the Euler

---

[1]Though there is no long-run growth in this model, funnily enough.

equation and the transversality condition for optimality in the canonical dynamic program (Theorem 4.15 in the book).

1.2   The neoclassical growth model

This section corresponds to sections 2.1 and 5.1 of Stokey, Lucas and Prescott (1989).

The economy is populated by identical households. Households value time-t consumption $c_t \geq 0$ according to a period utility function $u : \mathbf{R}_+ \to \mathbf{R}$, assumed strictly increasing, strictly concave, twice differentiable, once continuously differentiable, and satisfying the Inada conditions

$$\lim_{c \to 0} u'(c) = \infty \quad \text{and} \quad \lim_{c \to \infty} u'(c) = 0.$$

Consumers' preferences over consumption paths $c = (c_0, c_1, c_2, \dots)$ are the discounted sum of of period utilities, i.e. $U : \mathbf{R}_+^\infty \to \mathbf{R}$ is given by

$$U(c) = \sum_{t=0}^{\infty} \beta^t u(c_t) \quad \text{for any } c \in \mathbf{R}_+^\infty$$

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

The economy has a capital stock k_t at time t. Technology allows $\phi(k_t)$ units of output to be produced when the capital stock is k_t, where $\phi : \mathbf{R}_+ \to \mathbf{R}_+$ is the production function. This output may be used for consumption c_t or investment i_t. Hence the economy's resource constraint is

$$c_t + i_t \leq \phi(k_t) \quad \text{for all } t \in \mathbf{Z}_+.$$

The capital stock evolves according to

$$k_{t+1} = i_t + (1 - \delta)k_t \quad \text{for all } t \in \mathbf{Z}_+,$$

where i_t is investment and $\delta \in (0, 1)$ captures depreciation. The exogenous initial level of capital is $k_0 \geq 0$. Combining the resource constraint with the law of motion gives

$$k_{t+1} \leq \phi(k_t) - c_t + (1 - \delta)k_t.$$

Define $f : \mathbf{R}_+ \to \mathbf{R}_+$ by $f(k) := \phi(k) + (1 - \delta)k$. This function gives the total amount of final good that can be produced from initial capital k, including new output $\phi(k)$ and leftover capital $(1 - \delta)k$. Then the combined resource constraint is

$$k_{t+1} \leq f(k_t) - c_t.$$

We also impose $k_{t+1} \geq 0 \ \forall t \in \mathbf{Z}_+$, i.e. the capital stock cannot be negative. But we do not impose any other lower bound on k_{t+1}. This means that when $k_{t+1} = 0$ is chosen, consumption is

$$c_t = f(k_t) = \phi(k_t) + (1 - \delta)k_t.$$

So the household can eat not just output $\phi(k_t)$, but also the capital stock $(1 - \delta)k_t$! We might therefore want to impose $i_t \geq 0$, so that $k_{t+1} \geq (1 - \delta)k_t$ (only output can be consumed). An exercise on homework 3 shows that all the results in the neoclassical model are retained when $k_{t+1} \geq 0$ is strengthened to $i_t \geq 0$.

The function f is assumed to be strictly increasing, strictly concave and twice differentiable, and to satisfy the Inada conditions

$$\lim_{k \to 0} f'(k) = \infty \quad \text{and} \quad \lim_{k \to \infty} f'(k) = 1 - \delta.$$

Moreover $f(0) = 0$: you can't produce something from nothing.[2]

This completes the description of the environment. We're interested in the social planner's problem of maximising household utility subject to the economy's constraints:

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$
$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad k_{t+1} \leq f(k_t) - c_t$$
$$c_t, k_{t+1} \geq 0.$$

Since u is strictly increasing, it will always be optimal for the resource constraint to hold with equality, so we can write the program as

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(f(k_t) - k_{t+1})$$
$$\text{s.t. } \quad k_{t+1} \in [0, f(k_t)] \quad \text{for all } t \in \mathbf{Z}_+.$$

We're interested in both qualitative and quantitative features of the solution to this program. A qualitative question is whether there is a $k^\star$, independent of k_0, to which the capital stock converges at the solution (a stable steady state). If so, what are the transitional dynamics like? A

---

[2] We could impose these conditions directly on ϕ if we wanted. The conditions would be the same, except that the second Inada condition would be $\lim_{k \to \infty} \phi'(k) = 0$.

quantitative question is what the rate of convergence is, i.e. how big $T \in \mathbf{Z}_+$ has to be to have $|k_t - k^\star| < \varepsilon$ for all $t \geq T$.

Initially, we'll be studying the program above: an optimisation problem in a sequence space. We'll begin by generalising beyond the confines of this model to study a general class of sequential optimisation problems.

1.3   The Stokey–Lucas canonical model

In period t, the state x_t is inherited while the state x_{t+1} is chosen; x_0 is given. Choice x_{t+1} is feasible at x_t iff $x_{t+1} \in \Gamma(x_t)$ where $\Gamma : X \Rightarrow X$ is called the constraint correspondence. Let A be the graph of Γ:

$$A := \left\{ (x, y) \in X^2 : y \in \Gamma(x) \right\}.$$

Period payoffs are given by the function $F : A \to \mathbf{R}$, and are aggregated across time according to

$$\sum_{t=0}^{\infty} \beta^t F\left(x_t, x_{t+1}\right).$$

where $\beta \in (0, 1)$ is a discount factor. Optimal plans $\{x_{t+1}\}_{t=0}^{\infty}$ maximise payoffs subject to $x_{t+1} \in \Gamma(x_t)$ for each $t \in \mathbf{Z}_+$.

The neoclassical growth model fits into this framework. The state variable x_t is the capital stock k_t. The constraint correspondence is given by

$$\Gamma(k) = [0, f(k)],$$

and $A = \{(k, k') \in \mathbf{R}_+^2 : k' \in \Gamma(k)\}$. Period payoffs are

$$F(k, k') = u\left(f(k) - k'\right).$$

We now give a list of assumptions on the general model.

Assumption 1 (SLP Ass'n 4.3). $X \subseteq \mathbf{R}^l$ is convex. Γ is continuous and nonempty- and compact-valued.

Convexity will be needed when we want to get concavity (or quasi-concavity) of the objective function, but actually the first few results below will not make use of this property.[3] The restriction that X is a subset of a Euclidean space can be dispensed with for most results, but requires some

---

[3]In particular, Theorems 4.6 and 4.7 in Stokey et al. (1989) do not require convexity of X in their proofs.

additional technicalities. Continuity of Γ will be needed to apply Berge's theorem to get a continuous value function. Compact-valuedness of Γ will combine with continuity of F to guarantee that a maximiser exists.

In the neoclassical growth model, Γ's values are closed and bounded intervals of $\mathbf{R}_+$. Moreover the upper and lower bounds of these intervals are continuous, so Γ is continuous.

Assumption 2 (SLP Ass'n 4.4). *F is continuous and bounded, and $\beta \in (0, 1)$.*

As already mentioned, continuity together with compactness guarantees the existence of a maximiser. Bounded F means that there are $\underline{F}, \overline{F} \in \mathbf{R}$ such that $\underline{F} \leq F(x, x') \leq \overline{F}$ for all $(x, x') \in A$. It follows that

$$\frac{\underline{F}}{1 - \beta} \leq \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \leq \frac{\overline{F}}{1 - \beta},$$

and hence the objective function is finite at any sequence $\{x_t\}_{t=0}^{\infty}$.[4]

In the neoclassical growth model, we can make $F(k, k') = u\left(f(k) - k'\right)$ as large as we want by setting $k' = 0$ and increasing k, so the period payoff function is not bounded if we take X to be all of $\mathbf{R}_+$. To get around this, let's find a way of restricting X to a compact subset of $\mathbf{R}_+$ without loss of generality. In particular, suppose we can find a $\overline{k} \in \mathbf{R}_+$ such that whenever $k_t \geq \overline{k}$, $k_{t+1} > k_t$ is technologically infeasible. Then it would be the case that every feasible sequence satisfies $k_t \leq \max\{k_0, \overline{k}\}$ for each $t \in \mathbf{Z}_+$. We could therefore set $X = [0, \max\{k_0, \overline{k}\}]$ without loss of generality. Since X is compact and Γ is compact-valued, A is compact. Hence the continuous function F has a maximum on A, and is therefore bounded above as desired.

Now let's show that there is such a $\overline{k}$. The argument is based on the intuition from Figure 1. By the Inada conditions, f has slope greater than unity at 0, and has slope less than unity for k sufficiently large (since $1-\delta < 1$). Moreover, f is stricly increasing, strictly concave, and continuous. It follows that f cuts the diagonal exactly once on the interior of $\mathbf{R}_+$, and that it cuts from above. (Again, look at Figure 1!) Call this intersection $\overline{k}$. If $k > \overline{k}$, then

---

[4]Many results can be extended to the case where the objective is well-defined but possibly infinite. The formal equivalence of the sequence problem and functional equation still holds (Stokey et al., 1989, Theorems 4.2–4.5). If two distinct sequences both produce infinite value, we can compare them using the overtaking criterion (or similar). But this is a hassle, both for the substantial reason that the overtaking criterion only defines a partial order on the set of feasible sequences (so we may have a large set of undominated sequences), and because there are extra technicalities involved.
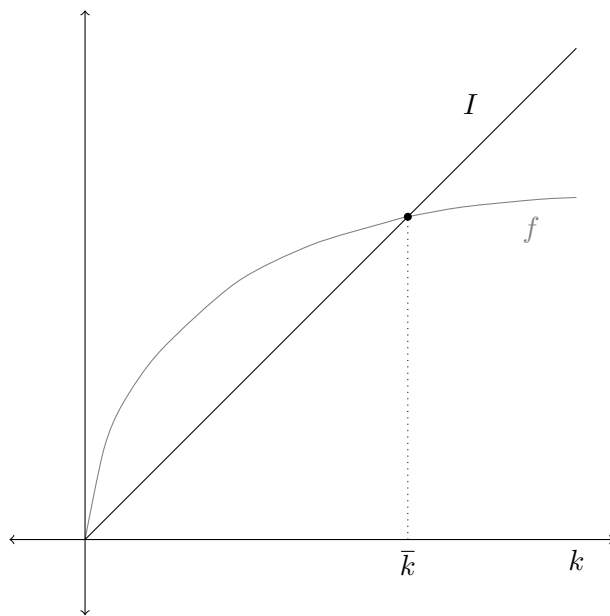
Figure 1 – 'Endogenous' bound on f. Drawn are f and the identity function I. When $k > k^\star$, even if we set $c = 0$, the capital stock falls until it stabilises at $k^\star$. Hence $k_t \leq \max\{k_0, \bar{k}\}$ at all $t \in \mathbf{Z}_+$.

$f(k) < k$, so any feasible k' (i.e. $k' \in [0, f(k)]$) satisfies $k' < k$. This is what we wanted to show.

Here are some more assumptions we might impose on the Stokey–Lucas canonical model.

Assumption 3 (SLP Ass'n 4.5)**.** For each $x' \in X$, $F(\cdot, x')$ is strictly increasing.

This is obviously satisfied in the neoclassical growth model.

Assumption 4 (SLP Ass'n 4.8)**.** (Γ is such that) A is convex.

Assumption 5 (SLP Ass'n 4.7)**.** F is strictly concave.

Strict concavity is easily verified in the neoclassical growth model by differentiating. Note that imposing Assumption 5 without Assumption 4 makes no sense becaues concavity is defined only for functions with convex domain.

Assumption 6 (SLP Ass'n 4.9)**.** F is continuously differentiable on the interior of A.

1.4 The variational approach

Unlike dynamic programming, the variational approach to dynamic optimisation works directly with the optimisation problem in sequence space. The approach is to consider a candidate solution $x^\star \in \mathbf{R}^\infty$, then perturb it. If the candidate solution is optimal, then perturbations should not have a first-order effect on the value. This reasoning yields a first-order condition, called an Euler equation, for each coordinate of the sequence. We pair this with a transversality condition that pins down the level of a sequence. Under (strict) assumptions, the optimal sequence is characterised by the Euler equation and transversality condition.

To say more, partition the vector $\nabla F(x, y)$ of partial derivatives as

$$\nabla F(x, y) = \begin{pmatrix} F_1(x, y) \\ F_2(x, y) \end{pmatrix}.$$

The following is called the Euler equation:

$$F_2(x_t, x_{t+1}) + \beta F_1(x_{t+1}, x_{t+2}) = 0 \quad \text{for each } t \in \mathbf{Z}_+. \tag{1}$$

This one is the transversality condition:

$$\lim_{T \to \infty} \beta^T F_1(x_T, x_{T+1}) \cdot x_T = 0. \tag{2}$$

13

Theorem 1 (SLP Th'm 4.15). Suppose Assumptions 4.3, 4.4, 4.5, 4.8, 4.7 and 4.9 hold and that $X \subseteq \mathbf{R}_+^l$. Let $\{x_{t+1}^\star\}_{t=0}^\infty$ be a sequence with $x_{t+1}^\star \in \operatorname{int} \Gamma(x_t^\star) \; \forall t \in \mathbf{Z}_+$ that satisfies the Euler equation (1) and the transversality condition (2). Then $\{x_{t+1}^\star\}_{t=0}^\infty$ is optimal.

Proof. Fix a feasible sequence $\{x_{t+1}\}$, and define

$$D := \lim_{T \to \infty} \sum_{t=0}^{T} \beta^t \left[F(x_t^\star, x_{t+1}^\star) - F(x_t, x_{t+1}) \right]$$

Assumption 4.4 guarantees that D is well-defined. We want to show that $D \geq 0$.

Since F is strictly concave, F lies below all its tangents, so for any (x, y) and $(x^\star, y^\star)$ in A,

$$F(x, y) \leq F(x^\star, y^\star) + \nabla F(x^\star, y^\star) \cdot \begin{pmatrix} x - x^\star \\ y - y^\star \end{pmatrix}.$$

Rearranging,

$$F(x^\star, y^\star) - F(x, y) \geq F_1(x^\star, y^\star) \cdot (x^\star - x) + F_2(x^\star, y^\star) \cdot (y^\star - y).$$

So for $(x, y) = (x_t, x_{t+1})$ and $(x^\star, y^\star) = (x_t^\star, x_{t+1}^\star)$,

$$D \geq \lim_{T \to \infty} \sum_{t=0}^{T} \beta^t \left[F_1(x_t^\star, x_{t+1}^\star) \cdot (x_t^\star - x_t) + F_2(x_t^\star, x_{t+1}^\star) \cdot (x_{t+1}^\star - x_{t+1}) \right]$$

$$= \lim_{T \to \infty} \sum_{t=0}^{T-1} \beta^t \left[\beta F_1(x_{t+1}^\star, x_{t+2}^\star) + F_2(x_t^\star, x_{t+1}^\star) \right] \cdot (x_{t+1}^\star - x_{t+1})$$

$$+ F_1(x_0^\star, x_1^\star) \cdot (x_0^\star - x_0) + \lim_{T \to \infty} \beta^T F_2(x_T^\star, x_{T+1}^\star) \cdot (x_{T+1}^\star - x_{T+1}).$$

The first term is zero by the Euler equation (1), and the second term is zero since $x_0^\star = x_0$ (because x_0 is fixed). Hence

$$D \geq \lim_{T \to \infty} \beta^T F_2(x_T^\star, x_{T+1}^\star) \cdot (x_{T+1}^\star - x_{T+1})$$

$$= -\lim_{T \to \infty} \beta^{T+1} F_1(x_{T+1}^\star, x_{T+2}^\star) \cdot (x_{T+1}^\star - x_{T+1})$$

$$\geq -\lim_{T \to \infty} \beta^{T+1} F_1(x_{T+1}^\star, x_{T+2}^\star) \cdot x_{T+1}^\star = 0$$

by the Euler equation (1), the fact that F_1 and x_{T+1} are nonnegative (since $F(\cdot, y)$ is increasing and since $X \subseteq \mathbf{R}_+^l$), and the transversality condition (2). ∎

Theorem 4.15 shows that under certain assumptions, the Euler equation and transversality condition are sufficient for an optimum. But note well that it does not establish that the Euler equation and transversality conditions are *necessary* for an optimum.

Below, I provide two results on when the Euler equation and transversality condition are necessary for optimality. They make use of material covered later in the course, so come back to them if this is your first pass through the material.

Proposition 1 (necessity of Euler eq'n)**.** Suppose Assumptions 4.3, 4.4, 4.8, 4.7 and 4.9 hold and that $X \subseteq \mathbf{R}^l$. Let $\{x^\star_{t+1}\}^\infty_{t=0}$ be an optimal sequence with $x^\star_{t+1} \in \operatorname{int} \Gamma(x^\star_t) \ \forall t \in \mathbf{Z}_+$. Then this sequence satisfies the Euler equation (1).

The conditions for this result are exactly the same as for the sufficiency theorem, except that we weakened $X \subseteq \mathbf{R}^l_+$ to $X \subseteq \mathbf{R}^l$.

Proof. By SLP Theorem 4.11, the value function is differentiable with

$$v'(x^\star_t) = F_1(x^\star_t, x^\star_{t+1}) \quad \forall t \in \mathbf{Z}_+.$$

Since v is differentiable and since each $x^\star_{t+1}$ is interior, the first-order condition must hold:

$$F_2(x^\star_t, x^\star_{t+1}) + \beta v'(x^\star_{t+1}) = 0 \quad \forall t \in \mathbf{Z}_+.$$

Combining yields the Euler equation. ∎

Proposition 2 (necessity of transversality cond'n)**.** Suppose Assumptions 4.3, 4.4, 4.8, 4.7 and 4.9 hold and that $X \subseteq \mathbf{R}^l_+$. Further suppose that Assumptions 4.5 and 4.6 hold and that $0 \in X$. Let $\{x^\star_{t+1}\}^\infty_{t=0}$ be an optimal sequence with $x^\star_{t+1} \in \operatorname{int} \Gamma(x^\star_t) \ \forall t \in \mathbf{Z}_+$. Then this sequence satisfies the transversality condition (2).

The conditions are the same as for the sufficiency theorem, except that we add Assumptions 4.5 and 4.6 (to get monotonicity of v) and $0 \in X$.

Proof. By SLP Theorem 4.11, v is differentiable with

$$v'(x^\star_T) = F_1(x^\star_T, x^\star_{T+1}).$$

Since $0 \in X$ and since v is concave by SLP Theorem 4.9,

$$v(0) - v(x^\star_T) \le v'(x^\star_T) \cdot (0 - x^\star_T) = -F_1(x^\star_T, x^\star_{T+1}) \cdot x^\star_T.$$

Multiplying by β^T and taking limits,

$$\lim_{T\to\infty} \beta^T \left[v(0) - v\left(x_T^\star\right)\right] \leq -\lim_{T\to\infty} \beta^T F_1\left(x_T^\star, x_{T+1}^\star\right) \cdot x_T^\star$$

Since v is bounded by SLP Theorem 4.6, it follows that

$$\lim_{T\to\infty} \beta^T F_1\left(x_T^\star, x_{T+1}^\star\right) \cdot x_T^\star \leq 0.$$

Now we make use of Assumptions 4.5 and 4.6: by SLP Theorem 4.7, v is increasing, so

$$F_1\left(x_T^\star, x_{T+1}^\star\right) = v'\left(x_T^\star\right) \geq 0.$$

Since $x_T^\star \geq 0$, it follows that

$$\lim_{T\to\infty} \beta^T F_1\left(x_T^\star, x_{T+1}^\star\right) \cdot x_T^\star \geq 0. \qquad \blacksquare$$

2   Wed 23 Sep 2015

2.1   Outline of today

Last time, we took a direct, variational approach to the sequential optim-isation problem, maximising the objective over the constraint set using first-order conditions for the optimal sequence.

We begin today by interpreting Theorem 4.15, which told us that the Euler equation and transversality condition are jointly sufficient for an interior optimum in the sequence problem.

We the examine the extent to which we can characterise the optimal paths of capital and consumption in the neoclassical growth model using Theorem 4.15. This will be a disappointment: without functional form assumptions, the Euler equation and transversality condition don't tell us much about the paths.

To get further, we reformulate the sequential optimisation problem in recursive form. We show that the sequential problem is equivalent to a functional equation. We prove theorems connecting the value functions and optimal sequences associated with the sequential problem and functional equation. We study how the optimal sequence can be found from the solution to the functional equation.

2.2   Interpreting the Euler equation and transversality condition

Recall (the sequential formulation of) the Stokey–Lucas canonical problem:

$$\max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F\left(x_t, x_{t+1}\right)$$

$$\text{s.t.} \quad x_{t+1} \in \Gamma\left(x_t\right) \quad \forall t \in \mathbf{Z}_+.$$

Theorem 4.15 told us that under assumptions on the primitives, the Euler equation

$$F_2(x_t, x_{t+1}) + \beta F_1(x_{t+1}, x_{t+2}) = 0 \quad \text{for each } t \in \mathbf{Z}_+$$

and transversality condition

$$\lim_{T \to \infty} \beta^T F_1(x_T, x_{T+1}) \cdot x_T = 0$$

are jointly sufficient for an interior solution.

In the neoclassical growth model, the sequence problem is

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(f(k_t) - k_{t+1})$$

$$\text{s.t.} \quad k_{t+1} \in [0, f(k_t)] \quad \text{for } t \in \mathbf{Z}_+.$$

So $F(k, k') = u(f(k) - k')$, and hence

$$F_2(k_t, k_{t+1}) = -u'(c_t)$$
$$F_1(k_{t+1}, k_{t+2}) = f'(k_{t+1})u'(c_{t+1}).$$

(Recall that $c_t = f(k_t) - k_{t+1}$ by the resource constraint with equality.) The Euler equation is therefore

$$u'(c_t) = \beta u'(c_{t+1})f'(k_{t+1}) \quad \forall t \in \mathbf{Z}_+,$$

or equivalently

$$\frac{\beta^t u'(c_t)}{\beta^{t+1} u'(c_{t+1})} = f'(k_{t+1}) \quad \forall t \in \mathbf{Z}_+.$$

The idea is that you equate the marginal rate of substitution between c_t and c_{t+1} (LHS) with the marginal rate of transformation $f'(k_{t+1})$. More verbosely, when you increase today's consumption c_t, you decrease saving and hence tomorrow's capital stock. This results in a marginal decrease of $f'(k_{t+1})$ in tomorrow's consumption.

If we define $w : \mathbf{R}_+^3 \to \mathbf{R}$ by

$$w(x, y, z) = F_2(x, y) + \beta F_1(y, z) \quad \forall (x, y, z) \in \mathbf{R}_+^3,$$

then the Euler equation is

$$w(k_t, k_{t+1}, k_{t+2}) = 0 \quad \text{for } t \in \mathbf{Z}_+.$$

k_0 is fixed. Fix $k_1 \in \mathbf{R}_+$. Then under regularity conditions on F_1 and F_2, k_2 is uniquely pinned down by $w(k_0, k_1, k_2) = 0$, and hence k_3 is pinned down by $w(k_1, k_2, k_3) = 0$, etc. So conditional on k_0 and k_1, there is only one sequence satisfying the Euler equation. But correspondingly, for each $k_1 \in \mathbf{R}_+$ there is a distinct solution to the Euler equation. Another way to look at this is that the Euler equation is a second-order difference equation, and we only have one initial condition, viz. k_0.

Since some choices of k_1 are clearly suboptimal, we definitely need another initial condition if we are to identify only optima. To motivate the condition we'll settle on, consider the finite-horizon problem

$$\max_{(k_1,\dots,k_{T+1})} \sum_{t=0}^{T} \beta^t F(k_t, k_{t+1}) \quad \text{s.t.} \quad k_{t+1} \in [0, f(k_t)] \quad \forall t \in \mathbf{Z}_+.$$

Obviously the optimal k_{T+1} is zero, since you get no future payoff from saving in the last period. So for the finite-horizon case, we have a second-order difference equation (the Euler equation) and two boundary conditions with the simple form k_0 given and $k_{T+1} = 0$. (The problem of solving a second-order difference equation with initial conditions of this form is called a two-point boundary value problem.)

One way to solve is to fix k_0 and solve the second-order difference equation as a function of k_1, then optimise w.r.t. k_1 at the end. This is called the 'shooting algorithm': you compute the trajectory of the shell for each possible angle, then choose the angle such that the final position of the shell is where you want it to land.

The second initial condition in the finite-horizon case is obviously equivalent to

$$\beta^{T+1} F_1(k_{T+1}, k_{T+2}) k_{T+1} = 0,$$

where we're being loose by using k_{T+2}, an undefined object. The limit as $T \to \infty$ of this expression turns out to be the correct infinite-horizon analogue of the finite-horizon boundary condition. We do *not* require k_{T+1} to converge to zero in the infinite-horizon case.

In the neoclassical growth model, the term $\beta^T F_1(k_T, k_{T+1})$ turns out to be the marginal utility value of the capital stock at T from the perspective of period 0. To see this, write $v(k_0)$ for the maximised value of the sequence problem at initial capital stock k_0:

$$v(k_0) := \max_{\{k_{t+1}\}_{t=0}^{\infty} \in \Pi(k_0)} \sum_{t=0}^{\infty} \beta^t u\left(f(k_t) - k_{t+1}\right)$$

where $\Pi(k_0)$ is the set of feasible sequences from k_0. Then by an envelope theorem,[5]

$$v'(k_t) = u'(f(k_t) - k_{t+1}) f'(k_t) = F_1(k_t, k_{t+1})$$

as claimed.

---

[5] The appropriate envelope theorem is the Benveniste–Scheinkman theorem, which will be introduced later.

2.3 Characterising solutions using the variational approach

The shooting algorithm described for the finite-horizon problem works equally well for the infinite-horizon sequence problem, provided the hypotheses of Theorem 4.15 hold. For each $x_1 \in X$, recursively compute the solution to the Euler equation with x_0 fixed. Then compute

$$\lim_{T \to \infty} \beta^T F_1(x_T, x_{T+1}) \cdot x_T$$

for each x_1. We seek solutions for which this expression is zero, so that the transversality condition is respected.

    If we do not have functional-form assumptions, we obviously can't explicitly compute the optimal sequence. But we still be able to obtain results with weaker assumptions on the primitives. Achieving this using the shooting algorithm is not easy, however. It is difficult to characterise the solution to a second-order difference equation with these sorts of initial conditions (in particular, the transversality condition is hard to work with). It will turn out that it is much easier to derive properties of optimal sequences without functional-form assumptions if we translate the sequence problem into functional space.

2.4 Moving into functional space

As in the previous section, let $v : \mathbf{R}_+ \to \mathbf{R}$ be given by

$$v(x_0) := \max_{\{x_{t+1}\}_{t=0}^\infty \in \Pi(x_0)} \sum_{t=0}^\infty \beta^t F(x_t, x_{t+1}),$$

where x_0 is given and $\Pi : X \to X^\infty$ is the set of feasible sequences, i.e. $\{x_{t+1}\}$ with $x_{t+1} \in \Gamma(x_t)$ for every $t \in \mathbf{Z}_+$. Notice that we are assuming that there is a maximiser of the objective function here, instead of using the supremum. The cautious approach starting with the supremum is pursued in section 4.1 of Stokey et al. (1989).

    Observe that we can break the optimisation problem into sub-problems:

$$
\begin{aligned}
v(x_0) &= \max_{x_1 \in \Gamma(x_0)} \max_{\{x_{t+1}\}_{t=1}^\infty \in \Pi(x_1)} \left[F(x_0, x_1) + \sum_{t=1}^\infty \beta^t F(x_t, x_{t+1}) \right] \\
&= \max_{x_1 \in \Gamma(x_0)} \left[F(x_0, x_1) + \max_{\{x_{t+1}\}_{t=1}^\infty \in \Pi(x_1)} \sum_{t=1}^\infty \beta^t F(x_t, x_{t+1}) \right] \\
&= \max_{x_1 \in \Gamma(x_0)} \left[F(x_0, x_1) + \beta \max_{\{x_{t+2}\}_{t=0}^\infty \in \Pi(x_1)} \sum_{t=0}^\infty \beta^t F(x_{t+1}, x_{t+2}) \right].
\end{aligned}
$$

Also notice that the optimisation problem is stationary in the sense that the problem starting at time t with state x_t is the same as the problem starting at time s with state x_s iff $x_t = x_s$: the problem does not depend on the date. Hence

$$v(x_0) = \max_{x_1 \in \Gamma(x_0)} \left[F(x_0, x_1) + \beta v(x_1) \right],$$

or more generally

$$v(x) = \max_{x' \in \Gamma(x)} \left[F\left(x, x'\right) + \beta v\left(x'\right) \right]. \tag{3}$$

So we've shown that when there is a sequence that attains the optimum in the sequence problem for every initial state $x_0 \in X$, so that the function v is well-defined, v solves the functional equation (3).

The dynamic programming approach to the sequential problem begins by finding a solution to the functional equation (3), i.e. a function v that satisfies (3) at every $x \in X$. (Under regularity conditions, the value function is the only solution to (3) that satisfies certain properties, so we can be sure that we've found the value function.) From v we find the optimal policy correspondence $G : X \Rightarrow X$ given by

$$
\begin{aligned}
G(x) &:= \arg\max_{x' \in \Gamma(x)} \left\{ F\left(x, x'\right) + \beta v\left(x'\right) \right\} \\
&= \left\{ x' \in \Gamma(x) : v(x') = v(x) \right\} \tag{4}
\end{aligned}
$$

Finally, compute the optimal sequences as the sequences $\{x_{t+1}\}_{t=0}^{\infty}$ in X that satisfy $x_{t+1} \in G(x_t)$ for every $t \in \mathbf{Z}_+$. When G is singleton-valued, the optimal sequence is unique.

2.5 The formal equivalence between the SP and FE

A more formal statement of the connections between the sequence problem and the functional equation is given here. The treatment is cursory and proofs are omitted; see section 4.1 of Stokey et al. (1989) for details.

In this section, we do not take a stand on whether suprema are attained. To this end, we define the sequence problem as

$$
\sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1})
$$
$$
\text{s.t.} \quad x_{t+1} \in \Gamma(x_t) \quad \forall t \in \mathbf{Z}_+. \tag{5}
$$

We similarly define the functional equation without requiring that the maximum be attained:

$$v(x) = \sup_{x' \in \Gamma(x)} \left[F\left(x, x'\right) + \beta v\left(x'\right) \right] \quad \text{for each } x \in X. \tag{6}$$

Assumption 7 (SLP Ass'n 4.1). Γ is nonempty-valued.

Assumption 8 (SLP Ass'n 4.2). $\lim_{T \to \infty} \sum_{t=0}^{T} \beta^t F(x_t, x_{t+1})$ exists for each $x_0 \in X$ and $\{x_{t+1}\}_{t=0}^{\infty} \in \Pi(x_0)$.

We first define the connection between the value of the sequence problem and solutions to the functional equation.

Theorem 2 (SLP Th'm 4.2). Suppose Assumptions 4.1 and 4.2 hold. If v is the value of the sequence problem (5), then v satisfies the functional equation (6).

Theorem 3 (SLP Th'm 4.3). Suppose Assumptions 4.1 and 4.2 hold. If v satisfies the functional equation (6) and $\lim_{t \to \infty} \beta^t v(x_t) = 0$ for every feasible sequence $\{x_{t+1}\}_{t=0}^{\infty} \in \Pi(x_0)$, then v is the value of the sequence problem (5).

Theorem 4.2 establishes that the value function is one solution to the functional equation. Theorem 4.3 shows that the value function is the only solution to the functional equation that satisfies a certain boundedness condition. This means that we can obtain the value function by finding solutions to the functional equation, then checking that they satisfy the boundedness condition.

The next task is to relate the sequences that attain the supremum in the sequence problem and those that attain the supremum in the functional equation. Write v for the value of the sequence problem. We say that a sequence $\{x_{t+1}\}_{t=0}^{\infty}$ in X attains the supremum in the sequence problem from x_0 iff

$$\sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) = v(x_0).$$

Analogously, a sequence $\{x_{t+1}\}_{t=0}^{\infty}$ in X attains the supremum in the functional equation from x_0 iff

$$v(x_t) = F\left(x_t, x_{t+1}\right) + \beta v\left(x_{t+1}\right) \quad \text{for each } t \in \mathbf{Z}_+.$$

Of course our interest will be in *feasible* sequences that attain these suprema, i.e. sequences lying in $\Pi(x_0)$.

Theorem 4 (SLP Th'm 4.4)**.** Suppose Assumptions 4.1 and 4.2 hold. If $\{x_{t+1}\}_{t=0}^{\infty} \in \Pi(x_0)$ attains the supremum in the sequence problem (5) from x_0, then it attains the supremum in the functional equation.

Theorem 5 (SLP Th'm 4.5)**.** Suppose Assumptions 4.1 and 4.2 hold. If $\{x_{t+1}\}_{t=0}^{\infty} \in \Pi(x_0)$ attains the supremum in the functional equation and $\limsup_{t \to \infty} \beta^t v(x_{t+1}) \leq 0$, then it attains the supremum in the sequence problem (5) from x_0.

These theorems tell us that we can find v via the functional equation rather than by attacking the sequence problem directly. But how does one solve a functional equation? In this case, we can achieve this using the contraction mapping theorem. Let $C(X)$ be the space of bounded and continuous functions $f : X \to \mathbf{R}$ endowed with the sup metric, i.e.

$$\rho(f, g) := \sup_{x \in X} |f(x) - g(x)| \quad \text{for any } f, g \in C(X).$$

The contraction mapping theorem allows us to establish that the functional equation has a unique solution v that lives in $C(X)$. Moreover, it gives us a simple algorithm that allows us to compute v by starting with an initial guess $v_0 \in C(X)$ and applying a recursive procedure.

3 Mon 28 Sep 2015

3.1 Outline of today

We begin with a recap of how the functional equation can be used to solve the sequence problem. We then study how the value function v can be found, illustrating with a parametrised variant of the neoclassical growth model. Finally we derive properties v under assumptions on the primitives. In particular, v is strictly increasing under Assumptions 4.2, 4.3, 4.5 and 4.6, and strictly concave under Assumptions 4.2, 4.3, 4.8 and 4.7.

3.2 Recap: solving a dynamic program via the value function

The functional-equation approach to the sequence problem delivers a solution in three steps. For simplicity, we consider the case where the policy correspondence $G : X \Rightarrow X$ is single-valued, in which case it can be written as a function $g : X \to X$.

(1) Find the value function v as a solution to the functional equation

$$v(x) = \max_{x' \in \Gamma(x)} \left[F(x, x') + \beta v(x') \right].$$

(Note that there may be solutions besides the value function!)

(2) Compute the policy function g:

$$g(x) = \arg\max_{x' \in \Gamma(x)} \left[F(x, x') + \beta v(x') \right].$$

Under differentiability and interiority assumptions, g is a solution to the first-order condition

$$F_2(x, g(x)) + \beta v'(g(x)) = 0 \quad \forall x \in X.$$

(3) Compute the optimal sequence as the solution to the first-order difference equation $x_{t+1} = g(x_t)$ for $t \in \mathbf{Z}_+$ with initial condition x_0.

Since the optimal sequence follows a first-order difference equation with transition function g, we can study its properties by studying g. When $X \subseteq \mathbf{R}$, we can do this by plotting g and the 45-degree line and analysing the optimal dynamics graphically. If $X \subseteq \mathbf{R}^l$ we can linearise g around a steady state to obtain approximate optimal dynamics local to that steady state. (The latter is not really very interesting in its own right, but it does allow us to examine local stability.)

3.3 How to find v

The functional equation is

$$v(x) = \max_{x' \in \Gamma(x)} \left[F(x, x') + \beta v(x') \right] \quad \text{for every } x \in X.$$

One function that satisfies this equation is the value function.

Let the space $C(X)$ be the set of continuous and bounded functions $f : X \to \mathbf{R}$ endowed with the sup norm, i.e. $\|\cdot\|$ is defined by

$$\|f\| := \sup_{x \in X} |f(x)| \quad \text{for each } f \in C(X).$$

We restrict attention to $C(X)$ because (as it will turn out) the value function lies in this space, and moreover the value function is the *only* solution to the functional equation that lies in this space. So we can be sure of finding a solution in $C(X)$ to the functional equation, and once we get it, we can be sure that this solution is the value function. A second perk of working in $C(X)$ is that continuous and bounded functions are nice when you can get them!

Now define the operator T on $C(X)$ by

$$(Tw)(x) := \max_{x' \in \Gamma(x)} \left[F(x, x') + \beta w(x') \right] \quad \forall (x, w) \in X \times C(X) \qquad (7)$$

(We will ensure that T is well-defined by imposing that Γ is nonempty- and compact-valued and that F is continuous.) Provided that T maps $C(X)$ into itself (i.e. provided that T preserves boundedness and continuity), the problem of finding v is equivalent to finding a fixed point of T. The following assumptions will turn out to be sufficient to guarantee that T maps $C(X)$ into itself, that is has a fixed point, and that the fixed point is unique. Recall that A denotes the graph of the constraint correspondence Γ:

$$A := \left\{ (x, y) \in X^2 : y \in \Gamma(x) \right\}.$$

Assumption 9 (SLP Ass'n 4.3). X is a convex subset of $\mathbf{R}^l$, and $\Gamma : X \rightrightarrows X$ is nonempty- and compact-valued and continuous.

Assumption 10 (SLP Ass'n 4.4). $F : A \to \mathbf{R}$ is bounded and continuous, and $\beta \in (0, 1)$.

Theorem 6 (SLP Th'm 4.6). Suppose Assumptions 4.3 and 4.4 hold. Then T maps $C(X)$ into itself. Moreover T has a unique fixed point v, and for any $v_0 \in C(X)$ and all $n \in \mathbf{N}$,

$$\|T^n v_0 - v\| \leq \beta^n \|v_0 - v\|.$$

This is obviously an application of the contraction mapping theorem. The proof consists of showing that T preserves boundedness and continuity, then showing that it satisfies Blackwell's sufficient conditions for a contraction mapping.

The last part of Theorem 6 provides a globally convergent algorithm for finding v. The existence of a globally convergent algorithm is actually fairly rare in computational economics. However, globally convergent algorithms are usually slow, and this one is no exception.

For computational purposes, there are two main algorithms that are used, both are much faster. The projection method approximates v by a polynomial. (The metric used to measure the closeness of approximation is important.) The other is the perturbation method, which will be discussed later.

3.4 Example: the neoclassical growth model

This example is simple and doesn't actually quite fit the assumptions, but hey. Consider the neoclassical growth model with parametric assumptions on u and f:

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln c_t$$
$$\text{s.t.} \quad k_{t+1} \le k_t^{\alpha} - c_t \qquad \qquad \text{for } t \in \mathbf{Z}_+$$
$$c_t, k_{t+1} \ge 0 \qquad \qquad \text{for } t \in \mathbf{Z}_+.$$

As usual, the resource constraint binds because u is strictly increasing, so we can substitute to obtain

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln(k_t^{\alpha} - k_{t+1})$$
$$\text{s.t.} \quad k_{t+1} \in [0, k_t^{\alpha}] \qquad \qquad \text{for } t \in \mathbf{Z}_+.$$

Define $\Pi : \mathbf{R}_+ \to \mathbf{R}_+^{\infty}$ by

$$\Pi(x_0) := \{\{k_{t+1}\}_{t=0}^{\infty} : k_{t+1} \in [0, k_t^{\alpha}] \quad \forall t \in \mathbf{Z}_+\} \quad \text{for each } k_0 \in \mathbf{R}_+.$$

Then the value function $v : \mathbf{R}_+ \to \mathbf{R}$ is defined as

$$v(k_0) := \max_{\{k_{t+1}\}_{t=0}^{\infty} \in \Pi(k_0)} \sum_{t=0}^{\infty} \beta^t \ln(k_t^{\alpha} - k_{t+1}) \quad \text{for each } k_0 \in \mathbf{R}_+.$$

Assumptions 4.1 and 4.2 hold, so v satisfies the functional equation

$$v(k) = \max_{k' \in [0, k^\alpha]} \left[\ln(k^\alpha - k') + \beta v(k') \right] \quad \text{for each } k \in \mathbf{R}_+.$$

Define the mapping T on $C(\mathbf{R}_+)$ by

$$(Tw)(k) := \max_{k' \in [0, k^\alpha]} \left[\ln(k^\alpha - k') + \beta w(k') \right] \quad \text{for each } (k, w) \in \mathbf{R}_+ \times C(\mathbf{R}_+).$$

If we are to use Theorem 4.6 of Stokey et al. (1989) to establish the existence and uniqueness of a fixed point of T, we must verify Assumptions 4.3 and 4.4. $X = \mathbf{R}_+$, and $\Gamma : X \Rightarrow X$ is given by

$$\Gamma(k) := [0, k^\alpha] \quad \text{for all } k \in \mathbf{R}_+,$$

and $F : A \to \mathbf{R}$ is given by

$$F(k, k') := u(f(k) - k') \quad \text{for all } (k, k') \in \mathbf{R}_+^2 \text{ s.t. } k' \in \Gamma(k).$$

While Assumption 4.3 holds, Assumption 4.4 fails because F is not bounded below. Theorem 4.6 is therefore inapplicable.

Nevertheless, it turns out that the conclusions of Theorem 4.6 still obtain in this setting. We will therefore use the inequality

$$\|T^n v_0 - v\| \le \beta^n \|v_0 - v\| \quad \text{for all } v_0 \in C(X),$$

to find the value function v. We could start with any function $v_0 \in C(X)$, but that would give very complicated results. We will instead make a clever guess (i.e. cheat) and let v_0 be log-linear:

$$v_0(k) := a_0 + b_0 \ln k.$$

Applying T to v_0 gives

$$(Tv_0)(k) = \max_{k' \in [0, k^\alpha]} \left[\ln \left(k^\alpha - k' \right) + \beta a_0 + \beta b_0 \ln k' \right] \quad \forall k \in \mathbf{R}_+.$$

The derivative $m(k, k')$ of the maximand w.r.t. k' satisfies

$$\lim_{k' \downarrow 0} m(k, k') = \infty \quad \text{and} \quad \lim_{k' \uparrow k^\alpha} m(k, k') = -\infty,$$

so the argmax $g_0 : \mathbf{R}_+ \to \mathbf{R}_+$ is interior. g_0 is therefore a solution to the first-order condition

$$\frac{1}{k^\alpha - g_0(k)} = \frac{\beta b_0}{g_0(k)} \quad \forall k \in \mathbf{R}_+,$$

which rearranges to

$$g_0(k) = \frac{\beta b_0}{1 + \beta b_0} k^\alpha \quad \forall k \in \mathbf{R}_+.$$

(As a check, observe that $g_0(k) \in [0, k^\alpha]$ at every $k \in \mathbf{R}_+$.)

Substituting back gives, for all $k \in \mathbf{R}_+$,

$$
\begin{aligned}
(Tv_0)(k) &= \ln\left(k^\alpha - \frac{\beta b_0}{1 + \beta b_0} k^\alpha\right) + \beta a_0 + \beta b_0 \ln\left(\frac{\beta b_0}{1 + \beta b_0} k^\alpha\right) \\
&= \ln\left(\frac{(\beta b_0)^{\beta b_0}}{(1 + \beta b_0)^{1 + \beta b_0}}\right) + \beta a_0 + \alpha\left(1 + \beta b_0\right) \ln k \\
&= a_1(a_0, b_0) + b_1(b_0) \ln k.
\end{aligned}
$$

This is again log-linear, so provided there is a solution $(a, b) \in \mathbf{R}^2$ to $b_1(b) = b$ and $a_1(b, a) = a$, the value function v is $v(k) = a + b \ln k$ for $k \geq 0$. There is in fact a solution, with b obviously given by $b = \alpha/(1 - \alpha\beta)$.

The idea of 'iterating on the contraction' to find v from an initial guess v_0 can be illustrated graphically as follows.[6] We study the contracting effect on the slope b_0 of applying the contraction T. For a given initial guess $b_0 \in \mathbf{R}_+$, $b_1(b_0) = \alpha(1 + \beta b_0)$, as drawn in Figure 2. If b_0 is below the fixed point then $b_1(\beta, b_0) > b_0$, and the opposite occurs if b_0 is below the fixed point.

Now that we have v, we can derive the optimal policy function g from the first-order condition:

$$g(k) = \frac{\beta b}{1 + \beta b} k^\alpha = \alpha\beta k^\alpha \quad \forall k \in \mathbf{R}_+.$$

With g in hand, we can analyse the dynamics of the optimal capital stock sequence graphically, since this sequence is the solution to the first-order difference equation $k_{t+1} = g(k_t)$ with initial condition k_0. By plotting $k_{t+1} = g(k_t)$ together with $k_{t+1} = k_t$, we can study the behaviour of k_{t+1} over time, in particular whether it converges toward a steady state $k^\star$ such that $k^\star = g(k^\star)$. This is done in Figure 3.

Finally, let's relate the dynamic programming approach we took here to the variational approach of Theorem 4.15. Since g is always interior in this example, the optimal sequence is always interior, so we'd expect the

---

[6]I remind the reader that Theorem 4.6, the only formal justification we have for this iteration, does not apply in this setting, so we have to take it on faith that this procedure works.
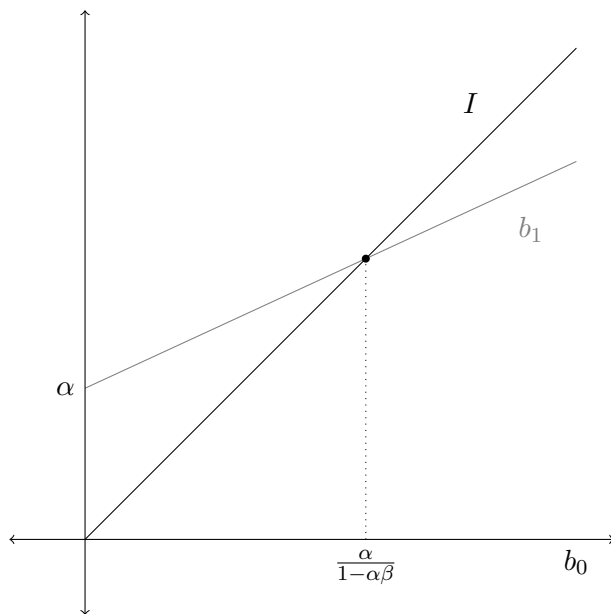
Figure 2 – The identity function I and the function $b_1 : \mathbf{R}_+ \to \mathbf{R}_+$ given by $b_1(b_0) := \alpha(1 + \beta b_0) \ \forall b_0 \in \mathbf{R}_+$, defined as part of the contraction mapping from a log-linear v_0 in the parametric neoclassical growth model. b_1 has absolute slope $\alpha\beta < 1$, so is a contraction mapping on $\mathbf{R}_+$. Since $\mathbf{R}_+$ is complete, b_1 must have a unique fixed point, as indeed it does.
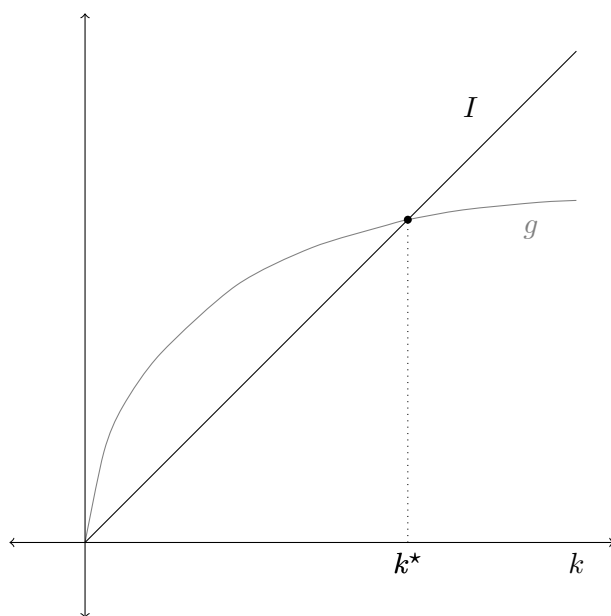
Figure 3 – Dynamics of the solution in the parametrised neoclassical growth model, assuming $k_0 > 0$. Drawn are the policy function g $(g(k) = \alpha\beta k^\alpha)$ and the identity function I. When $k < k^\star$, k increases, and vice versa. So in either case, k converges to $k^\star$, the fixed point of g.

Euler equation to hold. In the unparametrised neoclassical model, the Euler equation is

$$u'(f(k_t) - k_{t+1}) = \beta u'(f(k_{t+1}) - k_{t+2})f'(k_{t+1}) \quad \forall t \in \mathbf{Z}_+,$$

so with our parametric assumptions,

$$\frac{1}{k_t^\alpha - k_{t+1}} = \beta \frac{1}{k_{t+1}^\alpha - k_{t+2}} \cdot \alpha k_{t+1}^{\alpha-1} \quad \forall t \in \mathbf{Z}_+,$$

Rearranging yields

$$1 - (k_{t+2}/k_{t+1}^\alpha) = \alpha\beta \left((k_{t+1}/k_t^\alpha)^{-1} - 1 \right) \quad \forall t \in \mathbf{Z}_+,$$

a first-order difference equation in (k_{t+2}/k_{t+1}^α). At a stationary solution, i.e.

$$k_{t+2}/k_{t+1}^\alpha = k_{t+1}/k_t^\alpha = z,$$

we obtain a polynomial in z with roots 1 and $\alpha\beta$, giving us the two stationary solutions for $\{z_{t+1}\}$, each of which is a first-order difference equation in k_{t+1}:

$$k_{t+1} = k_t^\alpha \quad \text{and} \quad k_{t+1} = \alpha\beta k_t^\alpha.$$

The first solution violates the transversality condition since

$$\lim_{T \to \infty} \beta^T u'(k_T^\alpha - k_{T+1}) \cdot \alpha k_T^{\alpha-1} \cdot k_T = \alpha u'(0) \lim_{T \to \infty} \beta^T k_T^\alpha$$

and $u'(0) = 1/0$ is undefined, and certainly not equal to zero. The second solution satisfies the transversality condition, however:

$$\lim_{T \to \infty} \beta^T u'(k_T^\alpha - k_{T+1}) \cdot \alpha k_T^{\alpha-1} \cdot k_T = \lim_{T \to \infty} \beta^T \frac{1}{(1-\alpha\beta)k_T^\alpha} \cdot \alpha k_T^{\alpha-1} \cdot k_T$$

$$= \frac{1}{1-\alpha\beta} \lim_{T \to \infty} \beta^T = 0.$$

So the solution $k_{t+1} = \alpha\beta k_t^\alpha$ (with initial condition k_0) identified by our functional equation argument is the unique sequence that satisfies the Euler equation, transversality condition and initial condition k_0. Happy times!

3.5   Properties of v

Recall Assumptions 4.2 and 4.3, which allowed us to apply Berge's Theorem and the Contraction Mapping Theorem to the functional equation. These assumptions gave us existence and uniqueness of v, and gave us an algorithm for computing it. But it also gave us properties of v: it is continuous and bounded. Can we say more?

A natural question is whether v is increasing. As you might expect, v is increasing when F is increasing and Γ is also increasing in an appropriate sense.

Assumption 11 (SLP Ass'n 4.5)**.** For each $x \in X$, $F(\cdot, x)$ is strictly increasing.

Assumption 12 (SLP Ass'n 4.6)**.** Γ is monotone in the sense that for $x, x' \in X$, $x \leq x'$ implies $\Gamma(x) \subseteq \Gamma(x')$.

In the context of the neoclassical growth model, Assumption 4.6 says that a capital-rich planner can achieve any next-period capital stock that a capital-poor planner can. The way we've set up the neoclassical growth model, this assumption is satisfied: $k < k'$ implies $[0, f(k)] \subseteq [0, f(k')]$ because f is increasing.

Note, however, that monotonicity in the neoclassical growth model is an artifact of imposing $k_{t+1} \geq 0$ for every $t \in \mathbf{Z}_+$. If instead we imposed $i_{t+1} \geq 0$ for every $t \in \mathbf{Z}_+$, then

$$\Gamma(k) = [(1 - \delta)k, f(k)] \quad \forall k \in \mathbf{R}_+.$$

This correspondence is not monotone! Although our proof that v is strictly increasing will rely on monotonicity of Γ, v is in fact strictly increasing in the neoclassical growth model with $i_{t+1} \geq 0$ despite the failure of monotonicity. (This was a homework exercise.)

Theorem 7 (SLP Th'm 4.7)**.** Under Assumptions 4.3, 4.4, 4.5 and 4.6, v is strictly increasing.

Proof. Take $x' \geq x$ with $x' \neq x$ in X, and let $g : X \to X$ be any selection from the optimal policy correspondence G. Then

$$\begin{aligned}
v(x') &= F(x', g(x')) + \beta v(g(x')) \\
&\geq F(x', g(x)) + \beta v(g(x)) \\
&> F(x, g(x)) + \beta v(g(x)) = v(x)
\end{aligned}$$

where the weak inequality used the optimality of g and fact that $g(x) \in \Gamma(x')$ by monotonicity, and the strict inequality holds since $F(\cdot, y)$ is strictly increasing for each $y \in X$. ∎

Remark 1. If we assume only that $F(\cdot, y)$ is weakly increasing then the same proof establishes that v is weakly increasing.

Next, we wish to find conditions that guarantee that v is strictly concave. Recall that

$$A := \{(x, y) \in X^2 : y \in \Gamma(x)\}$$

is the graph of Γ. Unsurprisingly, this will require a concavity condition on F and a convexity condition on Γ.

Definition 1. $F : A \to \mathbf{R}$ for A convex is strictly concave iff for every $(x, y) \neq (x', y')$ in A and every $\lambda \in (0, 1)$,

$$F(\lambda x + (1 - \lambda)x', \lambda y + (1 - \lambda)y') \geq \lambda F(x, y) + (1 - \lambda)F(x', y').$$

Assumption 13 (SLP Ass'n 4.8). The graph A of Γ is convex.

In the neoclassical growth model, this requires the upper contour f of Γ to be concave, which means that the marginal product of capital is nonincreasing. We require convexity of A so that our concavity assumption on F makes sense:

Assumption 14 (SLP Ass'n 4.7). F is strictly concave.

Theorem 8 (SLP Th'm 4.8). Under Assumptions 4.3, 4.4, 4.8 and 4.7, v is strictly concave.

Proof. Let $C'(X) \subseteq C(X)$ denote all concave, continuous and bounded functions $f : X \to \mathbf{R}$, and let $C''(X) \subseteq C'(X)$ denote all *strictly* concave, continuous and bounded functions $f : X \to \mathbf{R}$. Note that $C'(X)$ is closed. By Corollary 1 in ch. 3 of Stokey et al. (1989), it therefore suffices to show that $T(C'(X)) \subseteq C''(X)$.

To that end, take $w \in C'(X)$, and let g_w be a selection from the argmax correspondence. Fix $x, x' \in X$ and $\lambda \in (0, 1)$. Then

$$
\begin{aligned}
&(Tw)(\lambda x + (1 - \lambda)x') \\
&\quad = F[\lambda x + (1 - \lambda)x', g_w(\lambda x + (1 - \lambda)x')] + \beta w[g_w(\lambda x + (1 - \lambda)x')] \\
&\quad \geq F[\lambda(x, g_w(x)) + (1 - \lambda)(x', g_w(x'))] + \beta w[\lambda g_w(x) + (1 - \lambda)g_w(x')] \\
&\quad > \lambda[F(x, g_w(x)) + \beta w(g_w(x))] + (1 - \lambda)[F(x', g_w(x')) + \beta w(g_w(x'))] \\
&\quad = \lambda(Tw)(x) + (1 - \lambda)(Tw)(x')
\end{aligned}
$$

where the weak inequality used the optimality of g_w and the convexity of A (so that $\lambda g_w(x) + (1 - \lambda) g_w(x')$ is feasible at $\lambda x + (1 - \lambda)x'$), and the strict inequality used the strict convexity of F. So $Tw \in C''(X)$ whenever $w \in C'(X)$, i.e. $T(C'(X)) \subseteq C''(X)$. ∎

4 Wed 30 Sep 2015

4.1 Outline of today

Today we'll provide sufficient conditions for the differentiability of the value function v (Theorems 4.10 and 4.11 in Stokey et al. (1989)).

We'll then apply all of our technical machinery to study the properties of optimal paths $\{x_{t+1}\}_{t=0}^{\infty}$ defined by $x_{t+1} = g(x_t)$ where g is an optimal policy function. It turns out that we can't say much: any smooth transition function g is the optimal policy function of some well-behaved Stokey–Lucas dynamic program. This anything-goes theorem is called the Boldrin-Montrucchio theorem.

We next turn to the neoclassical growth model to show that all is not lost. With the additional structure present there, we can characterise the optimal path of the capital stock quite sharply. In particular, it displays montone convergence to the unique positive steady-state level from any positive initial level.

4.2 Differentiability of v

Assumption 15 (SLP Ass'n 4.9)**.** For each $y \in X$, $F(\cdot, y)$ is continuously differentiable on int X.

Results establishing differentiability and formulae for derivatives of value functions with respect to parameters in parametrised optimisation problems are called envelope theorems. A fairly general envelope theorem that suits our purposes is the following.

Theorem 9 (Benveniste and Scheinkman (1979))**.** Let $X \subseteq \mathbf{R}^l$ be convex, and let $V : X \to \mathbf{R}$ be concave. Take $x_0 \in$ int X, and let D be an open neighbourhood of x_0. Suppose there is a concave and differentiable function $W : D \to \mathbf{R}$ that satisfies $W(x_0) = V(x_0)$ and $W(x) \leq V(x)$ for all $x \in D$. Then V is differentiable at x_0, and $V'(x_0) = W'(x_0)$.[7]

It looks from the proof in Stokey et al. (1989, p. 84) like we actually only need differentiability of W at x_0, not differentiability in the whole neighbourhood D. It's easy enough to convince yourself that this theorem is true by drawing pictures of attempted counterexamples.

Theorem 10 (SLP Th'm 4.11)**.** Suppose Assumptions 4.3, 4.4, 4.8, 4.7 and 4.9 hold, and take $x_0 \in$ int X such that $g(x_0) \in$ int $\Gamma(x_0)$. Then v is

---

[7]Theorem 4.10 in Stokey et al. (1989).

differentiable at x_0, and its derivative satisfies the envelope condition

$$v'(x_0) = F_1(x_0, g(x_0)).$$

Note that by Theorem 4.8, G is singleton-valued here, so g is unique. Also note that in Stokey et al. (1989), Theorem 4.11 states that v is *continuously* differentiable! Not sure what to make of this.

We might be unhappy with the assumption that $g(x_0) \in \operatorname{int} \Gamma(x_0)$, which is not directly on primitives. In most economic environments, we make assumptions to avoid boundaries, and then this is unproblematic. But note well that v need not be differentiable at all $x \in \operatorname{int} X$: we also require $g(x) \in \operatorname{int} \Gamma(x)$.

Proof. The functional equation gives us

$$v(x_0) = F(x_0, g(x_0)) + \beta v(g(x_0)).$$

Since x_0 and $g(x_0)$ are interior, $(x_0, g(x_0))$ is interior to A, so there is a neighbourhood D of x_0 such that $g(x_0) \in \Gamma(d)$ for every $d \in D$. (Larry also used the fact that Γ is continuous to establish this, but I'm not sure why that's needed.)

Define $W : D \to \mathbf{R}$ by

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

W is differentiable and concave since $F(\cdot, g(x_0))$ is, and satisfies $W(x_0) = v(x_0)$ and $W(x) \leq v(x)$ for every $x \in D$ by optimality of $g(x_0)$ at x_0. Hence by the Benveniste–Scheinkman theorem, v is differentiable at x_0 with derivative

$$v'(x_0) = W'(x_0) = F_1(x_0, g(x_0)). \qquad \blacksquare$$

We have not showed that g is differentiable! But to build intuition, suppose that g is differentiable. Then we can easily derive the envelope condition in Theorem 4.11 as follows: take $(x, g(x)) \in \operatorname{int} A$ and differentiate the functional equation using the chain rule:

$$v'(x) = F_1(x, g(x)) + \left[F_2(x, g(x)) + \beta v'(g(x)) \right] g'(x) = F_1(x, g(x))$$

since $g'(x) = 0$ by the first-order condition. Proving differentiability of g is harder, but it has been done: see Santos (1991).

4.3 The Boldrin-Montrucchio theorem

Taking stock: under the maintaned assumptions, we have shown that v is strictly increasing, strictly concave and differentiable, and that g is unique (G single-valued) and continuous. Can we say more about g, and by extension the optimal sequences?

Theorem 11 (Boldrin and Montrucchio (1986))**.** Suppose that $X \subseteq \mathbf{R}$ is compact and that $g : X \to X$ is twice differentiable. Then there is a Stokey–Lucas canonical program (X, Γ, F, β) satisfying Assumptions 4.3, 4.4, 4.8, 4.7 and 4.9 whose optimal policy function is g.[8]

Example 1. Take $X = [0, 1]$ and consider the 'tent map' $g(x) = 4x(1 - x)$. (If you draw it, the origin of the name should be clear.) 3/4 a steady state, but the dynamics are cyclical elsewhere! The Boldrin-Montrucchio theorem says that even this function is the policy function of some Stokey–Lucas dynamic program satisfying most of our assumptions.

4.4 Back to the neoclassical growth model

Although we can say little about the dynamics of optimal solutions to the general Stokey–Lucas dynamical program, all is not lost. It turns out that with more structure in place, we can sometimes obtain sharp characterisations of g, and by extension of the optimal sequence. The neoclassical growth model is one such case.

Recall that we take $X := \max\{k_0, \bar{k}\}$ where $\bar{k}$ is the unique strictly positive solution to $k = f(k)$. $\Gamma : X \rightrightarrows X$ is given by $\Gamma(x) := [0, f(k)]$ for each $k \in X$. Denoting the graph of Γ by A, $F : A \to \mathbf{R}$ is defined by $F(k, k') := u(f(k) - k')$. We also assume that $\beta \in (0, 1)$. The primitive functions u and f satisfy the restrictions imposed previously.

$X \subseteq \mathbf{R}$ is convex, and Γ is nonempty- and compact-valued as well as continuous (since f is continuous), so Assumption 4.3 holds. F is continuous since f and u are, and bounded since A is compact (since X is compact and u and f are continuous), and $\beta \in (0, 1)$, so Assumption 4.4 holds. $F(\cdot, k')$ is strictly increasing for every feasible k' since u and f are stricty increasing, so Assumption 4.5 holds. Since f is increasing, $k' \geq k$ implies $\Gamma(k') = [0, f(k')] \supseteq [0, f(k)] = \Gamma(k)$, so Assumption 4.6 holds. A is convex

---

[8]Theorem 6.1 in Stokey et al. (1989). In fact, we can take $\Gamma(x) = X \ \forall x \in X$ for this result.

since f is concave, so Assumption 4.8 holds. For any $k' > k$ and $\lambda \in (0, 1)$,

$$
\begin{aligned}
F(\lambda(k, y) + (1 - \lambda)(k', y')) &= u\left[f(\lambda k + (1 - \lambda)k') - (\lambda y + (1 - \lambda)y')\right] \\
&> u\left[\lambda(f(k) - y) + (1 - \lambda)(f(k') - y')\right] \\
&> \lambda u(f(k) - y) + (1 - \lambda)u(f(k') - y') \\
&= \lambda F(k, y) + (1 - \lambda)F(k', y'),
\end{aligned}
$$

where the first inequality holds since f is strictly concave and u strictly increasing, and the second inequality holds because u is strictly concave. So F is strictly concave, and thus satisfies Assumption 4.7. Finally, F is continuously differentiable on int A since u and f are continuously differentiable, so Assumption 4.9 holds.

Since Assumptions 4.3 through 4.9 hold, it follows that the value function v of this program is bounded, (continuous,) strictly increasing, strictly concave, and differentiable. Moreover the optimal policy is uniquely given by a continuous function g, and the derivative of the value function is $v'(k) = F(k, g(k))\ \forall k \in X$. Finally, v is the unique fixed point of the contraction mapping $T : C(X) \to C(X)$, giving us a way of computing it numerically from an initial guess.

With housekeeping done, we want to go further in characterising g than was possible in the general Stokey–Lucas model in which the Boldrin-Montrucchio theorem holds.

Proposition 3. In the neoclassical growth model, the following hold.

(1) $g(0) = 0$.

(2) For $k > 0$, $g(k) \in \text{int}\,\Gamma(k)$.

(3) g is strictly increasing.

(4) There is a unique $k^\star > 0$ that satisfies $g(k^\star) = k^\star$ (and $\beta f'(k^\star) = 1$).

(5) For $k < k^\star$, $g(k) \in (k, k^\star)$; for $k > k^\star$, $g(k) \in (k^\star, k)$.

(6) For any $k_0 > 0$, the sequence $\{k_{t+1}\}$ defined by $k_{t+1} = g(k_t)\ \forall t \in \mathbf{Z}_+$ converges to $k^\star$.

Proof. (1): $\Gamma(0) = \{0\}$, so $g(0) = 0$.

(2): Take $k > 0$. The derivative of the maximand in the functional equation is

$$
m(k, k') = F_2(k, k') + \beta v'(k') = -u'(f(k) - k') + \beta v'(k').
$$

38

By an Inada condition on u, $\lim_{k' \uparrow f(k)} m(k, k') = -\infty$, so $g(k) < f(k)$. By the envelope condition,

$$v'(k') = F_1(k', g(k')) = u'(f(k') - g(k'))f'(k').$$

By two Inada conditions on f and u, $\lim_{k' \downarrow 0} F_1(k', g(k')) = \infty \times \infty$. Hence $\lim_{k' \downarrow 0} m(k, k') = \infty$, so $g(k) > 0$.[9]

(3): Since g is interior by (2), it solves the first-order condition

$$u'(f(k) - g(k)) = \beta v'(g(k)).$$

u and v are strictly concave, so u' and v' are strictly decreasing, and f is strictly increasing. When k increases, $f(k)$ increases. $g(k)$ must then increase to preserve the equality, both by increasing the LHS and by lowering the RHS.[10]

(4): This is a little more involved. We first derive a necessary condition for $k^\star > 0$ to be stationary, then verify this condition is also sufficient. By interiority of g we have the first-order and envelope conditions

$$u'(f(k) - g(k)) = \beta v'(g(k))$$
$$u'(f(k) - g(k))f'(k) = v'(k).$$

Combining, we obtain

$$u'(f(k) - g(k)) = \beta u'(f(g(k)) - g^2(k))f'(g(k))$$

(the Euler equation). Now suppose that $k^\star > 0$ is such that $g(k^\star) = k^\star$. Then since u is strictly increasing, so that u' is strictly positive, we obtain $1 = \beta f'(k^\star)$, or $k^\star = (f')^{-1}(1/\beta)$. This is a necessary condition for $k^\star > 0$ to be a stationary point. With our conditions on f (strict concavity, continuous differentiability), there is exactly one $k^\star > 0$ that satisfies this necessary condition.

Now we show that the unique $k^\star > 0$ that satisfies $1 = \beta f'(k^\star)$ is in fact a stationary point. Observe that by strict concavity of v, for any $k > 0$ we have

$$[v'(k) - v'(g(k))][k - g(k)] \leq 0 \quad \text{with equality iff } k = g(k).$$

---

[9]This argument can be illustrated by a picture plotting $F_2(k, \cdot)$ against $\beta v'(\cdot)$. The one is the marginal cost today of saving, the other is the marginal benefit starting tomorrow. The intersection occurs at $g(k)$.

[10]Again we can draw a picture here: $F_2(k, \cdot)$ is strictly increasing and $\beta v'(\cdot)$ is strictly decreasing, and increasing k shifts the former down. Hence the intersection shifts to the right.
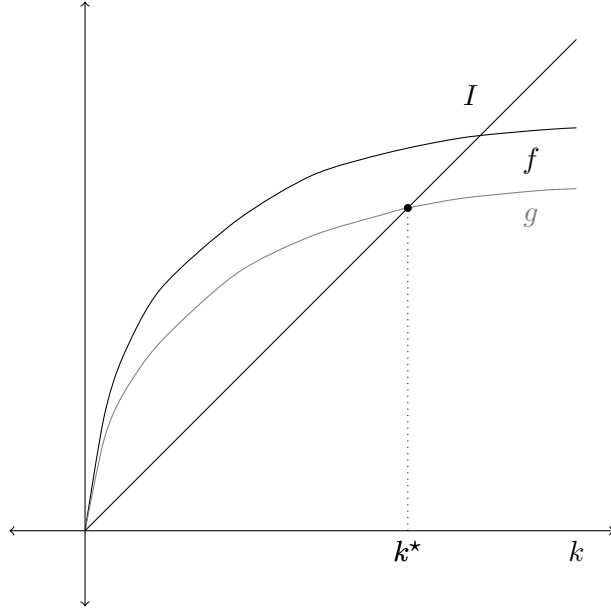
Figure 4 – Dynamics in the neoclassical growth model. The optimal policy function g has the depicted shape by Proposition 3. While g happens to be concave in this picture, that is not guaranteed by Proposition 3.

Using the first-order and envelope conditions (respectively) to substitute,

$$[u'(f(k) - g(k))f'(k) - \beta^{-1}u'(f(k) - g(k))][k - g(k)] \leq 0$$

$$\text{with equality iff } k = g(k).$$

Since u is strictly increasing, we can cancel u' to obtain

$$[\beta f'(k) - 1][k - g(k)] \leq 0 \quad \text{with equality iff } k = g(k). \tag{8}$$

So we've established that for $k^\star > 0$, $\beta f'(k^\star) = 1$ is sufficient for $k^\star = g(k^\star)$.

(5): Consider again the inequality (8). Take $k \in (0, k^\star)$. Then $\beta f'(k) - 1 > 0$ since f is strictly concave, so $g(k) > k$. Moreover $g(k) < g(k^\star) = k^\star$ since g is strictly increasing by part (3). A similar argument for $k > k^\star$ delivers $g(k) \in (k^\star, k)$.

(6): Since $\{k_{t+1}\}$ is monotone and bounded, it converges to some $k_{\lim}$. Hence $\{k_{t+1}\}$ is Cauchy, so $|g(k_{t+1}) - g(k_t)| \to 0$, which yields $g(k_{\lim}) = k_{\lim}$ by continuity of g. By part (4), it must be that $k_{\lim} = k^\star$. ∎

Figure 4 illustrates these properties. By (1), g goes through the origin. By (2), g is strictly between f and 0 except at the origin. By (3) it's strictly

increasing. By (4) it cuts the diagonal exactly once except at the origin. By (5), it cuts the diagonal from above.

Final remarks. g is also differentiable in the neoclassical growth model. Larry does not know conditions under which it is concave, however (so don't let the pictures of a concave g lead you astray). Although we've proved monotone convergence to $k^\star$, we have not said anything about the speed of convergence, an important quantitative question.

5 Mon 5 Oct 2015

5.1 Outline of today

We've been trying to characterise the optimal sequences $\{x_{t+1}\}_{t=0}^{\infty}$ induced by the optimal policy function g and initial condition x_0 under the assumptions that imply strict monotonicity, strict concavity and differentiability of v. The Boldrin-Montrucchio theorem showed us that we can say almost nothing without making more specific assumptions. We then looked at the neoclassical growth model (where we have additional assumptions), and gave a sharp characterisation of the optimal paths.

The finding that we could obtain a sharp characterisation of g with some additional assumptions may appear to have taken the sting out of Boldrin-Montrucchio. With just a few plausible assumptions, the anything-goes world is replaced with a list of strong properties for g! This would be too optimistic, however: it turns out that small adjustments to the assumptions of the neoclassical growth model give wildly different g. We will illustrate this with a simple two-sector model in which we can also sharply characterise g. This model also offers an opportunity to discuss the price of capital.

After today, we'll move on to market equilibria. We'll discuss constant returns to scale, the factor price frontier, and the indeterminacy of firm size under constant returns. Afterwards, we'll study three distinct concepts of equilibrium in this kind of model. This material is in ch. 6.1 in Stokey et al. (1989). Afterwards we'll move on to growth theory.

5.2 A two-sector growth model

There are consumption goods $c_t \geq 0$ and capital goods $k_t \geq 0$. Consumption goods are produced according to $c_t = k_t^{\theta} n_t^{1-\theta}$, where $n_t \in [0, 1]$ is labour employed for consumption goods production. Labour is supplied inelastically, and the other $1 - n_t$ workers are producing capital goods according to $k_{t+1} = 1 - n_t$. (Think of c_t as bread and k_{t+1} as ovens.) Capital depreciates fully in each period. Preferences are described by $u(c) = c^{\alpha}/\alpha$. We take θ and α to lie in $(0, 1)$.

We can write this in canonical form as

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(k_t, k_{t+1})$$

$$\text{s.t.} \quad k_{t+1} \in \Gamma(k_t) \quad \text{for } t \in \mathbf{Z}_+.$$

where $\Gamma(k) := [0, 1]$ and

$$F(k, k') := \left[k^\theta (1 - k')^{1-\theta} \right]^\alpha \Big/ \alpha.$$

Assumptions 4.1 and 4.2 hold, so the functional equation and sequence problem are equivalent in the appropriate sense. The value function v satisfies the functional equation

$$v(k) = \max_{k' \in [0,1]} \left[\left(k^\theta (1 - k')^{1-\theta} \right)^\alpha + \beta v(k') \right],$$

and we define the optimal policy function g by

$$g(k) = \arg\max_{k' \in [0,1]} \left[\left(k^\theta (1 - k')^{1-\theta} \right)^\alpha + \beta v(k') \right].$$

When $k = 0$, consumption will be zero no matter how much labour you pour into producing it, so clearly $g(k) = 1$. This is the opposite property to the neoclassical growth model, where $k = 0$ is an absorbing state.

From the constraint, consumption at (k, k') is

$$c(k, k') := k^\theta (1 - k')^{1-\theta},$$

with derivative

$$c_2(k, k') = -(1 - \theta)(k/(1 - k'))^\theta.$$

$c_2(k, k')$ is the marginal utility cost of next-period capital at (k, k'). Notice that the marginal cost of capital is increasing in k', whereas in the neoclassical growth model the marginal cost of capital is constant and equal to 1. (This is because the technology in that model is such that the same technology f produces both consumption goods and capital goods.) Also notice that marginal cost of capital is increasing in k. This might be thought a strange property: it is more costly to produce capital when the capital stock is large already.

Let's assume that g is interior. Then the first-order condition is

$$-F_2(k, g(k)) = \beta v'(g(k)),$$

where

$$F_2(k, g(k)) = \alpha \left(k^\theta (1 - g(k))^{1-\theta} \right)^{\alpha-1} \times -(1 - \theta)(k/(1 - g(k)))^\theta$$

$$= -\alpha \left(k^\theta (1 - g(k))^{1-\theta} \right)^{\alpha-1} c_2(k, g(k)).$$

43

The marginal cost of capital plays the same role in the first-order conditions as in the neoclassical growth model.

When k increases, the RHS of the first-order condition remains unchanged while the LHS shifts. The term $\alpha\big(k^{\theta}(1-g(k))^{1-\theta}\big)^{\alpha-1}$ decreases, while $c_2(k, k')$ increases. Apparently $\alpha > 0$ guarantees that the net effect is for the LHS to increase. This means that the equality can be preserved only if $g(k)$ decreases; so g is strictly decreasing. This is again the opposite of what happens in the neoclassical growth model.

There is a $k^{\star} \in (0, 1)$ such that $g(k^{\star}) = k^{\star}$. But there are parameter values here such that g has slope steeper than -1 at $k^{\star}$. This means that the steady state is locally unstable. Moreover, neither 0 nor 1 is a steady state. So when $k_0 \neq k^{\star}$, the solution does not display convergence to some steady state, unlike in the neoclassical growth model. Instead the linearised dynamics local to the steady state 'spiral outwards'. The intuition is that when capital is scarce you build lots, but then you have lots it's expensive, so you build little.

When the slope of g at $k^{\star}$ is steeper than -1, there is a stationary cycle: $k_1 \neq k_2$, $k_2 = g(k_1)$ and $k_1 = g(k_2)$. It can be shown that this stationary cycle is a local attractor. Larry conjectures that from any $k_0 \neq k^{\star}$, the system converges to this cycle. But we don't actually know: the linearised dynamics around the stationary point and the stationary cycle tell us nothing about behaviour in other parts of the state space.

This model demonstrates the Boldrin-Montrucchio theorem very starkly. With somewhat different assumptions, we obtain the opposite properties to those derived in the neoclassical growth model!

5.3   The speed of convergence

In the neoclassical growth model, we never reach $k^{\star}$ unless we start there. But a more interesting question is how many periods it takes to get within ε of $k^{\star}$ for any fixed $\varepsilon > 0$. (It would be a mistake to interpret the fact that $k^{\star}$ is never reached too literally.) In practice the question usually asked is how long it takes until $(k_t - k^{\star})/(k_0 - k^{\star}) \leq \varepsilon$, often for $\varepsilon = 0.05$.

Since $k_t = g^t(k_0)$, this is a question about the g function. To be able to answer this question, the obvious approach is to parametrise. We need to parametrise f so that we can solve $\beta f'(k^{\star}) = 1$ for $k^{\star}$, and we also need to parametrise u so that we can solve for g. In general we won't have a closed form for g even then, but we can use the algorithm in Theorem 4.6 (or a better one) to get a sequence of functions that converges uniformly to g.

Another method would be to use the Euler equation and the shooting

algorithm to obtain g. Larry thinks this doesn't give you good economic intuition. And it isn't easily implementable without tight parametric structure anyway.

We will instead use a perturbation method. Assume that g is n times differentiable (we have *not* proved this, but the method turns out to work here). We use the implicit function theorem to develop a Taylor approximation to g around $k^\star$. We'll illustrate this first in a trivial example, then in the neoclassical model.

Example 2. Let x be exogenous, and let y be endogenous, determined by $h(x, y) = 0$ here h is as differentiable as required. Assume that there is a solution for y for each x, i.e. there is a g such that $h(x, g(x)) = 0$ for every x. This is in general a hard problem: it's a functional equation in the unknown function g.

One issue is that there are in general multiple solutions. To gain traction, we assume that wherever there are multiple solutions, these are isolated points. For example if h is a circle

$$h(x, y) = -\tfrac{1}{2}\left(x - \tfrac{1}{2}\right)^2 - \tfrac{1}{2}\left(y - \tfrac{1}{2}\right)^2$$

then the solutions are isolated on $(-1/2, 1/2)$.

Assume that there is a $x^\star \in X$ such that I know $g(x^\star)$. (In the neoclassical growth model, use the interior steady state.) Let R be the error function $R(x, f) = h(x, f(x))$. Obviously $R(x, g) = 0$, and hence $R_1(x, g) = 0$. Moreover

$$R_1(x, g) = h_1(x, g(x)) + h_2(x, g(x))g'(x),$$

so we have

$$h_1(x, g(x)) + h_2(x, g(x))g'(x) = 0$$

at every x. Assuming that $h_2(x, g(x)) \neq 0$, rearranging yields

$$g'(x) = -\frac{h_1(x, g(x))}{h_2(x, g(x))}$$

for every x. (You will recognise this as the implicit function theorem.) Note $h_2(x, g(x)) \neq 0$ fails at two points on the circle! This is why we need the restriction that solutions are 'isolated'. In the case where x is a vector, we need the Jacobian of $h(x, \cdot)$ to have full rank, so the restriction is sometimes called a rank restriction.

By the same method, we can get g'' from h and g'. And again, we only need to solve a linear equation to get g''. We can continue to higher orders in the same way.

So under the full rank and differentiability assumptions, you can recursively and very quickly calculate all the derivatives of g at a point $x^\star$, provided you know the value of g at $x^\star$.

We can then approximate g in the vicinity of $x^\star$ by the Taylor expansion

$$g(x) \simeq \sum_{n=0}^{N} \frac{g^{(n)}(x^\star)}{n!}(x - x^\star)^n.$$

But we need to treat this with caution! There are always neighbourhoods in which the Taylor approximation is very poor. (This is called Runge's phenomenon, illustrated well by the Runge function $g(x) = 1/(1 + x^2)$.)

6 Wed 7 Oct 2015

6.1 Outline of today

Last time, we covered the qualitative properties of the neoclassical and two-sector growth models. The solutions were very different! Next on the agenda are quantitative properties, in particular the speed of convergence.

We've discussed three approaches to this kind of question: dynamic programming (i.e. iterating on the contraction in Theorem 4.6), the shooting algorithm (based on Theorem 4.15), and perturbation (Taylor series approximations). There are a lot of other algorithms. Note that perturbation requires g to be infinitely differentiable, or else the Taylor approximation breaks down. (This occurs in e.g. monetary models in which the interest rate cannot go negative in equilibrium.) Some of these methods are designed to deal with participation constraints (which sometimes do, sometimes don't bind on the equilibrium path), e.g. dynamic programming is. Participation constraints generate non-differentiable g, so perturbation won't work in these contexts! Another note: for most functions, Taylor approximations are locally good but globally poor due to Runge's phenomenon. For example, the Taylor approximation to $\ln(x)$ around 1 gets increasingly *bad* at 2 as you increase the order of the expansion!

6.2 Convergence speed in the neoclassical growth model

For the neoclassical growth model, let R map capital $k \in \mathbf{R}_+$ and candidate policy functions $z : \mathbf{R}_+ \to \mathbf{R}_+$ with $z(k) \in [0, f(k)]$ at each $k \in \mathbf{R}_+$ into the extent to which they violate the Euler equation:

$$R(k, z) := u'(f(k) - z(k)) - \beta u' \left(f(z(k)) - z^2(k) \right) f'(z(k)).$$

Since the optimal policy function g is interior-valued at all $k > 0$, we have $R(k, g) = 0$ at every $k > 0$. Hence $R_1(k, g) = 0$ for every $k > 0$. We will assume differentiability of everything, including g.[11] In order to use the perturbation method, we require a $k > 0$ at which $g(k)$ is known. The interior steady state does the trick: $g(k^\star) = k^\star$ where $k^\star := (f')^{-1}(1/\beta)$. Define steady-state consumption by $c^\star := f(k^\star) - k^\star$.

To obtain the (exact) value of the derivative g' at $k^\star$, we differentiate $R(\cdot, g)$ with respect to its first argument using the chain rule, use the previously-established fact that $R_1(k, g) = 0$ at every $k > 0$ (which therefore

---

[11]Christiano (1989) computes g very accurately for a model like this. It turns out to be very smooth! But Larry is not aware of a theorem that shows it to be differentiable.

holds at $k^\star$), and rearrange. To save space, we'll drop the arguments of all functions on the understanding that everything is evaluated at $(k^\star, g)$.

$$
\begin{aligned}
0 = R_1(k^\star, g) &= u'' \cdot (f' - g') - \beta u'' \cdot \left(f'g' - (g')^2 \right) f' - \beta u' f'' g' \\
&= u'' f' - \left[u'' + \beta u'' \cdot (f')^2 + \beta u' f'' \right] g' + \beta u'' f' \cdot (g')^2 . \\
&= \frac{u''}{\beta} - \left[u'' + \frac{u''}{\beta} + u' \frac{f''}{f'} \right] g' + u'' \cdot (g')^2 .
\end{aligned}
$$

where the final equality used $f' = f'(k^\star) = 1/\beta$ (to replace f' with $1/\beta$ in three places, and to replace β with $1/f'$ in one place). Dividing through by $u'' < 0$,

$$
\begin{aligned}
0 &= \frac{1}{\beta} - \left[1 + \frac{1}{\beta} + \frac{u'}{u''} \frac{f''}{f'} \right] g' + (g')^2 \\
&= \beta^{-1} - \phi g' + (g')^2
\end{aligned}
$$

where

$$
\phi := 1 + \beta^{-1} + \frac{u'}{u''} \frac{f''}{f'} = 1 + \beta^{-1} + \left| \frac{u'}{u''} \right| \left| \frac{f''}{f'} \right| .
$$

The final equality holds because $u'/u'' < 0$ and $f''/f' < 0$ since u and f are strictly increasing and strictly concave.

So $g'(k^\star)$ is the solution to a quadratic equation. Rewrite it as

$$
\beta^{-1} \left(g' \right)^{-1} + g = \phi .
$$

The LHS is a convex parabola with minimum at β^{-1}. The value at the minimum is $1 + \beta^{-1} < \phi$, so our quadratic has two real roots λ_1 and λ_2 with $\lambda_1 < \beta^{-1} < \lambda_2$. Since we know that $k^\star$ is locally stable and that $g' > 0$, it must be that $g' < 1 < \beta^{-1}$, so $g = \lambda_1$, the smaller root. It is given in closed form by the quadratic formula as

$$
g'(k^\star) = g' = \frac{\phi - \sqrt{\phi^2 - 4\beta^{-1}}}{2} .
$$

Why do we need to solve a quadratic to obtain $g'(k^\star)$ when it was given by a linear equation in the example of the previous section? The reason is that there, $R(x, z) = h(x, z(x))$, so that

$$
0 = R_1(x, g) = h_1(x, g(x)) + h_2(x, g(x)) g'(x),
$$

a linear equation. In the neoclassical growth model, we instead have

$$
R(x, z) = h \left(x, z(x), z^2(x) \right)
$$

since the Euler equation involves $g^2(x)$. Hence

$$0 = R_1(x, g) = h_1\left(x, g(x), g^2(x)\right) + h_2\left(x, g(x), g^2(x)\right) \cdot g'(x)$$
$$+ h_3\left(x, g(x), g^2(x)\right) \cdot g'(g(x)) \cdot g'(x).$$

In general, this is a complicated equation, but we consider only a value $x^\star$ of x at which $g(x^\star) = x^\star$. Then we obtain a quadratic in $g'(x^\star)$:

$$0 = h_1(x^\star, x^\star, x^\star) + h_2(x^\star, x^\star, x^\star) \cdot g'(x^\star) + h_3(x^\star, x^\star, x^\star) \cdot (g'(x^\star))^2.$$

This raises the important point that a perturbation method will in general deliver multiple solutions for the derivatives of g. We saw than when $R(x, z)$ depends only on $(x, z(x))$, the solution is unique, whereas when it depends on $(x, z(x), z^2(x))$ and we consider an $x^\star$ at which $g(x^\star) = x^\star$, we obtain a quadratic. If R were to depend on other features of z then we'd have an even more complicated equation, e.g. a cubic when R depends on $(x, z(x), z^2(x), z^3(x))$ and we consider $x^\star$ such that $g(x^\star) = x^\star$.

In dynamic programming, R will always be the error in the Euler equation. The Euler equation is a second-order difference equation, so R will always depend only on $(x, z(x), z^2(x))$, yielding a quadratic in $g'(x^\star)$ when $x^\star$ satisfies $g(x^\star) = x^\star$. In this context, is unsurprising that there are multiple solutions, since we know that the Euler equation has a continuum of solutions. The solutions to the Euler equation that we're after are all the ones that satisfy the transverality condition, which could in principle include both roots. In the neoclassical growth model, our lives are made easier by the fact that we know that there is a unique solution g to the Euler equation that satisfies transversality, so we know that we're after only one of the roots. Moreover we didn't have to check directly which root satisfies transversality: instead we made use of the previously-established fact that $k^\star$ is locally stable under g, so that $g'(k^\star) < 1$. (I'm guessing that this stability-based argument can be extended to a much broader class of dynamic programs.)

Having obtained $g'(k^\star)$, we can continue to obtain higher derivatives in the same manner. To obtain $g''(k^\star)$, solve $R_{11}(k^\star, g) = 0$; to obtain the ℓth derivative, set the ℓth derivative of $R(k^\star, g)$ to zero and solve. Larry says that whereas we needed to solve a quadratic to obtain $g'(k^\star)$, the equations for higher derivatives of g are all linear. This isn't obvious to me, but I'm guessing it relies on the fact that when we have two 'initial conditions' $g(k^\star)$ and $g'(k^\star)$ there is no 'indeterminacy'. Larry also says that this is a very general feature: beyond the first derivative, all the equations are linear.

With $g'(k^\star)$ in hand, we can compute a first-order approximation to g around $k^\star$:

$$g(k) \simeq g(k^\star) + g'(k^\star)(k - k^\star) = k^\star + g'(k^\star)(k - k^\star).$$

This approximation will be good near $k^\star$, and possibly very poor away from $k^\star$. If we went to to compute higher-order derivatives, we can compute higher-order Taylor approximations.

Using the first-order Taylor approximation, we can study the quantitative question of how long it takes to close $100 \cdot (1 - \theta)\%$ of the gap between k_0 and $k^\star$. (We're mainly interested in $\theta = 0.05$, i.e. the time taken to close 95% of the gap.) Formally, we seek a $T \in \mathbf{R}_+$ such that

$$\frac{\left|k_{\lceil T \rceil - 1} - k^\star\right|}{|k_0 - k^\star|} > \theta \quad \text{and} \quad \frac{\left|k_{\lceil T \rceil} - k^\star\right|}{|k_0 - k^\star|} \leq \theta.$$

(This only makes sense because k_t converges monotonically to $k^\star$.) Since $g(k) - k^\star \simeq g'(k^\star)(k - k^\star)$ and $\text{sign}(g(k) - k^\star) = \text{sign}(k - k^\star)$ by monotone convergence,

$$\frac{|g(k) - k^\star|}{|k - k^\star|} \simeq g'(k^\star)$$

for all $k > 0$. So for any $s \in \mathbf{Z}_+$, it holds for $k = g^{s-1}(k_0)$:

$$\frac{|g^s(k_0) - k^\star|}{|g^{s-1}(k_0) - k^\star|} \simeq g'(k^\star).$$

Hence

$$\frac{|k_t - k^\star|}{|k_0 - k^\star|} = \frac{\left|g^t(k_0) - k^\star\right|}{|k_0 - k^\star|}$$

$$= \frac{\left|g^t(k_0) - k^\star\right|}{\left|g^{t-1}(k_0) - k^\star\right|} \times \frac{\left|g^{t-1}(k_0) - k^\star\right|}{\left|g^{t-2}(k_0) - k^\star\right|} \times \cdots \times \frac{\left|g^2(k_0) - k^\star\right|}{|g(k_0) - k^\star|} \times \frac{|g(k_0) - k^\star|}{|k_0 - k^\star|}$$

$$= \prod_{s=1}^{t} \frac{|g^s(k_0) - k^\star|}{|g^{s-1}(k_0) - k^\star|} \simeq \prod_{s=1}^{t} g'(k^\star) = (g'(k^\star))^t.$$

So (one valid) T solves $(g'(k^\star))^T = \theta$, which rearranges to

$$T = \frac{\ln(\theta)}{\ln(g'(k^\star))}.$$

We can verify by differentiating (and recalling that $g'(k^\star) < 1$) that T is strictly *increasing* in $g'(k^\star)$. This is (geometrically) intuitive.

Recall from previously that

$$\phi = 1 + \beta^{-1} + \left| \frac{u'(c^\star)}{u''(c^\star)} \right| \left| \frac{f''(k^\star)}{f'(k^\star)} \right|.$$

ϕ is strictly decreasing in the relative curvature $|u''/u'|/|f''/f'|$ of the utility and production functions. Recall further that $g'(k^\star)$ is given by

$$g'(k^\star) = \frac{\phi - \sqrt{\phi^2 - 4\beta^{-1}}}{2}.$$

We can easily verify (by differentiation) that this expression is strictly decreasing in ϕ; hence it is strictly increasing in $|u''/u'|/|f''/f'|$. Hence T is strictly increasing in $|u''/u'|/|f''/f'|$, so convergence is fast when $|u''/u'|$ is *small* relative to $|f''/f'|$.

The economics of this can be seen from the Euler equation

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

Suppose f is highly curved and that u is not. If we decrease k_{t+1}, $f'(k_{t+1})$ goes way up, so we have to compensate by decreasing c_t relative to c_{t+1}. Since u is not very curved, we have to decrease c_t quite a lot relative c_{t+1} to move $u'(c_t)$ and $u'(c_{t+1})$ by a substantial amount. This means that the saving rate is high, and hence that the capital stock will converge quickly to the steady-state level.

We've now studied the determinants and comparative statics of the convergence speed. If we want an numerical estimate of T, we have to parametrise. Let $f(k) = k^\alpha + (1-\delta)k$ with $\alpha = 0.36$ and $u(c) = (1-\sigma)^{-1}c^{1-\sigma}$ with $\sigma = 1/1.03$. Also set $\beta = 0.97$ and $\delta = 0.1$.[12] With this parameterisation,

$$\frac{u''(c^\star)}{u'(c^\star)} = \frac{(c^\star)^{-\sigma}}{-\sigma(c^\star)^{-(\sigma+1)}} = -\frac{c^\star}{\sigma} = -\frac{f(k^\star) - k^\star}{\sigma}.$$

and

$$\frac{f''(k^\star)}{f'(k^\star)} = -\frac{\alpha(1-\alpha)(k^\star)^{\alpha-2}}{\alpha(k^\star)^{\alpha-1} + 1 - \delta}.$$

Neither curvature coefficient is a free parameter, since they depend on the fundamental parameters α, σ, β and δ (directly and via $k^\star$). Possibly interesting is that a higher δ increases the curvature of f.

---

[12]This calibration is for when time periods are one year long. To convert to quarterly time periods, use $\delta_Q = 1 - (1 - \delta)^{1/4}$ and $\beta_Q = \beta^{1/4}$. The calibration of the other parameters α and σ does not depend on the length of time period.

For the annual calibration, we obtain $T \simeq 23$, $g'(k^\star) \simeq 0.88$ and $\phi \simeq 2.0524$. For the quarterly calibration, we get $T_Q \simeq 87$, and $T_Q/4 \simeq 22$, so not that different. (I should think that they would coincide exactly if we used the continuous-time variant of the model.) At this point, Larry sprinted through a bunch of other calibrations and associated convergence speeds.

6.3 Preliminaries for competitive equilibrium

There are three conceptual issues we have to deal with as preliminary steps on the way to sensible definitions of competitive equilibrium. These are the importance of constant returns to scale, the factor price frontier, and the indeterminacy of firm size under constant returns (hence the existence of a representative firm). We will start on this today and wrap up in the next lecture.

The idea behind competitive equilibrium is to examine how decentralised markets allocate resources. We let production take place in profit-maximising competitive (atomistic) firms, and let the representative consumer buy goods from firms. For all the notions of competitive equilibrium we use, it will turn out that the competitive allocation coincides with the planner's optimum studied earlier (the first welfare theorem).

The seemingly natural way to study competitive equilibria is as follows. Firms have technology that converts capital into output, captured by a production function f. (There is a change of notation here: f no longer includes remaining undepreciated capital.) Households own the capital stock k, and firms rent it at a rental rate r that they take as given. Firms choose the profit-maximising amount of capital to rent. Normalising the price of the consumption good to unity,[13] this means that the firm chooses k to solve

$$\max_{k \in \mathbf{R}_+} \left[f(k) - rk \right].$$

The problem is that since f is strictly concave, for any $k > 0$ we have

$$2f\left(k/2\right) - 2(k/2) = 2f\left(k/2\right) - k > f\left(k\right) - k,$$

so that a firm can always increase its profits by splitting into two separate entities. Hence there is no positive firm size that can ever be optimal. But neither can firm sizes be zero, since then output is zero, so each firm would

---

[13]This is without loss of generality in formal terms, but it does change the interpretation of r. When the price of consumption is normalised to unity, r is the *real* rental rate.

do better by entering the market with strictly positive k.[14] When f is strictly convex, there is a similiar problem: exactly one firm produces at the optimum, but with just one firm the assumption of price-taking is highly implausible.

This discussion suggests that the marginal product of capital must be constant for competitive equilibrium to make sense, but what we actually need is just constant returns to scale. In the one-factor model just considered, constant returns and constant marginal product are the same thing, but we can retain decreasing marginal product with constant returns by adding a second factor of production.

To this end, add labour n to the production technology. The household supplies $n = 1$ inelastically to firms. We can still have f strictly concave in k, corresponding to (something slightly stronger than) a decreasing marginal product of capital. To dodge paradoxes, we do impose that returns to scale are constant, which corresponds of f being homogeneous of degree 1. Now each firm solves

$$\max_{(k,n)\in\mathbf{R}^2_+} [f(k,n) - rk - wn].$$

We assume the following Inada conditions on f to ensure an interior optimum:

$$\lim_{k\to 0} f_1(k,n) = \infty \quad \text{and} \quad \lim_{n\to 0} f_2(k,n) = \infty.$$

Then provided that a solution exists, it satisfies the first-order conditions $\nabla f(k,n) = (r,w)^\top$, or $\nabla f(k/n, 1) = (r,w)^\top$ by homogeneity. But for arbitrary (r,w), these equations cannot both hold. Explicitly, they requires that $(f_1)^{-1}(r) = k/n = (f_2)^{-1}(w)$, which means they can hold only if $(f_1)^{-1}(r) = (f_2)^{-1}(w)$. We call the set of (r,w) that satisfy $(f_1)^{-1}(r) = (f_2)^{-1}(w)$ the factor price frontier, a downward-sloping curve in $r - w$ space. Clearly a necessary condition for competitive equilibrium is that (r,w) lies on the factor price frontier.

What goes wrong when (r,w) are off the factor price frontier? By Euler's formula, profit at (r,w) and (k,n) is

$$f(k,n) - rk - wn = [f_1(k/n, 1) - r]k + [f_2(k/n, 1) - w]n.$$

Whenever the FOCs hold, profit is zero. But consider a point (r,w) above the factor price frontier. Here profit is negative whatever (k,n) is chosen,

---

[14]I really don't see why this should be a problem. The whole point of measure theory is to circumvent exactly this issue. We have just define a measure μ such that for a *measurable* subset B of firms, $\mu(B)$ is the total output of these firms. We can choose it to be zero for all singletons, but strictly positive when B includes all firms. (We could let μ be Lesbegue measure.) There are bigger fish to fry, but it's worth a footnote.

so no output is produced. But there is positive demand for output from households, this this cannot happen in equilibrium. Now consider a point (r, w) below the factor price frontier. Then profit is unbounded above since profit is linear in k and n with at least one positive coefficient, so there's no solution to the firm's maximisation problem.

To me at least, this argument is confusing. It assumes the first-order conditions hold, then shows that this implies a parametric restriction! Another way to see that the 'first-order conditions' must hold is to use the argument in the previous paragraph directly. Whenever we are not on the factor price frontier, at least one coefficient is always nonzero, leading to either zero output or no solution to the problem. Neither can happen in equilibrium, so we must be on the factor price frontier in equilibrium. It's an equilibrium condition, not an implication of optimal behaviour by firms! When we are on the factor price frontier, the firm can make at most zero profit. It earns zero profit by setting $\nabla f(k/n, 1) = (r, w)^\top$, so this condition characterises a firm's factor demands.

7 Mon 12 Oct 2015

7.1 Outline of today

Last time, we argued that competitive equilibrium only makes sense with constant returns to scale and factor prices on the factor price frontier. We need constant returns to scale because with decreasing returns, firms will split endlessly, and with increasing returns we'll get monopoly, which makes price-taking implausible. We need to be on the factor price frontier because all firms shut down above it and want to produce unbounded output above it.

We begin today by studying another implication of competitive equilibrium: the indeterminacy of firm size. This will lead us to consider a representative firm. We will contrast this with the idea of a representative household.

Having cleared these preliminaries, we'll define several notions of competitive equilibrium for infinite-horizon dynamic economies. We'll study three definitions. The first is Arrow–Debreu equilibrium, which is super-weird. The second is sequence-of-markets equilibrium, which is weird. These first two are both analogous to the sequence problem. The third equilibrium concept is recursive competitive equilibrium, which is analogous to the functional equation approach. Larry thinks this notion of equilibrium is natural, though it looks odd at first.

7.2 The indeterminacy of firm size under constant returns

Suppose there are $M \in \mathbf{N}$ firms. (We'll consider a finite number of firms for simplicity.) Let (r, w) lie on the factor-price frontier. Price-taking profit-maximisation by firm i yields first-order conditions $\nabla f(k_i, n_i) = (r, w)^\top$. So clearly every firm i chooses the same ratio k_i and n_i. Define 'aggregate' capital and labour as $k := \frac{1}{M} \sum_{i=1}^M k_i$ and $n := \frac{1}{M} \sum_{i=1}^M n_i$. (We choose to use the average as our aggregator, but that is merely a choice of convenience.) Then obviously $k_i/n_i = k/n$ for each i. Let aggregate output be $y := \frac{1}{M} \sum_{i=1}^M y_i$ where $y_i = f(k_i, n_i)$. By Euler's formula,

$$ y = \frac{1}{M} \sum_{i=1}^M \nabla f(k_i, n_i) \cdot (k_i, n_i)^\top = \nabla f(k, n) \cdot (k, n)^\top = f(k, n). $$

So aggregate output, capital and labour are as if there were a single price-taking firm producing output y with technology f by renting capital k and hiring labour n. The theory takes no stand on the distribution of output

production and input use across firms. So without loss of generality, we can consider a single representative firm, or M identical firms, and obtain results that hold equally well if firms are not all of the same size.

The existence of a representative firm is very different from the existence of a representative household. Whereas we showed that there exists a representative firm whenever all firms use an identical unit homogeneous technology, there are no mild conditions under which a collection of households can be represented by a single household. Even with identical utility functions is not enough: you also need homogeneous endowments of capital and labour. The weakest known conditions for aggregation across households are the Gorman conditions, which are strong.

7.3 Arrow–Debreu equilibrium

An equilibrium consists of a profile of actions for each agent, as well as (possibly) certain pure equilibrium objects (usually prices) that enter the agents' decision problems but are not chosen by anyone. In equilibrium, each agent's action is optimal taking the other agents' actions and the equilibrium objects as given. Moreover we (sometimes) impose additional consistency requirements (such as market clearing) that ensure that the actions of the agents do not conflict. The consistency conditions can be thought of as pinning down the pure equilibrium objects. In the case of competitive equilibrium, we can think of the market-clearing conditions as pinning down the equilibrium prices, i.e. 'prices adjust to clear markets'. In the case of rational-expectations equilibrium (with a common prior given exogenously), the belief is a pure equilibrium object, and we impose the consistency requirement that beliefs be updated according to Bayes's rule following positive-probability events.[15]

Our first concept of competitive equilibrium in an infinite-horizon dynamic economy is Arrow–Debreu. Debreu (1959) pointed out that you can reinterpret the commodities in classical general equilibrium theory as date-indexed commodities. In our simple model, we have a countably infinite collection $\{c_t, n_t\}_{t=0}^{\infty}$ of commodities, with corresponding prices $\{p_t, p_t w_t\}_{t=0}^{\infty}$. Similarly, there is an infinite sequence of asset holdings $\{k_t\}_{t=0}^{\infty}$ and returns $\{p_t r_t\}_{t=0}^{\infty}$.

To shoehorn our dynamic economy into the classical static literature on general equilibrium, we can only allow trading to take place once and for

---

[15]Outside the context of rational expectations, we still have consistency conditions. When priors are exogenous but heterogeneous, we still impose Bayes's rule where applicable to *each* agent's belief. In the 'learning' literature we impose that beliefs evolve according to some exogenous learning rule, e.g. running OLS regressions.

all. So markets for consumption, labour and capital open and close at date 0; we can think of households and firms as trading futures. (When there's also uncertainty, there are securities markets for each state-date, and the securities are called Arrow securities.)

As usual we assume that there is a competitive representative household which chooses $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$ at date 0 subject to budget and nonnegativity constraints, taking prices $\{p_t, p_t w_t, p_t r_t\}_{t=0}^{\infty}$ as given. Formally, the household's problem is

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

$$\text{s.t.} \quad \sum_{t=0}^{\infty} p_t(c_t + [k_{t+1} - (1-\delta)k_t] - w_t n_t - r_t k_t - \pi_t) \leq 0$$

$$c_t \in \mathbf{R}_+, \quad n_t \in [0,1], \quad k_{t+1} \in \mathbf{R}_+ \quad \text{for each } t \in \mathbf{Z}_+.$$

where $[k_{t+1} - (1-\delta)k_t]$ is date-t investment and π_t is the date-t profit of all firms (since they are owned by the household).[16] Since n_t does not enter the utility function, the household inelastically supplies $n_t = 1$ at each date t.

Since firms are owned by the household, it would be nice to justify the assumption that firms maximise profit rather than, say, maximising equilibrium household utility! Apparently, we can get profit-maximisation when ownership is highly dispersed. On another note, we have not allowed shares in firms to be traded. Since there is just one household (or several identical households), this is without loss of generality.

In the general equilibrium setting, the firms here are actually producing several (infinitely many!) goods. Since firms act competitively, they do not take into account the substitution patterns across goods by the household, so maximising undiscounted profits is equivalent to maximising profits period-by-period:

$$\arg\max_{\{n_t, k_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t \left[f(k_t, n_t) - (w_t n_t + r_t k_t) \right]$$

$$= \prod_{t=0}^{\infty} \arg\max_{n_t, k_t} p_t \left[f(k_t, n_t) - (w_t n_t + r_t k_t) \right].$$

(This is a Cartesian (set) product, not an ordinary one!) (We suppressed the nonnegativity constraints for simplicity.) Denote the maximised value

---

[16] Implicitly, the households own a technology that turns one consumption unit of investment into one unit of new capital. In some later models, we'll have firms operating this technology. This provides a place where we can build in financial frictions.

of profit at prices (p, w, r) by $\pi(p, w, r)$. In the equilibrium definition below, the π_t in the household's problem will be identically equal to $\pi(p_t, w_t, r_t)$.

Note well that both the household and the firm choose values for capital, labour and output. What's really going on is that for given prices, the solution to the household's problem gives us demand for the output good and supply of labour and capital services, while the solution to the firm's problem gives us the supply of the output good and demand for labour and capital services. In equilibrium, these will have to coincide, giving us the market-clearing consistency conditions.

Definition 2. Price and quantity sequences $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ and $\{c_t, n_t, k_t\}_{t=0}^{\infty}$ are an Arrow–Debreu equilibrium iff

(1) $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$ solves the household's problem at prices $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ starting at k_0.

(2) At each date t, (k_t, n_t) solves the firm's problem at prices (p_t, w_t, r_t) and $f(k_t, n_t) = c_t + k_{t+1} - (1 - \delta)k_t$.

As I've written it, the market-clearing conditions are baked in. To make them explicit, we can consider quantities $\{c_t, n_t, k_t, y_t, \widetilde{n}_t, \widetilde{k}_t\}_{t=0}^{\infty}$. We require that $(\widetilde{k}_t, \widetilde{n}_t)$ solves the firm's problem and that $y_t = f(\widetilde{k}_t, \widetilde{n}_t)$ (technological constraint). We then explicitly impose goods-market clearing

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

and factor-market clearing $(\widetilde{k}_t, \widetilde{n}_t) = (k_t, n_t)$.

The following are properties of any Arrow–Debreu equilibrium in the neoclassical growth model with labour.

(1) An equilibrium exists. We will not prove this directly, but it follows from the second welfare theorem: there is an optimum, and it must be supportable in an equilibrium.

(2) Any Arrow–Debreu equilibrium is 'zero-homogeneous in prices'. Formally, $(\{p_t, w_t, r_t\}_{t=0}^{\infty}, \{c_t, n_t, k_t\}_{t=0}^{\infty})$ is an Arrow–Debreu equilibrium iff $(\{\theta p_t, \theta w_t, \theta r_t\}_{t=0}^{\infty}, \{c_t, n_t, k_t\}_{t=0}^{\infty})$ for every $\theta > 0$ is an Arrow–Debreu equilibrium. This is a direct implication of the zero-homogeneity of demands and one-homogeneity of technology.

(3) A class of properties we will not study directly is uniqueness. But in the present case, there is a unique allocation $\{c_t, n_t, k_t\}_{t=0}^{\infty}$ that can be part of an Arrow–Debreu equilibrium (though there is a continuum of prices

that will support this allocation by property (2)). Uniqueness here follows from the first welfare theorem: the efficient allocation is unique, and only efficient allocations can be part of equilibria. (Obviously, the previous property implies that we can never hope for uniqueness of equilibria, only uniqueness of equilibrium allocations.)

(4) All equilibrium prices $\{p_t, r_t, w_t\}_{t=0}^{\infty}$ are strictly positive. For suppose some price is nonpositive. Then either the household's or the firm's optimisation problem has no solution, since the demand-side wants an unbounded amount of whatever dated commodity is free.[17]

(5) $\pi_t = \pi(p_t, w_t, r_t) = 0$ at each date t. For suppose $\pi_t > 0$; then by unit homogeneity of the function π (due to Euler's formula), firm can double profit again, so the putative equilibrium quantities do not solve the firm's problem. If $\pi_t < 0$ then the firm would make more profit (viz. zero) by producing nothing, so again the putative equilibrium quantities do not solve the firm's problem.

(6) The first welfare theorem holds: every Arrow–Debreu equilibrium is efficient. Since there's a single household, this just means that the allocation part of every Arrow–Debreu equilibrium coincides with one solution to the planner's problem.

For suppose there were another feasible allocation $\left\{\bar{c}_t, \bar{n}_t, \bar{k}_t\right\}_{t=0}^{\infty}$ with

$$\sum_{t=0}^{\infty} \beta^t u(\bar{c}_t) > \sum_{t=0}^{\infty} \beta^t u(c_t).$$

Then this new allocation can't have been affordable at the equilibrium prices or it would have been chosen by the household:

$$\sum_{t=0}^{\infty} p_t \left(\bar{c}_t + \left[\bar{k}_{t+1} - (1-\delta)\bar{k}_t\right] - w_t \bar{n}_t - r_t \bar{k}_t\right) > 0.$$

By feasibility of $\left\{\bar{c}_t, \bar{n}_t, \bar{k}_t\right\}_{t=0}^{\infty}$, we must have

$$\bar{c}_t + \left[\bar{k}_{t+1} - (1-\delta)\bar{k}_t\right] = f\left(\bar{k}_t, \bar{n}_t\right) \quad \text{at each } t \in \mathbf{Z}_+.$$

---

[17] A common but incorrect way to say this is that 'demand is infinite, which violates market clearing'. But really market clearing has nothing to do with it. Rather, the problem is that an optimisation problem has no solution, so in particular the putative equilibrium quantities cannot solve both the household's and the firm's problems.

Using this together with the fact that profits are zero in equilibrium,

$$\sum_{t=0}^{\infty} p_t[f(\bar{k}_t, \bar{n}_t) - w_t\bar{n}_t - r_t\bar{k}_t]$$

$$> 0 = \sum_{t=0}^{\infty} \pi(p_t, r_t, w_t) = \sum_{t=0}^{\infty} p_t[f(k_t, n_t) - w_t n_t - r_t k_t].$$

But this means that at prices $\{p_t, r_t, w_t\}_{t=0}^{\infty}$, the firm can make strictly higher profits by choosing $\{\bar{k}_t, \bar{n}_t\}_{t=0}^{\infty}$ instead of $\{k_t, n_t\}_{t=0}^{\infty}$. So the original equilibrium can't have been an equilibrium after all.

7.4 Sequence-of-markets equilibrium

Arrow–Debreu equilibrium is super-weird because all trade takes place ahead of time. Consumers do not purchase consumption goods and rent out capital and labour services once in their lifetimes; actual trading arrangements are more like spot markets. Even if we thought that people contracted on all future quantities in advance, we're left with the question on what period 0 might be. A second issue is that the informational requirements are high. Consumers need to anticipate the entire infinite sequence of prices in order to make their optimal choice.

Sequence-of-markets equilibrium addresses the first concern, but not the second. In this setup, markets for t-dated goods and services open (and close) at date t. But the household's problem is still inherently dynamic because it involves saving decisions. We therefore still require the household to solve a sequence problem each time it trades. Now it trades at each date t, so the household has to solve a sequence problem at each date instead of a single sequence problem! In particular, at date t, the household solves the Arrow–Debreu household problem with prices truncated to start at date t and initial capital k_t. While only date-t quantities are chosen at t, the household still forms a 'plan' for future dates. Sequence-of-markets equilibrium requires that subsequent choices coincide with these plans.[18]

Definition 3. Price and quantity sequences $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ and $\{c_t, n_t, k_t\}_{t=0}^{\infty}$ are a sequence-of-markets equilibrium iff

(1) At each date t, $\{c_{t+j}, n_{t+j}, k_{t+1+j}\}_{j=0}^{\infty}$ solves the household's problem at prices $\{p_{t+j}, w_{t+j}, r_{t+j}\}_{j=0}^{\infty}$ starting at k_t.

---

[18]This means that a sequence-of-markets equilibrium will fail to exist whenever preferences are time-inconsistent.

(2) At each date t, (k_t, n_t) solves the firm's problem at prices (p_t, w_t, r_t) and $f(k_t, n_t) = c_t$.

Whenever we have time consistency, the set of sequence-of-markets equilibria coincides with the set of Arrow–Debreu equilibria. As mentioned above, in the absence of time consistency, sequence-of-markets equilibria will not exist. It is unsurprising that the two equilibrium notions fail to coincide in the presence of time inconsistency, since time-0 trading is a commitment technology that is not available when there are only spot markets.

8 Wed 14 Oct 2015

8.1 Outline of today

First we'll define and discuss recursive competitive equilibrium. We'll then say a bit about complete markets, which is relevant for Arrow–Debreu equilibrium under uncertainty. This leads naturally to a discussion of insurance, which requires both uncertainty and ex-post heterogeneity of agents.

Finally we'll talk about time consistency and commitment vs. discretion. In the Arrow–Debreu model, all agents can (must!) commit at date 0 to their choices for future variables. In contrast, there is no commitment technology in the basic spot-market trading model in which we defined sequence-of-markets equilibrium. This absence of commitment technology can be viewed as a form of market incompleteness, since markets for future goods and services are missing. In the context of insurance, we will see that market incompleteness leads to incomplete insurance in equilibrium.

8.2 Recursive competitive equilibrium

Two problems with Arrow–Debreu equilibrium are that it involves all trade taking place in advance at a mysterious date 0, and that agents solving dynamic problems have to explicitly take into account the entire future sequence of prices when making their time-t decisions.

To rid ourselves of the first problem, we replaced the time-0 complete-markets model of Arrow–Debreu with a sequence of spot markets, and defined a sequence-of-markets equilibrium in this context. The mysterious date 0 vanished, as did the need to commit in advance. But the informational problem is at least as bad: the household has to take into account the entire future sequence of prices every time it makes a decision, and now it has to make an infinite sequence of decisions as well!

The motivation for introducing recursive competitive equilibrium is to fix this flaw in the sequence-of-markets equilibrium concept. But there's a bit of charlatanry here, as Larry admitted later. The household has to know the value function v, which may be an infinite-dimensional object. Even if it is not, computing v is surely just as hard as solving a sequence problem! So it's really not clear that the informational requirements of this approach are any more lenient than those Arrow–Debreu model. Other authors motivate recursive equilibrium in other ways, e.g. ease of computation and the fact that it restricts attention to Markov decision rules.

Let's return to and modify the Stokey–Lucas canonical model for the household in a generic competitive economy. We assume that households

are homogeneous, so they can be represented by a single household solving a Stokey–Lucas dynamic program. This household acts competitively, not taking into account the effect of her actions on the aggregate state of the economy. (She fails to interalise her effect not just on prices, but also on other aggregate quantities such as output.) To formalise this, denote by $\widetilde{x}_t \in X$ the 'aggregate state' that the household views as outside her control. In the neoclassical growth model, $\widetilde{x}_t$ is the aggregate capital stock in the economy. Write $x_t \in X$ for the 'individual state' that the individual chooses. In the neoclassical growth model, this is the capital stock chosen by the representative household. Our equilibrium concept will impose $\widetilde{x}_t = x_t$, but *after* the maximisation step.[19]

Another tweak to the Stokey–Lucas framework is that the household's set of feasible actions depends on the prices $q_t \in \mathbf{R}^n$ of n commodities. In the neoclassical growth model, $q_t = (r_t, w_t)^\top$. We allow the household's constraint to depend on prices in an arbitrary way. In the neoclassical growth model, q_t affects the household's income, and hence constrains her expenditure via the time-t budget constraint.

So far, we have not strayed from the recursive formulation of a sequence-of-markets equilibrium. But the next step is special to recursive competitive equilibrium. We assume that exogenous features of the environment evolve according to a first-order dynamic system. (With uncertainty, it'll be a Markov chain.) We have already assumed this by using the stationary dynamic program from chapter 4 of Stokey et al. (1989). To illustrate, there are no exogenously evolving features of the environment in the neoclassical growth model. But everything that will be said subsequently extends to the case where the environment evolves deterministically according to a first-order dynamic system, and to the case where it follows a first-order Markov chain.[20]

Sequence-of-markets and Arrow–Debreu equilibria can still be history-dependent in this stationary environment, despite the fact that the fundamentals are history-independent. Recursive competitive equilibrium (like Markov-perfect equilibrium) rules out history-dependent allocations and prices. In particular, we suppose that the household believes that the aggregate state transitions according to the first-order law of motion $\widetilde{x}_{t+1} = G(\widetilde{x}_t)$ (the 'perceived law of motion'), and that today's prices q_t are pinned down by

---

[19]This separation of aggregate and individual is sometimes called the 'big K, little k' trick. The name comes from using capital K for the aggregate state and little k for the individual state.

[20]The neoclassical growth model with disembodied exogenous technical change exemplifies the former; the same model with stochastic technology (i.e. the basic RBC model) is an example of the latter.

today's aggregate state according to $q_t = q(\widetilde{x}_t)$. The functions $G : X \to X$ and $q : X \to \mathbf{R}^n$ will obviously be equilibrium objects, but will be left exogenous for now.

Having ruled out history-dependence in the objects $\widetilde{x}_t$ and q_t that the household views as exogenous, we recover a recursive structure for the household's problem. Write $\Gamma : X^2 \times \mathbf{R}^n \to X$ for the household's constraint, which now depends both on the individual choice x_t, the aggregate $\widetilde{x}_t$, and the price q_t. The household's value function v, which now depends on both the individual and aggregate states, satisfies the functional equation

$$v(x, \widetilde{x}) = \max_{x' \in \Gamma\left(x, \widetilde{x}, q(\widetilde{x})\right)} \left[F(x, x') + \beta v(x', G(\widetilde{x}))\right] \quad \forall (x, \widetilde{x}) \in X^2. \quad (9)$$

Observe that v depends on q and G, which the household treats as given and which will be determined in equilibrium.[21] As usual, we'll write $g : X^2 \to X$ for an optimal policy function, i.e. a function that satisfies

$$v(x, \widetilde{x}) = F(x, g(x, \widetilde{x})) + \beta v(g(x, \widetilde{x}), G(\widetilde{x})) \quad \forall (x, \widetilde{x}) \in X^2. \quad (10)$$

We assume that firms use n inputs priced at q to produce an output good with price normalised to unity, and that their problem is time-separable. In the neoclassical growth model, inputs are capital and labour services with prices r and w. Profit-maximising behaviour will then imply n optimality conditions (usually first-order conditions) that relate prices to the aggregate state $\widetilde{x}$. We have to make sure to define the aggregate state so that this is true! In the neoclassical growth model, there are two prices, and the price function is

$$q(\widetilde{k}) := \begin{pmatrix} f_1(\widetilde{k}, 1) \\ f_2(\widetilde{k}, 1) \end{pmatrix} \quad \forall \widetilde{k} \in \mathbf{R}_+.$$

This involves only $\widetilde{k}$, validating our choice of the capital stock alone as the state variable. The optimality conditions then imply a function $q : X \to \mathbf{R}^n$ that maps values for the aggregate state into prices consistent with profit-maximisation by competitive firms. We say that this function 'solves the firm's problem'.

We are now ready to define recursive competitive equilibrium. The definition below is based on the original definition of Prescott and Mehra (1980); Larry gave a looser version.

Definition 4. A value function v, policy function g, transition function G and price function q form a recursive competitive equilibrium iff

---

[21]I think that we can allow F to depend on $\widetilde{x}$ and $G(\widetilde{x})$ without causing trouble.

(1) v is the value function of the household's problem at price function q and perceived transition function G, i.e. v solves (9).

(2) g is an optimal policy for the household's problem at price function q and perceived transition function G, i.e. g solves (10).

(3) $g\left(\widetilde{x}, \widetilde{x}\right) = G\left(\widetilde{x}\right) \ \forall \widetilde{x} \in X$ ('rational expectations').

(4) q solves the firm's problem.

A slightly funny feature of this equilibrium definition is that it says nothing about sequences of prices and quantities! In contrast, Arrow–Debreu and sequence-of-markets equilibria were explicitly defined as sequences. To remedy this, the following concept formalises the notion that a sequence of prices and quantities 'is an RCE'.

Definition 5. Price and quantity sequences $\{q_t\}_{t=0}^{\infty}$ and $\{x_t\}_{t=0}^{\infty}$ are an RCE sequence iff there is an RCE (v, g, G, q) such that

(1) At each date t, $x_{t+1} = G(x_t)$.

(2) At each date t, $q_t = q(x_t)$.

In any model in which the environment has the requisite Markov property and preferences are time-consistent, the set of price and quantity sequences induced by recursive competitive equilibria are obviously a subset of the set of sequence-of-markets and Arrow–Debreu equilibria. For a recursive competitive equilibrium then induces the same price and quantity sequence as any history-independent sequence-of-markets equilibrium, and these coincide with the Arrow–Debreu equilibria under time consistency. So under these conditions, all recursive competitive equilibria are efficient.

When time-consistency fails, we know already that sequence-of-markets equilibria may fail to exist, and certainly won't coincide with Arrow–Debreu equilibria. I believe that the set of sequences induced by recursive competitive equilibria need not coincide with either of the others. Arrow–Debreu equilibria give us the full-commitment solution, sequence-of-markets equilibria give us no solution since no plan is time-consistent, whereas recursive equilibria induce the behaviour we would expect from a (non-myopic) time-inconsistent planner with no commitment technology.

In some models, including the neoclassical growth model, the set of sequences induced by recursive competitive equilibria coincides exactly with

the set of sequence-of-markets equilibria.[22] To begin to see this intuitively, we could follow the usual steps (first-order condition plus envelope condition) to show that if (v, G, q) is a recursive competitive equilibrium then G satisfies the Euler equation. This is not sufficient, however, since we also need the transversality condition to be respected.

8.3 Arrow–Debreu equilibrium under uncertainty

We will now introduce a simple variety of uncertainty. At each t we have a random variable s_t taking values in the finite set $S = \{s(1), \ldots, s(m)\}$. A time-$t$ history is $h^t = (s_0, s_1, \ldots, s_t)$.[23] Note that the set of all possible histories is S^t, the t-fold Cartesian product of the state space S. We assume that s_0 is degenerate (so that we can focus on the transition probabilities). For each t, let μ_t be the distribution of histories on S^t. We will sloppily drop the subscript and write μ instead. We assume that $\mu(h^t) > 0$ for every $h^t \in S^t$. This makes conditional probabilities easy; we write these as (for $t > \tau$)

$$\mu(h^t | h^\tau) := \frac{\mu(h^t)}{\mu(h^\tau)}.$$

At an optimum or equilibrium, all quantities and prices will depend on the history h^t, so we write these as functions c, k, n, r and w. (Note that $k_{t+1} = k(h^t)$ since it is chosen at time t.) The difference from before is that we index by history rather than by date, and that we use functions rather than subscripts because it's more convenient.

We assume that μ is both the objective probability distribution and the (commonly known) belief of every agent. One loose justification is that the environment has been stable for a long time, so that agents have all Bayes-updated their way close to the truth. As we know from time series, this will only work under ergodicity assumptions on μ! When we impose the Markov property later on, we get ergodicity for free.[24]

---

[22]More broadly, we have the following. If the model has the requisite Markov property and preferences are time-consistent, the sequence-of-markets equilibrium is unique, and there exists a recursive competitive equilibrium, then obviously we have coincidence. This is the case in the neoclassical growth model.

[23]There's an unfortunate habit in macro of writing a history as s^t. Larry did this in the lecture, and so do e.g. Ljungqvist and Sargent (2012). I find this impossibly confusing, so I use h^t. (This is common in game theory.)

[24]Ergodicity is by no means enough for this to be justified. We're going to need something like the 'grain of truth' assumption of Kalai and Lehrer (1993), i.e. that all agents' priors assign strictly positive probability to the truth; otherwise they will not be able to Bayes-update their way to the truth. This is a very strong assumption!

We look at Arrow–Debreu equilibria first. Suppose we have complete markets: at date 0, there are markets for consumption goods, capital services and labour services for each current and future *history* h^t. As before, the household operates the one-for-one technology that transforms consumption goods into investment goods. The household's budget constraint at the time of trading is therefore

$$\sum_{t=0}^{\infty} \sum_{h^t \in S^t} p(h^t) \left[c(h^t) + k(h^t) - (1-\delta)k(h^{t-1}) \right.$$

$$\left. -r(h^t)k(h^t) - w(h^t)n(h^t) - \pi\left(p(h^t), r(h^t), w(h^t)\right) \right] \leq 0.$$

There are no probabilities here because these quantities and prices are all chosen at date 0. The budget constraint has nothing to do with the consequences of these choices in the future, so probabilities are beside the point.

The initial capital stock $k_0 = k\left(s^{-1}\right)$ is given. We assume that preferences are additively separable across time and states:

$$\sum_{t=0}^{\infty} \sum_{h^t \in S^t} \beta^t \mu(h^t) u\left(c(h^t)\right).$$

We make all the nice assumptions needed for the optimum to be characterised by the first-order conditions. Letting the multiplier be $\lambda \geq 0$, the first-order condition w.r.t. $c(h^t)$ is then

$$\beta^t \mu(h^t) u'\left(c(h^t)\right) = \lambda p(h^t).$$

The first-order condition w.r.t. $k(s^t)$ is

$$\lambda p(h^t) = \lambda \sum_{h^{t+1} \in S^{t+1}|h^t} p(h^{t+1}) \left[r(h^{t+1}) + 1 - \delta \right].$$

Combining the two to eliminate $\lambda p(h^t)$ yields

$$\beta^t \mu(h^t) u'\left(c(h^t)\right)$$
$$= \sum_{h^{t+1} \in S^{t+1}|h^t} \left[\beta^{t+1} \mu(h^{t+1}) u'\left(c(h^{t+1})\right) \right] \left[r(h^{t+1}) + 1 - \delta \right],$$

which rearranges to

$$u'\left(c(h^t)\right) = \beta \sum_{h^{t+1} \in S^{t+1}|h^t} \frac{\mu(h^{t+1})}{\mu(h^t)} u'\left(c(h^{t+1})\right) \left[r(h^{t+1}) + 1 - \delta \right].$$

We can write this in shorthand as

$$u'_t = \beta \mathbf{E}_t u'_{t+1} \left[r_{t+1} + 1 - \delta \right]$$
$$= \beta \mathbf{E}_t u'_{t+1} \left[f_{1,t+1} + 1 - \delta \right]$$

where we used one of the firm's first-order conditions. In a loose sense (since we won't check transversality), this suggests that the equilibrium allocation remains efficient despite the introduction of uncertainty.

9 Mon 19 Oct 2015

9.1 Outline of today

We'll first finish off our discussion of Arrow–Debreu equilibrium under uncertainty. We will then consider ex-post heterogeneous agents, and see that complete markets lead to complete insurance in equilibrium when agents are risk-averse. We'll then drop market incompleteness and study the degree of insurance in equilibrium.

9.2 Back to Arrow–Debreu equilibrium under uncertainty

Recall that in the time-0 trading model with complete markets introduced last time, the representative household's problem is

$$
\max_{c(\cdot),k(\cdot)} \sum_{t=0}^{\infty} \sum_{h^t \in H^t} \beta^t \mu(h^t) u\left(c(h^t)\right)
$$

$$
\text{s.t.} \quad \sum_{t=0}^{\infty} \sum_{h^t \in S^t} p(h^t) \left[c(h^t) + k(h^t) - (1-\delta)k(h^{t-1}) \right.
$$

$$
\left. - r(h^t)k(h^t) - w(h^t) - \pi\left(p(h^t), r(h^t), w(h^t)\right) \right] \leq 0
$$

$$
k(h^t), c(h^t) \geq 0 \quad \forall h^t \in S^t, \quad \forall t \in \mathbf{Z}_+.
$$

Since $n(h^t)$ does not enter utility but expands the budget constraint, $n(h^t) = 1$ is optimal at each $h^t \in S^t$ and $t \in \mathbf{Z}_+$. So we say that (c,k) solves the household's problem at prices (p,w,r) starting at k_0 iff (c,k) attains the maximum in the program above and $n(h^t) = 1$ everywhere.

Now we move on to the firm's problem. As usual the firm's problem is separable, so at each $h^t \in S^t$ for every $t \in \mathbf{Z}_+$ it solves

$$
\max_{(k,n)\in\mathbf{R}_+^2} p(h^t) \left[f(k,n) - r(h^t)k - w(h^t)n \right].
$$

The function π of prices (already used in the household's budget constraint) is defined as the value of this problem. Having defined the household's and firm's problems, we are ready to define equilibrium.

Definition 6. Price and quantity functions (p,w,r) and (c,n,k) are an Arrow–Debreu equilibrium iff

(1) (c,n,k) solves the household's problem at prices (p,w,r) starting at k_0.

(2) At each history $h^t \in S^t$ for each date t, $(k(h^t), n(h^t))$ solves the firm's problem at prices (p, w, r), and $f(k(h^t), n(h^t)) = c(h^t)$.

There's a funny feature of the uncertainy here: it doesn't affect fundamentals at all! In order for s_t to affect fundamentals, it would have to be an argument in f or u. When s_t is unrelated to fundamentals as here, we call it a sunspot. An equilibrium is a sunspot equilibrium iff it depends on the realised path h^t. In this model, there are no sunspot equilibria, so in every equilibrium, every sample path is identical to one of the equilibria of the Arrow–Debreu equilibria in the deterministic time-0 trading model studied last time.[25]

9.3 Complete insurance in Arrow–Debreu equilibrium

People often say that Arrow–Debreu equilibria exhibit complete insurance. To investigate this claim, we obviously need idiosyncratic risk, so consider the following elaboration of the model. There is a unit measure of agents, half of them type 1 and the other half type 2. Let the random variable s_t take values in $\{1/2 - \sigma, 1/2 + \sigma\}$ where $\sigma \in (0, 1/2)$.

There is no production. Instead, type i is endowed with $e^i(h^t)$ units of the consumption good at history h^t. In particular, we let $e^1(h^{t-1}, s_t) = s_t$ and $e^2(h^{t-1}, s_t) = 1 - s_t$, so that there is no aggregate risk. So when s_t is high, type-1 agents get a boost of σ at the expense of type-2 agents, and vice versa.

We're going to look at ex-ante Pareto-efficient allocations. We assume that u is concave. It follows that all and only Pareto efficient allocations are saddle-points of the Lagrangean

$$\sum_{t=0}^{\infty} \sum_{h^t \in S^t} \beta^t \mu(h^t) \left\{ \theta_1 u\left(c_1(h^t)\right) + \theta_2 u\left(c_2(h^t)\right) \right.$$
$$\left. + \lambda(h^t) \left[1 - c_1(h^t) - c_2(h^t) \right] \right\}$$

for some $(\theta_1, \theta_2) > 0$, where $\lambda(h^t)$ is the constraint on the history-h^t resource constraint. (Note that the constraint is node-by-node!) The first-order conditions are

$$\beta^t \mu(h^t) \left[\theta_i u'\left(c_i(h^t)\right) - \lambda(h^t) \right] = 0$$

---

[25] The result that sunspot equilibria are not possible in the Arrow–Debreu framework is due to Cass and Shell (1983). More later on sunspots and this paper!

for $i = 1, 2$ and each $h^t \in H^t$ and $t \in \mathbf{Z}_+$. It follows that

$$\theta_1 u' \left(c_1(h^t) \right) = \theta_2 u' \left(c_2(h^t) \right).$$

Together with the resource constraint (with equality since u is strictly increasing), this implies that

$$\theta_i u' \left(c_i(h^t) \right) = \theta_j u' \left(1 - c_i(h^t) \right),$$

for $i \in \{1, 2\}$ and $j \neq i$. So each agent's consumption is constant across all histories, meaning that every Pareto optimum involves complete insurance. As usual, Pareto efficiency puts no demands on the distribution of consumption; we can obtain any distribution by varying the Pareto weights θ_1 and θ_2.

Now let's look at decentralised allocations. We consider Arrow–Debreu equilibria in which markets open at date -1, i.e. before the realisation of any uncertainty. Household i's problem is as before, but with the budget constraint

$$\sum_{t=0}^{\infty} \sum_{h^t \in H^t} p(h^t) \left[e^i(h^t) - c_i(h^t) \right] \geq 0.$$

Let λ_i be the multiplier. The first-order condition for $c_i(h^t)$ is

$$\beta^t \mu(h^t) u' \left(c_i(h^t) \right) = \lambda_i p(h^t),$$

Hence for $i \in \{1, 2\}$ and $j \neq i$,

$$\lambda_i u' \left(c_i(h^t) \right) = \lambda_j u' \left(1 - c_i(h^t) \right).$$

So we must have complete insurance in any Arrow–Debreu equilibrium. This is unsurprising since the first welfare theorem holds here!

The distribution of income in equilibrium depends on the multipliers λ_i and λ_j. Notice that λ_1/λ_2 determines the distribution of consumption in the same way as θ_1/θ_2 did for the Pareto optima. Since the only ex-ante heterogeneity among agents is the distribution of their lifetime endowments (governed by μ), we can vary λ_1/λ_2 by varying μ. As we vary μ to favour one agent over the other, we trace out the set of Pareto optima, so the second welfare theorem holds as well.

9.4   Incomplete insurance

The foregoing analysis highlights a difficulty. To make it particularly stark, consider the case in which $\lambda_1 = \lambda_2$ (which will happen iff the transition

distribution is symmetric). Then each agent consumes $1/2$ in equilibrium. If $s_0 = 1/2 + \sigma$ then agent 1 has to give up σ at date 0 because she agreed to do so in advance. But now that she has a higher lifetime income path, it may be that agent 1 does not find it in her interest to respect the contracts she signed at the time of trading. In the Arrow–Debreu model, contracts are unbreakable, but it might be more plausible to think that our newly rich agent can default on her obligations, or at least renegotiate a better deal. In the following, we will explicitly study a variant of Arrow–Debreu equilibrium in which default is allowed.

Suppose there is no commitment, i.e. you cannot sign binding contracts at date -1. Then it may be that at history h^t, the lucky types will default on their nonbinding promise from time -1 to pay σ to the unlucky types. Whether or not this is optimal depends on the future consequences of defaulting today. This depends entirely on the response of the other type of agent. We will suppose throughout that agents play grim trigger: if anyone ever defaults, everyone keeps defaulting forever. (This should provide us with a lower bound on the amount of insurance in equilibrium.) We can have the extremes of full insurance or no insurance, but there are also intermediate cases in which equilibrium insurance is partial. (The Pareto optimum still involves full insurance, of course.)

A lot technical preliminaries are required to make progress. For shorthand, write $s(1) := 1/2 + \sigma$ and $s(2) := 1/2 - \sigma$. We assume that $\{s_t\}$ is a Markov process:

$$\mathbf{P}\left(s_{t+1} = s(1)\middle|(h^{t-1}, s_t = s(1))\right) = \mathbf{P}\left(s_{t+1} = s(1)|s_t = 1/2 + \sigma\right) = \alpha$$

$$\mathbf{P}\left(s_{t+1} = s(2)\middle|(h^{t-1}, s_t = s(2))\right) = \mathbf{P}\left(s_{t+1} = s(2)|s_t = 1/2 + \sigma\right) = \delta.$$

(α and δ are just shorthand.) Then

$$\begin{pmatrix} \mathbf{E}(s_{t+1}|s_t = s(1)) \\ \mathbf{E}(s_{t+1}|s_t = s(2)) \end{pmatrix} = \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \delta & \delta \end{pmatrix} \begin{pmatrix} s(1) \\ s(2) \end{pmatrix} = \pi s$$

where π the the Markov transition matrix and $s := (s(1), s(2))^\top$. Observe

that by the law of iterated expectations and the Markov property,

$$
\begin{pmatrix} \mathbf{E}(s_{t+2}|s_t = s(1)) \\ \mathbf{E}(s_{t+2}|s_t = s(2)) \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{E}(s_{t+2}|s_{t+1} = s(1), s_t = s(1)) \\ (1-\delta)\mathbf{E}(s_{t+2}|s_{t+1} = s(1), s_t = s(2)) \end{pmatrix}
$$
$$
+ \begin{pmatrix} (1-\alpha)\mathbf{E}(s_{t+2}|s_{t+1} = s(2), s_t = s(1)) \\ \delta\mathbf{E}(s_{t+2}|s_{t+1} = s(2), s_t = s(2)) \end{pmatrix}
$$
$$
= \begin{pmatrix} \alpha\mathbf{E}(s_{t+2}|s_{t+1} = s(1)) \\ (1-\delta)\mathbf{E}(s_{t+2}|s_{t+1} = s(1)) \end{pmatrix}
$$
$$
+ \begin{pmatrix} (1-\alpha)\mathbf{E}(s_{t+2}|s_{t+1} = s(2)) \\ \delta\mathbf{E}(s_{t+2}|s_{t+1} = s(2)) \end{pmatrix}
$$
$$
= \begin{pmatrix} \alpha & 1-\alpha \\ 1-\delta & \delta \end{pmatrix} \begin{pmatrix} \mathbf{E}(s_{t+2}|s_{t+1} = s(1)) \\ \mathbf{E}(s_{t+2}|s_{t+1} = s(2)) \end{pmatrix}
$$
$$
= \pi \begin{pmatrix} \mathbf{E}(s_{t+2}|s_{t+1} = s(1)) \\ \mathbf{E}(s_{t+2}|s_{t+1} = s(2)) \end{pmatrix} = \pi^2 s.
$$

By continuing inductively, we obtain

$$
\begin{pmatrix} \mathbf{E}(s_{t+j}|s_t = s(1)) \\ \mathbf{E}(s_{t+j}|s_t = s(2)) \end{pmatrix} = \pi^j s \quad \text{for any } j \in \mathbf{N}.
$$

Let P be a matrix of right-eigenvectors of π, i.e. a matrix such that $\pi P = P\Lambda$ where Λ is the diagonal matrix of eigenvalues. Denote the columns of P by P_1 and P_2. Since π is row-stochastic, one of its eigenvalues is $\lambda_1 = 1$, and $P_1 = (1,1)^\top$ is a corresponding right-eigenvector. It can be shown that the other eigenvalue λ_2 lies strictly inside the unit circle.

Another feature is that $\pi^j = P\Lambda^j P^{-1}$. So as $j \to \infty$,

$$
\pi^j = \begin{pmatrix} P_1 & P_2 \end{pmatrix} \begin{pmatrix} 1^j & 0 \\ 0 & \lambda_2^j \end{pmatrix} \begin{pmatrix} (P^{-1})_1 \\ (P^{-1})_2 \end{pmatrix}
$$
$$
\to \begin{pmatrix} P_1 & P_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (P^{-1})_1 \\ (P^{-1})_2 \end{pmatrix}
$$
$$
= P_1 \left(P^{-1}\right)_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \left(P^{-1}\right)_1 = \begin{pmatrix} (P^{-1})_1 \\ (P^{-1})_1 \end{pmatrix}.
$$

Hence

$$
\begin{pmatrix} \mathbf{E}(s_{t+j}|s_t = s(1)) \\ \mathbf{E}(s_{t+j}|s_t = s(2)) \end{pmatrix} = \pi^j s \to \begin{pmatrix} (P^{-1})_1 \\ (P^{-1})_1 \end{pmatrix} s.
$$

So the process is ergodic: regardless of the initial state s_t, the expected value eventually converges to the same number, viz. $(P^{-1})_1 s$.

Since π has all eigenvalues on or inside the unit circle, $\beta\pi$ has all eigenvalues strictly inside the unit circle, so

$$\sum_{j=0}^{\infty} \beta^j \pi^j = (I - \beta\pi)^{-1}.$$

We can prove this using the Jordan decomposition:

$$\begin{aligned}
\sum_{j=0}^{\infty} \beta^j \pi^j &= P \left(\sum_{j=0}^{\infty} \beta^j \Lambda^j \right) P^{-1} \\
&= P \left(I - \beta\Lambda \right)^{-1} P^{-1} \\
&= \left[P \left(I - \beta\Lambda \right) P^{-1} \right]^{-1} \\
&= \left(I - \beta\pi \right)^{-1}.
\end{aligned}$$

We can now compute off-equilibrium payoffs under when grim trigger strategies are played. Let V^h and V^l denote the expected value of autarky forever, starting with high and low endowment, respectively. We compute

$$\begin{pmatrix} V^h \\ V^l \end{pmatrix} = \sum_{j=0}^{\infty} \left[\beta^j \pi^j \begin{pmatrix} u(1/2 + \sigma) \\ u(1/2 - \sigma) \end{pmatrix} \right] = (I - \beta\pi)^{-1} \begin{pmatrix} u(1/2 + \sigma) \\ u(1/2 - \sigma) \end{pmatrix}.$$

Now look at the constrained efficient allocation problem. We'll take the distribution of s_0 to equal the stationary distribution, so that the endowment process is stationary. We maximise

$$\sum_{t=0}^{\infty} \beta^t \sum_{h^t \in S^t} \mu(h^t) \left[\theta_1 u \left(c_1(h^t) \right) + \theta_2 u \left(c_2(h^t) \right) \right],$$

where varying θ_1 and θ_2 allows us to trace out all the constrained Pareto optima since u is concave. The maximisation is subject to the node-by-node incentive-compatibility constraint that high-endowment people at node h^t don't renege:

$$u(1/2 + \sigma) + \sum_{j=1}^{\infty} \sum_{h^{t+j} \in S^{t+j}|h^t} \beta^j \frac{\mu(h^{t+j})}{\mu(h^t)} u \left(c_1(h^{t+j}) \right) \geq V^h.$$

We have a simliar no-reneging constraint for low-endowment people in terms of $u(1/2 - \sigma)$ and V^l. Denote the multipliers on the two constraints by $\lambda^h(h^t)$ and $\lambda^l(h^t)$.

The solutions turn out to depend on the specifics. In general the constrained optimum has partial insurance. To show this, compute the first-order condition w.r.t. $c_1(h^t)$, letting $s_t = s(1)$ without loss of generality:

$$0 = u'(c_1(h^t)) \left[\beta^t \mu(h^t) \theta_1 \right.$$

$$+ \beta^t \mu(h^t) \lambda^h(h^t) + \beta^{t-1} \mu(h^{t-1}) \lambda^h(h^{t-1}) \beta \frac{\mu(h^t)}{\mu(h^{t-1})}$$

$$\left. + \cdots + \beta^0 \mu(h^0) \lambda^h(h^0) \beta^t \frac{\mu(h^t)}{\mu(h^0)} \right]$$

$$= \beta^t \mu(h^t) u'\left(c_1(h^t) \right) \left[\theta_1 + \sum_{j=0}^{t} \lambda^1(h^j) \right].$$

The same holds for type 2. So efficiency requires that at all nodes h^t,

$$u'\left(c_1(h^t) \right) \left[\theta_1 + \sum_{j=0}^{t} \lambda^h(h^j) \right] = u'\left(1 - c_1(h^t) \right) \left[\theta_2 + \sum_{j=0}^{t} \lambda^l(h^j) \right]$$

I cannot for the life of me understand why we carried λ^l all this way since in equilibrium the no-reneging constraint does not bind for the low-endowment types, so that $\lambda^l = 0$. However, $\lambda^h > 0$ in general. So

$$\left[\theta_1 + \sum_{j=0}^{t} \lambda^h(h^j) \right] u'\left(c_1(h^t) \right) = \theta_2 u'\left(1 - c_1(h^t) \right)$$

holds at all nodes $h^t = (h^{t-1}, s(1))$ at which type 1 has the high endowment. Observe that having the high endowment today effectively increases type 1's Pareto weight, this yielding a higher amount of consumption for type 1 that we had in the unconstrainted Pareto optimum. This is because the planner needs to bribe the high-endowment type to convince her to not renege.

Observe further that the degree to which the high-endowment type's Pareto weight is nudged up depends on how many times in the past she has had the high endowment. The more times she has had it, the more she needs to be bribed. I think this is because the planner commits in advance, so a household with high lifetime wealth needs more bribing.

This means that how much consumption an agent gets at t depends on h^t, so insurance is incomplete. Further still, the degree of incompleteness is time-varying in general.

10   Mon 26 Oct 2015

10.1   Outline of today

Today we're starting growth theory. This material will not be coming from either Stokey et al. (1989) or Ljungqvist and Sargent (2012). Sources will be mentioned along the way, and many can be found on the syllabus. We'll first introduce the topic, then outline the models that we'll be covering. Finally we'll get started on the 'Solow I' or disembodied exogenous technical change model.

10.2   Growth theory

Growth theory is motivated by a puzzle. From about 50000BC to about 1750AC, the world was very poor and didn't get much richer. Then output skyrocketed (initially in Holland).

    We can refine the puzzle somewhat using the neoclassical 'growth' model. The only way to grow in that model is to accumulate capital. But first of all, we know that the capital stock can never grow beyond $\max\{k^\star, k_0\}$ no matter what savings policy is pursued, so perpetual growth is impossible. Moreover, equilibrium choices of consumption and capital satisfy the Euler equation

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

where $f'(k_{t+1})$ is the one-period rate of return on capital from period t to period $t + 1$ (marginal product of capital plus $1 - \delta$). As the capital stock grows, the rate of return $f'(k_{t+1})$ declines since f is strictly concave. As a result, the incentive to save diminishes over time, so we approach a steady-state level of capital $k^\star$. There only growth in this model comes from these transition dynamics; long-run growth is impossible. So the refined version of the growth puzzle is: why does capital accumulation persist?

    A related question is to what extent decentralised markets promote or hinder growth. The anecdotal evidence on this is mixed: western Europe's growth in the 18th and 19th centuries took place in the presence of fairly decentralised markets, whereas growth in Asian countries post-WWII occurred under a great deal of central planning.

    There's a related political economy issue here. Change creates winners and losers, winners acquire political power, and then have an incentive to maintain the status quo to retain their wealth and influence, which causes stagnation. We might conjecture that this will lead to endogeneous cycles of

growth and stagnation. Krusell and Ríos-Rull (1996) formalise this argument; Parente and Prescott (2000) is similar.

10.3   Overview of growth models

The basic models are exogenous growth models. This means that the economy's resource constraint expands exogenously over time. There are two variants: disembodied technical change and embodied technical change. In the former, the technology for turning input goods into output (consumption and capital goods) improves over time. In the latter, the technology for turning consumption goods (investment) into new capital improves over time. Larry calls the former 'Solow I' and the latter 'Solow II'.[26] Solow I is the discrete-time version of the Ramsey-Cass-Koopmans model (Ramsey, 1928; Cass, 1965; Koopmans, 1965); it is not clear to my why we should name it after Solow. (This is *not* the Solow (1956) model, where the saving policy is exogenous.) The Solow II model appears to have originated with Solow (1960).

The more interesting models exhibit endogenous growth. This means that the primitives of the model are fixed, so that growth is purely the result of equilibrium interactions. We'll start out with the *Ak* model, which lays bare the main issues. The next few models we'll look at will all turn out to be *Ak* models in disguise: these include a multi-sector model, the human capital model and the Romer (1987) specialisation model.

We'll pay special attention to the Romer (1987) model. Here specialisation in the production of intermediate goods drives growth. However, specialisation is costly, so only happens gradually. This provides one account of why endogenous growth should be incremental.

Another brace of models is concerned with the discovery and implementation of new ideas and how these factors drive growth. (For comparison, in Romer (1987), new ideas are out there the entire time. They are not immediate implemented only because implementation is costly.) In particular, firms discover new ideas by sinking resources into search (a bandit), and implement the idea they discover via costly investment. One paper along these lines is Comin and Gertler (2006), which we will spend a lot of time on.

We'll also talk about an overlapping-generations models of growth. When consumers have finite horizons, things change quite a lot. (Unless you do a Barro (1974)!) One difference is that in equilibrium, the return on capital may grow faster than wages. If capital holdings are heterogeneous across

---

[26]He said that these terms come from Greenwood, Hercowitz and Krusell (1997), but I couldn't find them there.

households and inherited preferentially, this will lead to a concentration of wealth in the hands of the owners of capital. (Piketty happy.)

For each model, we'll ask what the equilibrium growth rate is, both of output, consumption and the capital stock. We'll also ask about whether there is convergence in the following sense: if two countries start with different initial conditions, do their levels (of output, consumption, and capital) get closer to each other over time? Both of the exogenous growth models we'll look at exhibit this form of long-term path-independence, but endogenous growth models generally do not.

10.4   The Solow I model

In the neoclassical growth model, the technology f that produces output from capital services is fixed and satisfies satisfies the Inada condition

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

As we've seen, this leads to transition dynamics on the way to the interior steady state, but no long-run growth.

To allow for long-run growth, the Solow I model has the technology evolve deterministically over time. In the Cobb–Douglas case, we take

$$f_t(k) = z_t^{1-\alpha} k^\alpha + (1 - \delta)k$$

where total factor productivity z_t evolves according to

$$z_t = z_{t-1} \exp(\mu_z), \quad \mu_z > 0.$$

The drift in z_t constitutes technological change: as time passes, more output can be produced from a given quantity of capital services. Rather than micro-founding this law of motion, we treat it as given; hence technical change is exogenous here.

We assume log utility for simplicity.[27] We then have the optimal growth

---

[27]Balanced growth will require some restrictions on u, but there are functional forms other than log that will do fine.

problem

$$\max_{\{k_{t+1},c_t,i_t\}} \sum_{t=0}^{\infty} \beta^t \ln(c_t)$$
$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad c_t + i_t \leq k_t^\alpha z_t^{1-\alpha}$$
$$k_{t+1} = (1-\delta)k_t + i_t$$
$$z_{t+1} = \exp(\mu_z)z_t$$
$$k_{t+1}, c_t, i_t \geq 0$$

where k_0 and z_0 are given. As written, this problem does not have a recursive representation because the constraints evolve over time. (Moreover, we don't have e.g. boundedness.) But as we will show, we can transform this model to put it back into Stokey–Lucas form; in fact, we can transform it into the neoclassical growth model, in which there is no technological change. This trick generalises to some extent: any model written so as to allow for balanced growth can be 'rescaled' by the exogenous time-varying components (z_t here) so that it becomes stationary. In this case, the scaling will be simple, but in other models the correct scaling may be tricky to find.

In this case, the correct transformation is to scale all time-t choice variables by z_t as follows:

$$\widetilde{k}_{t+1} := \frac{k_{t+1}}{z_t}, \quad \widetilde{c}_t := \frac{c_t}{z_t} \quad \text{and} \quad \widetilde{i}_t := \frac{i_t}{z_t}.$$

The general rule for scalings is that we want the exogenously time-varying components (z_t here) to vanish from the objective and constraints of the problem. If this works, then we've got a stationary model in Stokey–Lucas form that is amenable to recursive methods. In the context of the neoclassical growth model, the necessary and sufficient conditions on f and u for the existence of such a scaling are the 'balanced growth conditions' of King, Plosser and Rebelo (1988).

In the reparametrised model, utility becomes

$$\sum_{t=0}^{\infty} \beta^t \ln(c_t) = \sum_{t=0}^{\infty} \beta^t \ln(\widetilde{c}_t) + \sum_{t=0}^{\infty} \beta^t \ln(z_t).$$

We don't actually need the second term to be finite, only well-defined. (If it is infinite, we can use the overtaking criterion.) But it's nice if it's finite, and it is: $\ln(z_t) = \ln(z_0) + t\mu_z$, so

$$\sum_{t=0}^{\infty} \beta^t \ln(z_t) = \sum_{t=0}^{\infty} \beta^t \left(\ln(z_0) + \mu_z t \right) = \frac{\ln(z_0)}{1-\beta} + \mu_z \sum_{t=0}^{\infty} \beta^t t.$$

Differentiating both sides of $\sum_{t=0}^{\infty} \beta^t = \frac{1}{1-\beta}$ yields

$$\beta^{-1} \sum_{t=0}^{\infty} t\beta^t = \frac{1}{(1-\beta)^2}.$$

We therefore have

$$\sum_{t=0}^{\infty} \beta^t \ln(z_t) = \frac{\ln(z_0)}{1-\beta} + \mu_z \frac{\beta}{(1-\beta)^2} < \infty.$$

So utility is $\sum_{t=0}^{\infty} \beta^t \ln(\widetilde{c}_t)$ plus a constant. We can obviously ignore this constant.

The resource constraint becomes

$$\widetilde{c}_t + \widetilde{i}_t \leq \left(\frac{k_t}{z_t}\right)^\alpha = \left(\frac{z_{t-1}}{z_t}\right)^\alpha \left(\frac{k_t}{z_{t-1}}\right)^\alpha = \exp(-\alpha\mu_z)\widetilde{k}_t^\alpha.$$

The law of motion becomes

$$\widetilde{k}_{t+1} = (1-\delta)\frac{k_t}{z_{t-1}}\frac{z_{t-1}}{z_t} + \widetilde{i}_t = (1-\delta)\exp(-\mu_z)\widetilde{k}_t + \widetilde{i}_t = \left(1 - \widetilde{\delta}\right)\widetilde{k}_t + \widetilde{i}_t$$

where $\widetilde{\delta} \in (0,1)$ is defined by

$$1 - \widetilde{\delta} = (1-\delta)\exp(-\mu_z).$$

We will solve the original problem by solving the scaled problem, then un-scaling the solution. The scaled problem is

$$\max_{\{\widetilde{k}_{t+1},\widetilde{c}_t,\widetilde{i}_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln(\widetilde{c}_t)$$

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad \widetilde{c}_t + \widetilde{i}_t \leq \exp(-\alpha\mu_z)\widetilde{k}_t^\alpha$$

$$\widetilde{k}_{t+1} = \left(1 - \widetilde{\delta}\right)\widetilde{k}_t + \widetilde{i}_t$$

$$\widetilde{k}_{t+1}, \widetilde{c}_t, \widetilde{i}_t \geq 0$$

where $\widetilde{k}_0 = k_0/z_{-1}$ is given. This is exactly the neoclassical growth model, so we know that there's a unique solution with lots of nice properties. One of these is that $\widetilde{k}_{t+1}$ converges monotonically to a steady state $\widetilde{k}^\star$ that satisfies

$$\alpha \exp(-\alpha\mu_z) \left(\widetilde{k}^\star\right)^{\alpha-1} + 1 - \widetilde{\delta} = \beta^{-1},$$

which rearranges to

$$\widetilde{k}^\star = \left(\frac{\alpha \exp(-\alpha\mu_z)}{\beta^{-1} - (1-\delta)\exp(-\mu_z)} \right)^{1/(1-\alpha)}.$$

So on the balanced growth path (the path of the unscaled model when the scaled model is in steady state), the capital stock evolves as

$$
\begin{aligned}
k_{t+1} &= \left(\frac{\alpha \exp(-\alpha\mu_z)}{\beta^{-1} - (1-\delta)\exp(-\mu_z)} \right)^{1/(1-\alpha)} z_t \\
&= \left(\frac{\alpha \exp(-\alpha\mu_z)}{\beta^{-1} - (1-\delta)\exp(-\mu_z)} \right)^{1/(1-\alpha)} z_0 \cdot \exp(\mu_z t).
\end{aligned}
$$

So the growth rate of the capital stock on the balanced growth path is μ_z. Off the balanced growth path, the path of the capital stock converges monotonically to the balanced growth path. It is not just that the growth rate converges to μ_z: every path converges to the balanced growth path in levels! So this model exhibits convergence.

Without the balanced-growth conditions, the problem becomes non-stationary. The shooting algorithm will still work, but we cannot use recursive methods. This loss of tractability is why we will not look at models without a balanced growth path. It may be that nonstationary models could provide insights that stationary models cannot!

Solow I is the workhorse growth model.[28] For example, Young (1992) applies it to study growth in Hong Kong and Singapore. He finds that while growth in one of them was driven by increasing total factor productivity (the kind of growth that the Solow I model can explain), growth in the other was not. Another example is Cheremukhin, Golosov, Guriev and Tsyvinski (2013), who study 'Stalinist growth', meaning that capital accumulation and hence growth is faster than optimal.

10.5   The Solow II model

The paper that became Solow (1960) was presented at a conference a little before 1960 with people like Ken Arrow in attendance. All present agreed that the new 'Solow II' model of embodied technical change was far more sensible than the disemobided technical change of 'Solow I'.

They considered Solow I unsatisfactory because in that model, the technology for transforming capital services into any kind of output, be it haircuts

---

[28] And as we will see in section 17, Solow I with a stochastic law of motion for z_t is the workhorse (real) business cycle model.

or laptops, grows at the same rate, which clearly isn't right! Rather, we think that economic growth has been driven by technological change in the production of certain goods (laptops), but not others (haircuts).

The new model featured embodied technical change and embodied investment.[29] There are several ways of modelling this, but the Solow II way is as follows. Unlike before, the technology for turning capital services into goods does not evolve over time, so the resource constraint remains $c_t + i_t \leq k_t^\alpha$ throughout. Instead, technical progress involves improvements in our ability to turn investment into productive capital goods. We capture this by letting the law of motion for capital be

$$k_{t+1} = (1 - \delta)k_t + q_t i_t,$$

where the 'investment efficiency' q_t evolves according to

$$q_{t+1} = \exp(\mu_q)q_t.$$

The equations suggest that we get a larger quantity of capital per dollar of investment as q_t grows. But it is perhaps more appealing to think of k_t as measuring capital in efficiency units; then we can think of growth in q_t as increasing the *quality* of the capital that can be obtained per dollar of investment. Either way, economic growth is not driven by improvements the technology for turning capital services into output.[30]

The optimal growth problem is then

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

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad c_t + i_t \leq k_t^\alpha$$
$$k_{t+1} = (1 - \delta)k_t + q_t i_t$$
$$q_{t+1} = \exp(\mu_q)q_t$$
$$c_t, i_t \geq 0.$$

As usual, we prefer to use k_{t+1} rather than i_t as a choice variable. Reformu-

---

[29]I still have no idea what 'embodied' and 'disembodied' are supposed to mean.

[30]Larry said that Basu and Fernald (1997) is relevant here. But he did not say why, and it's not obvious from looking at the paper.

lating the problem in this way yields

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

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad c_t + q_t^{-1} [k_{t+1} - (1-\delta)k_t] \leq k_t^{\alpha}$$

$$q_{t+1} = \exp(\mu_q) q_t$$

$$k_{t+1} \geq (1-\delta)k_t, \quad c_t \geq 0.$$

Ignoring the nonnegativity constraints (by appeal to the Inada conditions), the Lagrangean is

$$\sum_{t=0}^{\infty} \beta^t \{u(c_t) + \lambda_t (k_t^{\alpha} - c_t - q_t^{-1} [k_{t+1} - (1-\delta)k_t])\}.$$

The first-order conditions with respect to c_t and k_{t+1} are

$$u'(c_t) = \lambda_t$$

$$\beta^t \lambda_t q_t^{-1} = \beta^{t+1} \lambda_{t+1} \left[\alpha k_{t+1}^{\alpha-1} + q_{t+1}^{-1}(1-\delta) \right].$$

Combining them yields the Euler equation

$$u'(c_t) = \beta u'(c_{t+1}) \underbrace{\left(\frac{\alpha k_{t+1}^{\alpha-1} + (1-\delta)q_{t+1}^{-1}}{q_t^{-1}} \right)}_{=:R_t^k}.$$

R_t^k is the rate of return to holding capital from period t to period $t + 1$, denominated in units of period-t consumption. To see this, reason as follows. To purchase an extra unit of capital at date t, the household must give up q_t^{-1} units of consumption. In so doing, the household earns a return of $\alpha k_{t+1}^{\alpha-1}$ from renting the extra capital out to firms. After this, she is left with $(1 - \delta)$ units of extra capital, which she can convert back to $(1 - \delta)q_{t+1}^{-1}$ units of consumption.

Long-run growth in output will require long-run growth in the capital stock. From the Euler equation, it's clear that the capital stock can continue to grow in perpetuity only if the rate of return on capital remains high.

83

11   Wed 28 Oct 2015

11.1   Outline of today

We first study the Solow II model that we looked at briefly last time. We'll
scale it (twice) to get it into recursive form, then solve. Our focus will be
on the growth rate of the capital stock, consumption and the rate of return
on capital. We'll subsequently look at other properties of the solutions and
their implications. One important property is the convergence (in levels)
phenomenon that we saw in Solow I.

   We'll then move on to endogenous growth, starting with the Ak model.
We'll use a trick to write it in recursive form, solve for the efficient allocations,
and examine their properties. We will see that unlike the exogenous growth
models, the Ak model does not display convergence.

11.2   The Solow II model again

The optimal growth problem in the Solow II model (specialised to log utility)
is

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

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad c_t + q_t^{-1}\left[k_{t+1} - (1-\delta)k_t\right] \le k_t^{\alpha}$$

$$q_{t+1} = \exp(\mu_q)q_t$$

$$k_{t+1} \ge (1-\delta)k_t, \quad c_t \ge 0.$$

There is no TFP growth in this model. Nevertheless, we will have long-run
growth in k_t driven by the growth of investment efficiency q_t.

   To obtain a recursive formulation, we have to scale the model. Write
$\bar{k}_{t+1} := k_{t+1}/q_t$. Then the resource constraint becomes

$$c_t + \bar{k}_{t+1} - (1-\delta)\exp(-\mu_q)\bar{k}_t \le q_{t-1}^{\alpha}\bar{k}_t^{\alpha}.$$

Now define $z_t := q_{t-1}^{\alpha/(1-\alpha)}$, which evolves according to

$$z_t = q_{t-1}^{\alpha/(1-\alpha)} = (\exp(\mu_q)q_{t-2})^{\alpha/(1-\alpha)}$$

$$= \exp\left(\tfrac{\alpha}{1-\alpha}\mu_q\right)q_{t-2}^{\alpha/(1-\alpha)} = \exp\left(\tfrac{\alpha}{1-\alpha}\mu_q\right)z_{t-1} = \exp(\bar{\mu}_z)z_{t-1}$$

where $\bar{\mu}_z := \tfrac{\alpha}{1-\alpha}\mu_q$. The resource constraint then becomes

$$c_t + \bar{k}_{t+1} - \left(1-\bar{\delta}\right)\bar{k}_t \le z_t^{1-\alpha}\bar{k}_t^{\alpha}$$

84

where $\bar{\delta}$ is defined by $1 - \bar{\delta} = (1 - \delta)\exp(-\mu_q)$. With this scaling, we obtain the optimal growth problem

$$\max_{\{\bar{k}_{t+1}, c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln(c_t)$$

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad c_t + \bar{k}_{t+1} - \left(1 - \bar{\delta}\right)\bar{k}_t \leq z_t^{1-\alpha}\bar{k}_t^{\alpha}$$

$$z_{t+1} = \exp\left(\bar{\mu}_z\right)z_t$$

$$\bar{k}_{t+1} \geq \left(1 - \bar{\delta}\right)\bar{k}_t, \quad c_t \geq 0.$$

This program is precisely the optimal growth problem in the Solow I model! So to obtain a recursive formulation, define $\widetilde{k}_{t+1} := \bar{k}_{t+1}/z_t$ and $\widetilde{c}_t := c_t/z_t$ and rewrite the program as

$$\max_{\{\widetilde{k}_{t+1}, \widetilde{c}_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln\left(\widetilde{c}_t\right)$$

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad \widetilde{c}_t + \widetilde{k}_{t+1} - \left(1 - \widetilde{\delta}\right)\widetilde{k}_t \leq \widetilde{k}_t^{\alpha}$$

$$\widetilde{k}_{t+1} \geq \left(1 - \widetilde{\delta}\right)\widetilde{k}_t, \quad \widetilde{c}_t \geq 0,$$

where $\widetilde{\delta}$ is defined by

$$1 - \widetilde{\delta} = \left(1 - \bar{\delta}\right)\exp(-\bar{\mu}_z) = (1 - \delta)\exp(-\mu_q)\exp\left(-\tfrac{\alpha}{1-\alpha}\mu_q\right)$$

$$= (1 - \delta)\exp\left(-\tfrac{1}{1-\alpha}\mu_q\right).$$

This is the sequence problem for the neoclassical growth model, so we know that $\widetilde{k}_{t+1}$ converges monotonically to a steady state $\widetilde{k}^{\star} > 0$ (for which we have a formula). To convert back to k_{t+1},

$$k_{t+1} = q_t\bar{k}_{t+1}$$

$$= q_t z_t \widetilde{k}_{t+1}$$

$$= [\exp(\mu_q)q_{t-1}]\, q_{t-1}^{\alpha/(1-\alpha)}\widetilde{k}_{t+1}$$

$$= \exp(\mu_q)q_{t-1}^{1/(1-\alpha)}\widetilde{k}_{t+1}$$

$$= \exp(\mu_q)\left[\exp(\mu_q(t-1))q_0\right]^{1/(1-\alpha)}\widetilde{k}_{t+1}$$

$$= \exp(\mu_q)\exp\left(\tfrac{1}{1-\alpha}\mu_q(t-1)\right)q_0^{1/(1-\alpha)}\widetilde{k}_{t+1}$$

$$= \exp(\mu_q)\exp\left(\tfrac{1}{1-\alpha}\mu_q(t-1)\right)q_0^{1/(1-\alpha)}\widetilde{k}_{t+1}$$

$$= \exp\left(-\tfrac{\alpha}{1-\alpha}\mu_q\right)q_0^{1/(1-\alpha)}\cdot\exp\left(\tfrac{1}{1-\alpha}\mu_q t\right)\cdot\widetilde{k}_{t+1}.$$

85

Since $\widetilde{k}_{t+1}$ converges monotonically to $\widetilde{k}^\star$, k_{t+1} converges monotonically to the balanced growth path regardless of the initial condition. So this model displays convergence. The long-run growth rate of the capital stock is $\mu_q/(1-\alpha)$.

Consumption is

$$
\begin{aligned}
c_t &= z_t \widetilde{c}_t \\
&= q_{t-1}^{\alpha/(1-\alpha)} \widetilde{c}_t \\
&= [\exp(\mu_q(t-1))q_0]^{\alpha/(1-\alpha)} \widetilde{c}_t \\
&= \exp\left(\tfrac{\alpha}{1-\alpha}\mu_q(t-1)\right) q_0^{\alpha/(1-\alpha)} \widetilde{c}_t \\
&= \exp\left(-\tfrac{\alpha}{1-\alpha}\mu_q\right) q_0^{\alpha/(1-\alpha)} \cdot \exp\left(\tfrac{\alpha}{1-\alpha}\mu_q t\right) \cdot \widetilde{c}_t.
\end{aligned}
$$

Again, consumption converges to the balanced growth path. On the balanced growth path, it grows at rate $\alpha \cdot \mu_q/(1-\alpha)$, a slower rate than the capital stock. However, it does not really make sense to compare the growth rates of consumption and capital in this model. Whereas in Solow I, we could convert the output good into capital one-for-one, the marginal rate of transformation is changing over time in Solow II.

To rectify this, we introduce the price of capital p_t^k, which is the number of units of consumption required to buy an extra unit of capital. In the Solow II model, $p_t^k = q_t^{-1}$ is a purely technological object. We call it the price of capital because if took the technology for converting consumption goods into capital away from the household and instead let a competitive firm operate it, the firm would set the price p_t^k of capital equal to its marginal cost q_t^{-1}.[31]

To compare consumption and capital in the same units, we adjust by p_t^k to account for the fact that consumption and capital cannot be exchanged one-for-one. In the interpretation of the model in which growth in q_t reflects improvements in the quality rather than the quantity of capital produced, we can think of this adjustment as taking proper account of improving quality

---

[31]The term 'price of capital' is not used by very many people other than Larry, it seems. When it is used, people usually say call it the *relative* price of capital to emphasise that it is the price of capital in units of the consumption good (whose price need not be normalised to unity as here).

of the capital stock. So compute

$$
\begin{aligned}
p_t^k k_{t+1} &= q_t^{-1} k_{t+1} \\
&= \exp(-\mu_q t) q_0^{-1} k_{t+1} \\
&= \exp(-\mu_q t) q_0^{-1} \cdot \exp\left(-\tfrac{\alpha}{1-\alpha}\mu_q\right) q_0^{1/(1-\alpha)} \cdot \exp\left(\tfrac{1}{1-\alpha}\mu_q t\right) \cdot \widetilde{k}_{t+1} \\
&= \exp\left(-\tfrac{\alpha}{1-\alpha}\mu_q\right) q_0^{\alpha/(1-\alpha)} \cdot \exp\left(\tfrac{\alpha}{1-\alpha}\mu_q t\right) \cdot \widetilde{k}_{t+1}.
\end{aligned}
$$

So the long-run growth rate of capital in units of consumption is $\alpha \cdot \mu_q/(1-\alpha)$, the same as the growth rate of consumption.

In the previous lecture, we introduced the rate of return on capital as

$$
R_t^k = \frac{\alpha k_{t+1}^{\alpha-1} + (1-\delta) q_{t+1}^{-1}}{q_t^{-1}} = \frac{\alpha k_{t+1}^{\alpha-1} + (1-\delta) p_{t+1}^k}{p_t^k}.
$$

Rewriting a little yields

$$
\begin{aligned}
R_t^k &= \alpha \left(q_t^{1/(1-\alpha)} k_{t+1}^{-1} \right)^{1-\alpha} + (1-\delta) \frac{q_t}{q_{t+1}} \\
&= \alpha \left(q_t^{1/(1-\alpha)} k_{t+1}^{-1} \right)^{1-\alpha} + (1-\delta) \exp(-\mu_q).
\end{aligned}
$$

The growth rate of $q_t^{1/(1-\alpha)}$ is $\mu_q/(1-\alpha)$, and we've shown that k_{t+1} grows at this same rate on the balanced growth path. So on the balanced growth path R_t^k is a positive constant, maintaining the incentive to accumulate capital even in the long run.

Let's return to the intellectual history of this model. If Solow II was thought so much more sensible than Solow I when first introduced, why did Solow II not become the workhorse growth model? The answer lies with the fact that p_t^k declines steadily over time in Solow II, whereas it is constant in Solow I. When these models were first introduced, the empirical evidence seemed to suggest that p_t^k was in fact roughly constant, appearing to refute embodied technical change.

Many years later, Gordon (1990) argued convincingly that p_t^k has in fact been falling steadily over most of history, consistent with embodied technical change and inconsistent with growth driven by TFP. The problem with the earlier data was that it did not adequately take into account quality improvements in the capital stock. (In both Solow I and Solow II, the capital stock is measured in efficiency units, so the same number of higher-quality machines means a larger capital stock.) Greenwood et al. (1997) publicised this result and extended it. They authors showed that p_t^k has in fact been

steadily declining for a very long time, supporting the hypothesis that the main engine of growth is embodied rather than disembodied technical change.[32]

Another point in favour of Solow I over Solow II put forward by Greenwood et al. (1997) is its quantitative implications for growth. Using the plausible calibration $\mu_q \simeq 0.03$ and $\alpha \simeq 1/3$ the Solow II model predicts a growth rate in consumption of 1.5%, broadly consistent with US data for the last 100 years. On the other hand, the modern data suggests that TFP growth has been close to zero over the same period, in which case the Solow I model predicts almost no growth.

Others have argued that q_t drives not just growth but also the business cycle; see e.g. Greenwood, Hercowitz and Huffman (1988). Rather than a stochastic variant of Solow I (i.e. the RBC model), these authors propose that we ought to use a stochastic version of Solow II as our workhorse business cycle model. Instead of business cycles driven by TFP shocks, we get business cycles driven by shocks to investment efficiency. Many papers these days include both, and find that investment efficiency shocks explain a substantial fraction of variation in output, unemployment etc. at business cycle frequency. Another advantage of this formulation is that is able to explain why investment is so volatile (and procyclical): in a stochastic Solow II model, booms are driven by investment, which is in turn driven by shocks to investment efficiency.[33]

11.3   The Ak model

We now move on to endogenous growth models. These models try to explain growth in a 'deeper' way than exogenous growth models, since they generate growth in endogenous variables without recourse to growth in some exogenous variable. The simplest endogenous growth model is the Ak model, in which we simply postulate that the marginal product of capital is constant (and equal to A, hence the name). Then the marginal product of capital does not decline even as the capital stock grows large, so the rate of return remains high, maintaining the incentive to save.

---

[32]Larry is somewhat skeptical of Greenwood et al. (1997) because the authors only look at business equipment, where quality improvements have indeed been rapid. But other kinds of capital, in particular fixed structures (i.e. buildings), are known to have seen a slower pace of quality improvement.

[33]Christiano, Motto and Rostagno (2014) argue that there's something missing from this argument.

The optimal growth problem is[34]

$$\max_{\{k_{t+1}, c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} c_t^{1-\gamma}$$

$$\text{s.t. } \forall t \in \mathbf{Z}_+ \quad c_t + k_{t+1} - (1-\delta)k_t \leq Ak_t$$

$$k_{t+1}, c_t \geq 0.$$

Since the price of capital is unity, the rate of return is $A + 1 - \delta$. We therefore assume $A > \delta$ so that the rate of return is above unity.

Substituting,

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} \left((A + 1 - \delta)k_t - k_{t+1} \right)^{1-\gamma}$$

$$\text{s.t. } \quad k_{t+1} \in [0, (A + 1 - \delta)k_t] \quad \forall t \in \mathbf{Z}_+.$$

We scale the model using $\lambda_t := k_{t+1}/k_t$. The constraint becomes $\lambda_t \in [0, A + 1 - \delta]$. Using the 'accordion trick', we can write k_{t+1} as

$$k_{t+1} = \lambda_t k_t = \lambda_t \lambda_{t-1} k_{t-1} = \cdots = \left(\prod_{j=0}^{t} \lambda_j \right) k_0.$$

We can therefore rewrite the period return in terms of λ_t as follows:

$$\frac{1}{1-\gamma} \left((A + 1 - \delta)k_t - k_{t+1} \right)^{1-\gamma} = \frac{1}{1-\gamma} \left(A + 1 - \delta - \lambda_t \right)^{1-\gamma} k_t^{1-\gamma}$$

$$= \frac{1}{1-\gamma} \left(A + 1 - \delta - \lambda_t \right)^{1-\gamma} \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma} \right) k_0^{1-\gamma}.$$

The reformulated optimal growth problem is therefore

$$\max_{\{\lambda_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} k_0^{1-\gamma} \left(A + 1 - \delta - \lambda_t \right)^{1-\gamma} \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma} \right)$$

$$\text{s.t. } \quad \lambda_t \in [0, A + 1 - \delta] \quad \forall t \in \mathbf{Z}_+.$$

To avoid having to use the overtaking criterion or similar, we wish to show that the objective is finite. To achieve this, we will require the restriction

---

[34]We don't restrict $k_{t+1} \geq (1 - \delta)k_t$ (nonnegative investment) here, but I don't think it's a problem to do that. We allow for CRRA preferences here (rather than the more restrictive log), but we could have done this for the exogenous growth models, too.

$\beta(A+1-\delta)^{1-\gamma} < 1$. To show that the objective is bounded above, notice that of the two terms involving λ_t, the first is decreasing and the second is increasing. Hence we can bound the objective from above by letting $\lambda_t = 0$ $\forall t$ in the first term and $\lambda_t = A + 1 - \delta$ $\forall t$ in the second term (it doesn't matter that this is infeasible). This yields

$$\sum_{t=0}^{\infty} \beta^t \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A+1-\delta-\lambda_t)^{1-\gamma} \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma} \right)$$

$$\leq \sum_{t=0}^{\infty} \beta^t \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A+1-\delta)^{1-\gamma} \left(\prod_{j=0}^{t-1} (A+1-\delta)^{1-\gamma} \right)$$

$$= \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A+1-\delta)^{1-\gamma} \sum_{t=0}^{\infty} \left[\beta \left(A+1-\delta \right)^{(1-\gamma)} \right]^t$$

$$= \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A+1-\delta)^{1-\gamma} \left[1 - \beta \left(A+1-\delta \right)^{(1-\gamma)} \right]^{-1} < \infty$$

where the final equality used the boundedness condition $\beta(A+1-\delta)^{1-\gamma} < 1$.

To show boundedness below, we use the other restriction on A, viz. $A > \delta$. This implies that $A+1-\delta > 1$, so that $\lambda_t = 1$ $\forall t$ is feasible. We can therefore bound the payoff from below as

$$\sum_{t=0}^{\infty} \beta^t \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A+1-\delta-\lambda_t)^{1-\gamma} \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma} \right)$$

$$\geq \sum_{t=0}^{\infty} \beta^t \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A-\delta)^{1-\gamma} = \tfrac{1}{1-\beta} \tfrac{1}{1-\gamma} k_0^{1-\gamma} (A-\delta)^{1-\gamma} > -\infty.$$

Having proved boundedness of the objective, we're ready to reformulate the problem in recursive form. Begin by writing the time-t value function v_t as a function of k_t in the usual way. Then the time-t functional equation is

$$v_t \left(\{\lambda_j\}_{j=0}^{t-1}, k_0 \right) = \max_{\lambda_t \in [0, A+1-\delta]} \left[\tfrac{1}{1-\gamma} (A+1-\delta-\lambda_t)^{1-\gamma} \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma} \right) k_0^{1-\gamma} \right.$$

$$\left. + \beta v_{t+1} \left(\{\lambda_j\}_{j=0}^{t}, k_0 \right) \right].$$

Now write $w \in \mathbf{R}$ for the time- and history-independent part of v_t:

$$v_t \left(\{\lambda_j\}_{j=0}^{t-1}, k_0 \right) = \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma} \right) k_0^{1-\gamma} w.$$

Then the time-t functional equation becomes

$$\left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma}\right) k_0^{1-\gamma} w$$

$$= \max_{\lambda_t \in [0, A+1-\delta]} \left[\frac{1}{1-\gamma} (A + 1 - \delta - \lambda_t)^{1-\gamma} \left(\prod_{j=0}^{t-1} \lambda_j^{1-\gamma}\right) k_0^{1-\gamma} \right.$$

$$\left. + \beta \left(\prod_{j=0}^{t} \lambda_j^{1-\gamma}\right) k_0^{1-\gamma} w\right].$$

Cancelling terms, we obtain

$$w = \max_{\lambda_t \in [0, A+1-\delta]} \left[\frac{1}{1-\gamma} (A + 1 - \delta - \lambda_t)^{1-\gamma} + \beta \lambda_t^{1-\gamma} w\right].$$

There is no time- or history-dependence left here, so we can drop time subscripts to obtain the time-invariant equation

$$w = \max_{\lambda \in [0, A+1-\delta]} \left[\frac{1}{1-\gamma} (A + 1 - \delta - \lambda)^{1-\gamma} + \beta \lambda^{1-\gamma} w\right]. \qquad (11)$$

Once we've solved this equation for w, we can obtain the optimal policy $\lambda^\star$ as the argmax on the right-hand side. The optimal policy will clearly be time- and history-independent. We can then recover the optimal path of the capital stock as $k_t^\star = (\lambda^\star)^t k_0 \; \forall t$.

To solve for w, define the mapping T on $\mathbf{R}$ by

$$Tf := \max_{\lambda \in [0, A+1-\delta]} \left[\frac{1}{1-\gamma} (A + 1 - \delta - \lambda)^{1-\gamma} + \beta \lambda^{1-\gamma} f\right] \quad \text{for any } f \in \mathbf{R}.$$

w is obviously a fixed point of T. In fact, we cannot appeal the contraction mapping theorem to establish that T has a unique fixed point because T is not a contraction mapping. But by other methods (see homework 2), we establish that T does in fact have a unique fixed point.

The derivative of the maximand $m(\cdot)$ in (11) is

$$m'(\lambda) = -(A + 1 - \delta - \lambda)^{-\gamma} + \beta(1-\gamma)\lambda^{-\gamma} w.$$

Clearly $\lim_{\lambda \downarrow 0} m'(\lambda) = \infty$ and $\lim_{\lambda \uparrow A+1-\delta} m'(\lambda) = -\infty$, so $\lambda^\star$ is interior. It therefore solves the first-order condition

$$(A + 1 - \delta - \lambda^\star)^{-\gamma} = \beta(1-\gamma)(\lambda^\star)^{-\gamma} w.$$

Since $\lambda^\star$ attains the maximum in (11),

$$w = \tfrac{1}{1-\gamma}\left(A + 1 - \delta - \lambda^\star\right)^{1-\gamma} + \beta(\lambda^\star)^{1-\gamma}w,$$

which rearranges to

$$w = \frac{1}{1-\gamma}\frac{\left(A + 1 - \delta - \lambda^\star\right)^{1-\gamma}}{1 - \beta(\lambda^\star)^{1-\gamma}}.$$

Plugging this into the first-order condition yields

$$\left(A + 1 - \delta - \lambda^\star\right)^{-\gamma} = \beta(1-\gamma)(\lambda^\star)^{-\gamma}\frac{1}{1-\gamma}\frac{\left(A + 1 - \delta - \lambda^\star\right)^{1-\gamma}}{1 - \beta(\lambda^\star)^{1-\gamma}}.$$

Cancelling terms and rearranging,

$$\begin{aligned}
1 - \beta(\lambda^\star)^{1-\gamma} &= \beta(\lambda^\star)^{-\gamma}\left(A + 1 - \delta - \lambda^\star\right)\\
&= \beta(\lambda^\star)^{-\gamma}\left(A + 1 - \delta\right) - \beta(\lambda^\star)^{1-\gamma}.
\end{aligned}$$

Cancelling $-\beta(\lambda^\star)^{1-\gamma}$ from both sides and rearranging then yields

$$\lambda^\star = \left[\beta\left(A + 1 - \delta\right)\right]^{1/\gamma}.$$

We already verified interiority, but as a sanity check we can see directly from the two boundedness conditions $A > \delta$ and $\beta(A + 1 - \delta)^{1-\gamma} < 1$ that $\lambda^\star$ is interior.

$\lambda^\star$ is the growth rate in this economy. Note that it is the assumption $A > \delta$ which guarantees growth in equilibrium. This is what we would expect from the Euler equation, as $A > \delta$ ensures that the rate of return is above unity. Since $\lambda^\star$ doesn't depend on initial conditions, there is no convergence of the levels of two economies that start out with different capital stocks. To draw an intuitive parallel with the neoclassical growth model, there is no curvature at all in the production function, so the convergence time is infinite.[35]

---

[35]Larry drew another analogy here that I found less clear. He said that in models of small open economies, the rate of return on capital from investing abroad is totally insensitive to how much capital the small open economy invests.

12 Mon 2 Nov 2015

12.1 Outline of today

Today we will cover the Romer (1987) model. We'll begin by studying the partial equilibrium among firms (the interesting part of the economy), taking prices and the household's supply of labour and capital services as given. There are both final-goods firms and intermediate-goods firms, and their interaction is at the heart of the model. We'll then study the household's (very simple) problem, define sequence-of-markets equilibrium, and study the properties of the equilibrium. It will turn out that the reduced form of the model is essentially the *Ak* model.

The aim of the Romer model is to capture Smith's (1776) idea that wealth creation is driven by increasing specialisation in production. The model has another interesting feature: because intermediate-goods firms are monopolistic, the equilibrium is inefficient. In particular, the equilibrium growth rate is always lower than the optimal growth rate.

12.2 The Romer (1987) model

12.2.1 Production environment

Competitive final-goods firms produce a homogeneous good used for both consumption and investment.[36] They have access to a continuum of potential inputs indexed by $i \in [0, M]$, each of which is produced by a monopolistic intermediate-goods firm. M is an equilibrium object that is allowed to live anywhere in $\mathbf{R}_+$. In this section, we treat M as fixed, but we will pin it down when we study the general equilibrium.

The representative firm in the final-goods industry uses the production function

$$F(n, x) := n^{1-\alpha} \int_0^M x(i)^\alpha \mathrm{d}i,$$

where $x : [0, M] \to \mathbf{R}_+$ is such that $x(i)$ is the quantity (or intensity) used of input i.[37] Total usage of inputs is $N := \int_0^M x(i)\mathrm{d}i$. To get a feel for the

---

[36]Until now we've only had one good, so didn't have to worry about defining GDP. In this model, the quantity of final output is GDP.

[37]Although we will often refer to $x(i)$ as the 'quantity' used, it is more properly called the *intensity* of usage. Formally, $x(i)$ is the Radon-Nikodým derivative of quantity of usage. This raises a mathematical curiosity: we are implicitly restricting the final-goods firm to use only input configurations that possess a density w.r.t. Lesbegue measure. This actually rules out some input configurations that may be optimal in a general production problem, but no such configuration is optimal in the present model.

technology, set $n = 1$ and suppose that each input is used with equal intensity: $x(i) = N/M$. Then output is

$$\int_0^M x(i)^\alpha \mathrm{d}i = (N/M)^\alpha \int_0^M \mathrm{d}i = N^\alpha M^{1-\alpha}.$$

This is increasing in M, so for a given measure N of total input usage, we get more output when more input varieties are used.

Now suppose that (for whatever reason) N is allowed to increase to $N + \Delta N$. There are two polar ways in which we could change production: continue using M varieties at increased intensity $(N + \Delta N)/M$, or increase the number of varieties used to $M + \Delta M$ while keeping the intensity $(N + \Delta N)/(M + \Delta M)$ fixed. (Again, this is purely to illustrate the technology. Final-goods firms do *not* control M!) In the first case, output increases to

$$(N + \Delta N)^\alpha M^\alpha = N^\alpha M^{1-\alpha} \left(1 + \frac{\Delta N}{N}\right)^\alpha.$$

In the second case, ΔM solves

$$\frac{N + \Delta N}{M + \Delta M} = \frac{N}{M} \quad \Leftrightarrow \quad \Delta M = M\frac{\Delta N}{N},$$

so output is

$$(N + \Delta N)^\alpha (M + \Delta M)^{1-\alpha} = N^\alpha \left(1 + \frac{\Delta N}{N}\right)^\alpha \cdot M^{1-\alpha} \left(1 + \frac{\Delta N}{N}\right)^{1-\alpha}$$
$$= N^\alpha M^{1-\alpha} \left(1 + \frac{\Delta N}{N}\right).$$

This is the same expression, except that there's no α exponent on the final term. So output increases by more when the extra inputs ΔN are spread out over additional varieties of inputs instead of being used to stock up on more of the same inputs as before. This is simply a reflection of the fact that there are diminishing returns to each input, but constant returns to scale. This is important because in the Romer (1987) model, the source of economic growth will be growth in M.

The motivation for building a model of specialisation-driven growth is that historically, economic growth has gone hand-in-hand with an increasingly complex supply chain network in which producers use a dizzying number of inputs to produce their goods. For example, consider how many separate parts go into building a modern car or other appliance. The Romer (1987)

model suggests that this specialisation is linked fundamentally to economic growth.

A possible reaction to this story is to ask why, if specialisation is so fantastic, firms don't jump to full specialisation in period 0? The answer in the Romer (1987) model is that in order to produce an intermediate good, a firm must pay a fixed cost. Only in a rich economy in which demand is high will it be worthwhile for firms to enter lots of these industries. So we get a gradual, circular process of growth in which economies get richer by specialising and become more specialised in response to getting richer.

12.2.2 Partial equilibrium among firms

We normalise the price of the output good to unity. The representative competitive final-goods firm solves

$$\max_{n,x(\cdot)} \left[y - wn - \int_0^M p(i)x(i)\mathrm{d}i : y = n^{1-\alpha} \int_0^M x(i)^\alpha \mathrm{d}i \right].$$

The first-order conditions are

$$(1 - \alpha)y/n = w \tag{12}$$

$$\alpha[n/x(i)]^{1-\alpha} = p(i) \quad \text{for each } i \in [0, M]. \tag{13}$$

We will think of these as demand curves for labour and intermediate goods.

The real action in this model is with the intermediate-goods firms. Each intermediate good $i \in [0, M]$ is produced by a single firm, e.g. because of patents. Each of these firms optimises subject to demand from the final-goods industry rather than taking prices as given. However, we can ignore strategic interactions between monopolists because each monopolist is atomistic with respect to n and y in the final-goods firm's problem. (This is the virtue of assuming a continuum of inputs.) So each monopolist sets marginal revenue equal to marginal cost. Using the demand curve (13), monopolist i's marginal revenue is

$$\frac{\mathrm{d}}{\mathrm{d}x(i)} \left(\alpha[n/x(i)]^{1-\alpha} x(i) \right) = \alpha^2 [n/x(i)]^{1-\alpha}.$$

To get marginal cost, we first have to specify the cost structure of intermediate-goods firms. To produce $x(i)$, firm i needs to rent $h(x(i))$ units of capital.[38] h has the property that $h(0) = 0$, but that for some $B > 0$,

---

[38]There are variants of the model in which they hire labour or yet other factors of production.

$h(x) > B$ for all $x > 0$. That is, there's a discontinuity at the origin, so that any positive production requires more capital than some lower bound $B > 0$ (this translates into a fixed cost). We also assume that h is convex, corresponding to increasing returns to scale. To get closed-form solutions, we will consider a particular specification for h:

$$h(x) = \begin{cases} 0 & \text{for } x = 0 \\ \frac{1}{2}\left[1 + x(i)^2\right] & \text{for } x > 0. \end{cases}$$

The cost of renting a unit of capital is r, which intermediate-goods firms take as given. It follows that marginal cost for $x > 0$ is $rh'(x) = rx$. The monopolist's profit-maximising choice of $x(i)$ is therefore

$$\alpha^2[n/x(i)]^{1-\alpha} = rx(i),$$

the intersection of marginal revenue and marginal cost. Notice that every firm i faces the same problem, so will choose the same output $x(i) = \bar{x}$. By the demand curve (13), this will then lead to a homogeneous price $p(i) = \bar{p}$ for intermediate goods. So we can write the monopoly first-order condition as

$$\alpha^2 n^{1-\alpha} = r\bar{x}^{2-\alpha}. \tag{14}$$

We can also rewrite the first-order conditions (12) and (13) of final-goods firms as

$$(1-\alpha)y/n = w \tag{15}$$

$$\alpha\left[n/\bar{x}\right]^{1-\alpha} = \bar{p}. \tag{16}$$

Each intermediate-goods firm's capital demand is $\frac{1}{2}\left[1 + \bar{x}^2\right]$. Letting k stand for the aggregate stock of capital, equilibrium requires the capital market to clear:

$$\frac{1}{2}\left[1 + \bar{x}^2\right]M = k. \tag{17}$$

Although each intermediate-goods producer has market power, there is free entry into the intermediate-goods industry. Since there is no uncertainty, this means that the profits of each monopolist must be exactly zero in equilibrium, i.e. the fixed cost $1/2$ exactly offsets the monopoly rents earned by incumbents.[39] The zero-profit condition is $\bar{p}\bar{x} = rh(\bar{x})$, or

$$\bar{p}\bar{x} = r\frac{1}{2}\left[1 + \bar{x}^2\right].$$

---

[39]Matsuyama (1999) is a nice variation on this model in which the fixed cost is replaced by a sunk cost. This fits a little better with leading examples of innovation in monopolistic industries, e.g. the pharmaceutical industry.

Substituting for $\bar{p}$ using the demand curve (16),

$$\alpha \bar{x}^\alpha n^{1-\alpha} = r \frac{1}{2} \left[1 + \bar{x}^2 \right] \quad \text{for each } i \in [0, M]. \tag{18}$$

Finally, the goods market must clear:

$$y = n^{1-\alpha} \bar{x}^\alpha M. \tag{19}$$

This gives us six equations (14), (15), (16), (17), (18) and (19) in the six variables $\bar{x}$, $\bar{p}$, y, w, r and M. Note well that n and k are treated as given, since they are supplied by the household and so determined in general equilibrium. For later reference (when we define the general equilibrium), we say that quantities $(\bar{x}, y, M)$ and prices $(r, w, \bar{p})$ form a partial equilibrium among firms at (k, n) iff they solve equations (14) through (19) when the supply of capital and labour services from households is (k, n).

Now let's plug things into each other to obtain some properties of partial equilibria. Rearrange (14) to obtain:

$$\alpha \left[\alpha \bar{x}^\alpha n^{1-\alpha} \right] = r \bar{x}^2.$$

Substituting on the left-hand side using (18),

$$\alpha r \frac{1}{2} \left[1 + \bar{x}^2 \right] = r \bar{x}^2.$$

Rearranging yields

$$\bar{x} = \sqrt{\frac{\alpha}{2 - \alpha}}.$$

Hence
$$h(\bar{x}) = \frac{1}{2} \left[1 + \bar{x}^2 \right] = \frac{1}{2} \left[\frac{2}{2 - \alpha} \right] = \frac{1}{2 - \alpha},$$

whence (17) yields
$$M = (2 - \alpha)k.$$

Substituting for $\bar{x}$ and M in (19),

$$y = n^{1-\alpha} \cdot \left(\frac{\alpha}{2 - \alpha} \right)^{\alpha/2} \cdot (2 - \alpha)k = n^{1-\alpha} Ak,$$

where
$$A := [\alpha/(2 - \alpha)]^{\alpha/2} (2 - \alpha) = (\alpha/2)^{\alpha/2} (1 - \alpha/2)^{1-\alpha/2}$$

is a constant. Since $n = 1$ in general equilibrium, we have $y = Ak$, essentially an Ak model![40] This happy coincidence is a consequence of carefully chosen functional forms that have the knife-edge property that increasing variety exactly offets diminishing marginal returns.

Now that we have all three quantities $(\bar{x}, y, M)$, we can back out the prices if desired: w from (15), $\bar{p}$ from (16), and r from (14). The last is rather interesting:

$$
\begin{aligned}
r &= \alpha^2 n^{1-\alpha} \bar{x}^{-(2-\alpha)} \\
&= \alpha^2 n^{1-\alpha} \left(\frac{2-\alpha}{\alpha} \right)^{1-\alpha/2} \\
&= \alpha n^{1-\alpha}(2-\alpha) \left(\frac{2-\alpha}{\alpha} \right)^{-\alpha/2} \\
&= \alpha n^{1-\alpha} \left(\frac{\alpha}{2-\alpha} \right)^{\alpha/2} (2-\alpha) \\
&= \alpha n^{1-\alpha} A
\end{aligned}
$$

using the definition of A. Since $n = 1$ in general equilibrium, it follows that $r = \alpha A$ in equilibrium. But the marginal product of capital, and hence the efficient rental rate of capital, is A. So the rental rate is inefficiently low whenever $\alpha < 1$, precisely because intermediate-goods firms have market power in this case. Accordingly, as $\alpha \to 1$ (monopoly power vanishes), r converges to its efficient level A.

12.2.3   Households and general equilibrium

The household's problem at date t is

$$
\max_{\{k_{t+1+j}, c_{t+j}, n_{t+j}\}_{j=0}^{\infty}} \sum_{j=0}^{\infty} \beta^j (1-\gamma)^{-1} c_{t+j}^{1-\gamma}
$$
$$
\text{s.t. } \forall j \in \mathbf{Z}_+ \quad c_{t+j} + k_{t+1+j} - (1-\delta)k_{t+j} \le r_{t+j} k_{t+j} + w_{t+j} n_{t+j} + \pi_{t+j}
$$
$$
c_{t+j}, k_{t+1+j} \in \mathbf{R}_+, \quad n_{t+j} \in [0,1]
$$

where k_t is taken as given. We say that $\{k_{t+1+j}, c_{t+j}, n_{t+j}\}_{j=0}^{\infty}$ solves the household's problem from k_t at prices $\{r_{t+j}, n_{t+j}\}_{j=0}^{\infty}$ iff it attains the maximum in this program.

With households in place, we are ready to define equilibrium.

---

[40]I say 'essentially' because there is no market power in the Ak model.

Definition 7. Quantities $\{\bar{x}_t, y_t, M_t, k_t, c_t, n_t\}_{t=0}^{\infty}$ and prices $\{w_t, r_t, \bar{p}_t\}_{t=0}^{\infty}$ form a sequence-of-markets equilibrium iff

(1) At each date t, $\{k_{t+1+j}, c_{t+j}, n_{t+j}\}_{j=0}^{\infty}$ solves the household's problem from k_t at prices $\{r_{t+j}, n_{t+j}\}_{j=0}^{\infty}$.

(2) At each date t, quantities $(\bar{x}_t, y_t, M_t)$ and prices $(r_t, w_t, \bar{p}_t)$ form a partial equilibrium among firms at (k_t, n_t).

(3) At each date t, the goods market clears:

$$y_t = c_t + k_{t+1} - (1 - \delta)k_t.$$

As usual, a bunch of additional market-clearing conditions are imposed implicitly. Since any interior solution to the household's problem satisfies the Euler condition and transversality condition, we have a collection of equations that can (in principle) be solved for all the equilibrium quantities. In particular, any equilibrium in which $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$ is interior is characterised by (1) the Euler equation and transversality condition, (2) the six partial equilibrium conditions, and (3) the goods-market-clearing condition.

The following are some properties of the equilibrium.

(1) $n_t = 1$ at each t. (Obviously optimal for the household.)

(2) The price of capital is fixed at unity. This is because diminishing marginal returns are counteracted by the gradual growth of input varieties M_t.[41]

(3) There is no convergence since the reduced form the model is Ak.

(4) The equilibrium is inefficient since $r_t = \alpha A$ $\forall t$, whereas the marginal product of capital is A. We get efficiency only in the limit $\alpha \to 1$ of no market power. From the Euler equation, consumption growth is

$$\frac{c_{t+1}}{c_t} = \beta^{1/\gamma} \left(\alpha A + 1 - \delta\right)^{1/\gamma},$$

which is increasing in αA. So market power leads to an inefficiently low rate of growth.

---

[41]This is the explanation Larry gave, but it seems to me that the price of capital in this model has nothing to do with firms.

(5) We cannot improve welfare by forcing monopolists to price at marginal cost. Since they face a fixed cost, this would lead to no entry, and hence to zero output! But there are other 'realistic' policy interventions that could correct the inefficiency. The obvious one is to subsidise production of intermediate goods (see homework 7). Provided that the subsidy is financed by lump-sum taxation, this policy can implement the optimum if the right subsidy rate is chosen.[42]

I asked Larry why we should believe anything coming out of a model with so many restrictive functional-form assumptions. He said it should be viewed as a possiblity result: growth can be generated from specialisation alone, without the need for exogenous technical change.

---

[42]In the lecture, Larry said that another way to remedy the inefficiency is to give cash prizes (financed by lump-sum taxation) rather than patents to innovators. While that works in other models, it seems to me to be beside the point here, since we assumed that there is free entry into every industry! The source of the inefficiency is the fixed cost, which lead to increasing returns and hence to highly concentrated industries with monopoly power.

13   Wed 4 Nov 2015

13.1   Outline of today

Today we'll talk about another endogenous growth model, the human capital model. Here long-run growth is driven by investment in human capital. The model can be viewed as a way of endogenising TFP growth in the Solow I model.

What is human capital accumulation? There are two basic varieties: on-the-job (or learning-by-doing) and going-to-school.[43] We will focus on the going-to-school variety of human capital accumulation. As we'll see, it has some similarities with the Romer model: costly investment in accumulation gradually improves productivity, and the reduced for is Ak.

We would like to address some of the following questions. How does human capital accumulation give rise to long-run growth? Is there convergence as in the exogenous growth models, or path-dependence like in the Ak model?

We will only study the efficient allocations in this model. The technicalities are left for the reader to complete, as is the study of decentralised allocations. Some of this was covered in a question on homework 7.

13.2   The human capital accumulation model

We'll describe the model and scale it (section 13.2.1), prove boundedness (section 13.2.2) and provide a recursive formulation (section 13.2.3). It is left to the reader to complete the argument by using a perturbation method to study the optimal policy.

13.2.1   Model and scaling

As ever, preferences are ordered by

$$\sum_{t=0}^{\infty} \beta^t (1-\gamma)^{-1} c^{1-\gamma}$$

for $\gamma > 0$. The resource constraint is

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

We assume that investment cannot be negative, i.e. capital is inedible: $k_{t+1} \geq (1-\delta)k_t$. As usual, $c_t \geq 0$.

---

[43]The former links up with a commonly perceived benefit of foreign direct investment: that local workers become more productive via learning-by-doing at a technologically advanced foreign factory.

n_t is labour input in efficiency units. n_t is constrained to lie in $[0, h_t]$, where h_t is interpreted as human capital. Human capital evolves according to

$$h_{t+1} = h_t + \lambda(h_t - n_t),$$

where $\lambda > 0$ is the efficiency of investing in human capital, i.e. how much human capital you get per efficiency hour spent in schooling.[44] $k_0, h_0 > 0$ are given.

As in the Ak model, we make a boundedness assumption that rules out fast growth, thus guaranteeing that utility is finite. In this context, the restriction is $\beta(1 + \lambda)^{1-\gamma} < 1$. As before, some technical wizardry would allow us to solve the model even without this condition.

To put the problem in recursive form, we're going to scale everything by human capital h_t. Dividing both sides of the resource constraint by h_t yields

$$\frac{c_t}{h_t} + \frac{k_{t+1}}{h_t} - (1 - \delta)\frac{k_t}{h_t} \leq \left(\frac{k_t}{h_t}\right)^\alpha \left(\frac{n_t}{h_t}\right)^{1-\alpha}.$$

Define the choice variables $x_t := k_t/h_t$ and $y_t := h_{t+1}/h_t$. Then the reformulated resource constraint is

$$\frac{c_t}{h_t} + x_{t+1}y_t - (1 - \delta)x_t \leq x_t^\alpha \left(\frac{n_t}{h_t}\right)^{1-\alpha}.$$

The law of motion for human capital becomes

$$y_t = 1 + \lambda - \lambda\frac{n_t}{h_t}.$$

Since $n_t \in [0, h_t]$, $y_t \in [1, 1 + \lambda]$. Since $k_{t+1} \geq (1 - \delta)k_t$ (nonnegative investment), the lower bound on x_{t+1} is

$$x_{t+1} \geq \frac{(1 - \delta)x_t}{y_t}.$$

To get the upper bound on x_{t+1}, rearrange the the law of motion for y_t to obtain

$$\frac{n_t}{h_t} = \lambda^{-1}(1 + \lambda - y_t).$$

Substituting into the reformulated resource constraint,

$$x_{t+1}y_t \leq (1 - \delta)x_t - \frac{c_t}{h_t} + x_t^\alpha \left[\lambda^{-1}(1 + \lambda - y_t)\right]^{1-\alpha}.$$

---

[44]We don't allow human capital to depreciate, but that would be a nice variation.

Using $c_t \geq 0$ and dividing through by y_t yields the upper bound

$$x_{t+1} \leq \frac{(1-\delta)x_t + x_t^\alpha \left[\lambda^{-1}(1+\lambda-y_t)\right]^{1-\alpha}}{y_t}.$$

So x_{t+1} is constrained to lie in $\Gamma(x_t, y_t)$ where $\Gamma : \mathbf{R}_+^2 \to \mathbf{R}_+$ is defined by

$$\Gamma(x,y) := \left[\frac{(1-\delta)x}{y}, \frac{(1-\delta)x + x^\alpha \left[\lambda^{-1}(1+\lambda-y)\right]^{1-\alpha}}{y} \right]$$

for any $(x,y) \in \mathbf{R}_+^2$.

Period utility can be rewritten

$$(1-\gamma)^{-1}c_t^{1-\gamma} = (1-\gamma)^{-1}h_t^{1-\gamma} \left(\frac{c_t}{h_t}\right)^{1-\gamma}.$$

Rearranging the reformulated resource constraint and using the fact that it must hold with equality at an optimum,

$$\frac{c_t}{h_t} = -x_{t+1}y_t + (1-\delta)x_t + x_t^\alpha \left(\frac{n_t}{h_t}\right)^{1-\alpha}$$

$$= -x_{t+1}y_t + (1-\delta)x_t + x_t^\alpha \left[\lambda^{-1}(1+\lambda-y_t)\right]^{1-\alpha}.$$

So define $u : \mathbf{R}_+^3 \to \mathbf{R}$ by

$$u(x, x', y) := (1-\gamma)^{-1} \left(-x'y + (1-\delta)x + x^\alpha \left[\lambda^{-1}(1+\lambda-y)\right]^{1-\alpha} \right)^{1-\gamma}.$$

for every $(x, x', y) \in \mathbf{R}_+^3$. Using the 'accordion trick',

$$h_t = y_{t-1}h_{t-1} = y_{t-1}y_{t-2}h_{t-2} = \left(\prod_{j=0}^{t-1} y_j\right) h_0.$$

So we can write period utility as

$$h_0^{1-\gamma} \left(\prod_{j=0}^{t-1} y_j^{1-\gamma}\right) u\left(x_t, x_{t+1}, y_t\right).$$

We are now ready to formulate a recursive optimal growth problem:

$$\max_{\{y_t, x_{t+1}\}_{t=0}^\infty} h_0^{1-\gamma} \sum_{t=0}^\infty \beta^t \left(\prod_{j=0}^{t-1} y_j^{1-\gamma}\right) u\left(x_t, x_{t+1}, y_t\right)$$

$$\text{s.t.} \quad y_t \in [1, 1+\lambda], \quad x_{t+1} \in \Gamma(x_t, y_t) \quad \forall t \in \mathbf{Z}_+$$

where $x_0 := k_0/h_0$ is given. Write $h_0^{1-\gamma}w(x_0)$ for the value of the problem. (The value obviously doesn't depend on y_{-1}.)

13.2.2 Proving boundedness

Before proceeding to use our ordinary recursive methods, we have to show that w is bounded.[45] Since $c_t/h_t \geq 0$, $u(x, x', y)$ has the same sign as $1 - \gamma$. $\prod_{j=0}^{t-1} y_j^{1-\gamma}$ is always nonnegative since $y_j \geq 1 > 0$. Hence for $\gamma > 1$, $w(x_0) \leq 0 < \infty$, and for $\gamma < 1$, $w(x_0) \geq 0 > -\infty$. So we just need to establish a lower bound for when $\gamma > 1$ and an upper bound for when $\gamma < 1$.

Take $\gamma < 1$. Then $\prod_{j=0}^{t-1} (y_j)^{1-\gamma} \leq \prod_{j=0}^{t-1} (1 + \lambda)^{1-\gamma}$. Suppose we had an upper bound B for u. Then the objective is bounded above by

$$B \sum_{t=0}^{\infty} \left[\beta(1+\lambda)^{1-\gamma} \right]^t = B \left[1 - \beta(1+\lambda)^{1-\gamma} \right]^{-1} < \infty,$$

where the equality used our boundedness assumption $\beta(1+\lambda)^{1-\gamma} < 1$. So all we need to do is find an upper bound B for u. Begin with x_{t+1}, which is constrained to lie in $\Gamma(x_t, y_t)$, so in particular

$$x_{t+1} \leq \frac{(1-\delta)x_t + x_t^\alpha \left[\lambda^{-1}(1 + \lambda - y_t) \right]^{1-\alpha}}{y_t} \leq (1-\delta)x_t + x_t^\alpha$$

where we used the fact that $y_t \geq 1$. This is the bound from the neoclassical model (unsurprisingly, since this model with $y_t = 1\ \forall t$ is the neoclassical growth model). This difference equation exhibits monotone convergence to a steady state $x^\star$, so x_t can never exceed $\bar{x} := \max\{x^\star, x_0\}$. Moreover, $x_{t+1} \geq 0$. Hence for a given x_0, we can bound u as

$$u(x_t, x_{t+1}, y_t)$$

$$= (1-\gamma)^{-1} \left(-x_{t+1}y_t + (1-\delta)x_t + x_t^\alpha \left[\lambda^{-1}(1 + \lambda - y_t) \right]^{1-\alpha} \right)^{1-\gamma}$$

$$\leq (1-\gamma)^{-1} \left(-0 \cdot 1 + (1-\delta)\bar{x} + \bar{x}^\alpha \left[\lambda^{-1}(1 + \lambda - 1) \right]^{1-\alpha} \right)^{1-\gamma}$$

$$= (1-\gamma)^{-1} \left((1-\delta)\bar{x} + \bar{x}^\alpha \right)^{1-\gamma}$$

$$< \infty.$$

Now take $\gamma > 1$. As in the Ak model, we will find a lower bound for utility by constructing a feasible path that generates a finite payoff. To this end, take $y_t = 1\ \forall t$ (so that we're back in the neoclassical growth model). If $x_0 \leq x^\star$

---

[45]We won't explicitly cover the limit $\gamma \to 1$, but a continuity argument should do the trick.

then $x_{t+1} = x_0$ $\forall t$ is feasible, in which case consumption if $c_0 = h_0 x_0 > 0$ in every period, yielding utility

$$-(1-\beta)^{-1}(\gamma-1)^{-1}(h_0 x_0)^{-(\gamma-1)} > -\infty.$$

The case $x_0 > x^\star$ just requires a small tweak. Simply invest nothing until x_{t+1} hits $x^\star$, then stay at $x^\star$. Along the path to $x^\star$ consumption is higher than if $x_0 \le x^\star$, so this yields weakly higher utility, and hence a lower bound.

13.2.3 Recursive formulation

With boundedness in place, we can reformulate the sequence problem of the planner as a functional equation:

$$w(x) = \max_{y' \in [1, 1+\lambda], x' \in \Gamma(x, y')} \left\{ u(x, x', y') + \beta(y')^{1-\gamma} w(x') \right\}.$$

This model fits neatly into the Stokey–Lucas framework. Assumptions 4.1 and 4.2 hold, so the sequence problem and functional equation are equivalent in the sense of Theorems 4.2-4.5 of Stokey et al. (1989). As usual, we define an operator T on $C(\mathbf{R}_+)$ by

$$(Tf)(x) := \max_{y' \in [1, 1+\lambda], x' \in \Gamma(x, y')} \left\{ u(x, x', y') + \beta(y')^{1-\gamma} f(x') \right\}$$

for every $f \in C(\mathbf{R}_+)$ and $x \in \mathbf{R}_+$. Since Assumptions 4.3 and 4.4 hold, by Theorem 4.6 we have that T is a contraction mapping whose unique fixed point is the value function v of the sequence problem.

The rest of the argument is sketched as follows. (Larry asked us to fill in the details at home. Some of this material was covered in homework 7.) The Benveniste–Scheinkman theorem applies, so w is differentiable with derivative $w'(x) = u_1(x, f(x), g(x))$ where $(f(\cdot), g(\cdot))$ are the maximal choices as a function of x. It's straightforward to approximate f and g using the (first-order) perturbation method. Note that we need to find a steady state first to use the perturbation method.

The solution has the property that the growth rate is the same at every state. Hence there is no convergence. You'd expect the steady state to be stable because when the capital stock is large relative to the human capital stock, the rate of return to physical capital is high, so investment is pumped into physical capital.

14   Mon 9 Nov 2015

14.1   Outline of today

Today we'll study the overlapping-generations model of Jones and Manuelli (1997, sec. 7). We first lay out the model and study the technological elasticity of substitution. We'll then show that growth is technologically feasible regardless of the elasticity of substitution. Nevertheless, we will see that sustained growth is not possible in equilibrium except in the knife-edge case of unit elasticity of substitution (Cobb–Douglas technology). When the elasticity of substitution is higher, sustained growth leads labour's share of income to diminish over time, so that the young eventually cannot afford to save enough to sustain capital accumulation.

14.2   An overlapping-generations model

14.2.1   Environment

Each household lives for two periods. The household born at t consumes c_t^t when young and c_{t+1}^t when old. Utility is additive and discounted as usual:

$$u\left(c_t^t\right) + \beta u\left(c_{t+1}^t\right)$$

The young born at t inelastically supply one unit of labour at wage w_t, and have no other source of income (e.g. no bequests). The only savings technology they have access to is capital k_{t+1}. The price of consumption is normalised to unity in each period. So they face the budget constraint

$$c_t^t + k_{t+1} \leq w_t.$$

The old also supply one unit of labour, earning labour income w_{t+1}. In addition, they earn income $r_{t+1}k_{t+1}$ from their capital holdings. Finally, before dying they sell their undepreciated capital $(1-\delta)k_{t+1}$ to the young. Since consumption can be transformed into capital at a unit rate of marginal transformation, the price of capital is unity, so the income that the old earn from selling their assets is simply $(1-\delta)k_{t+1}$. (As usual, the household operates a technology in its back yard for converting output goods into fresh capital.) The budget constraint of the old born at t is therefore

$$c_{t+1}^t \leq w_{t+1} + (1 - \delta + r_{t+1})k_{t+1}.$$

Competitive firms produce consumption goods using a technology $f : \mathbf{R}_+^2 \to \mathbf{R}$ that uses capital and labour services as inputs. The representative

firm solves the competitive profit-maximisation problem

$$\max_{(k_t, n_t) \geq 0} \left[f(k_t, n_t) - r_t k_t - w_t n_t \right].$$

As usual, equilibrium will require (r, w) to lie on the factor price frontier. The zero-profit (or first-order) conditions for competitive firms are $f_1(k_t, n_t) = r_t$ and $f_2(k_t, n_t) = w_t$.

Demand for output goods at time t comes from three sources: the consumption c_t^t of the young, the consumption c_t^{t-1} of the old, and the net saving $k_{t+1} - (1 - \delta)k_t$ of the young. The last looks the way it does because the young buy $(1 - \delta)k_t$ directly from the time-t old, so only need to buy enough output goods to make up the difference.

The goods-market clearing condition is therefore

$$c_t^t + c_t^{t-1} + k_{t+1} - (1 - \delta)k_t = f(k_t, 2),$$

where we've imposed labour-market clearing $n_t = 2$.

We assume a particular form for f: for some $b > 0$ and $\rho \in (0, 1)$, $f(k, n) = bg(k, n) \ \forall (k, n) \in \mathbf{R}_+^2$, where

$$g(k, n) := [\alpha k^\rho + (1 - \alpha)n^\rho]^{1/\rho} \quad \forall (k, n) \in \mathbf{R}_+^2.$$

This is called constant-elasticity-of-substitution technology (it is obviously inspired by Dixit and Stiglitz (1977) preferences). To motiviate the name, write $k(r, w)$ and $n(r, w)$ for the input choices that satisfy the zero-profit (first-order) conditions at factor prices (r, w). The elasticity of substitution between capital and labour is defined as

$$\eta(r, w) := -\frac{\mathrm{d}\ln(k(r, w)/n(r, w))}{\mathrm{d}\ln(r/w)}.$$

We will show that for this technology, $\eta(r, w) = 1/(1 - \rho) \ \forall (r, w)$, a constant.

Begin by computing the marginal product of capital for inputs (k, n) as

$$f_1(k, n) = bg_1(k, n) = b \cdot \rho^{-1} g(k, n)^{1-\rho} \cdot \rho\alpha k^{\rho-1} = \alpha b(g(k, n)/k)^{1-\rho}.$$

By an analogous calculation, $f_2(k, n) = (1 - \alpha)b(g(k, n)/n)^{1-\rho}$. Dividing the zero-profit/first-order conditions by each other yields

$$\frac{f_1(k(r, w), n(r, w))}{f_2(k(r, w), n(r, w))} = \frac{r}{w}.$$

107

Substituting for the marginal products and cancelling, we obtain

$$\frac{\alpha}{1-\alpha} \left(\frac{n(r,w)}{k(r,w)} \right)^{1-\rho} = \frac{r}{w}.$$

Taking logs,

$$\ln \left(\frac{\alpha}{1-\alpha} \right) - (1-\rho) \ln \left(\frac{k(r,w)}{n(r,w)} \right) = \ln \left(\frac{r}{w} \right).$$

Finally, differentiate with respect to (r/w) and rearrange to obtain

$$\eta(r,w) = -\frac{\mathrm{d}\ln(k(r,w)/n(r,w))}{\mathrm{d}\ln(r/w)} = \frac{1}{1-\rho}$$

as claimed.

It is important (in my view) to understand that the elasticity of substitution is not a purely technological object. Rather, it is the outcome of profit-maximising behaviour subject to constraints and prices. As a result, it will in general depend on all the exogenous features of the profit-maximisation problem, viz. technology and factor prices (r,w). The fact that it does not depend on (r,w) in this case is a curiosity. (It's called CES technology precisely because no other technology has the property that the elasticity of substitution is constant in (r,w)!)

Observe that as $\rho \to 0$, the elasticity of substitution converges to unity. In fact, if we take $\rho \to 0$ for the production function, we obtain Cobb–Douglas technology with exponent α on k. (To verify this, you'll need l'Hôpital's rule.) So our new technology nests the Cobb–Douglas case.

To preview imminent developements, the Cobb–Douglas case will turn out to be a special knife-edge case in which growth is possible. Whenever $\rho > 0$, long-run growth will not be possible in equilibrium because as the capital stock grows, factor demand substitutes away from labour and towards capital, lowering labour's equilibrium share of income. This will eventually make young workers sufficiently income-poor that they cannot afford to buy enough capital to keep growth going.

14.2.2 Growth

The first order of the day is to show that growth is technologically feasible for some $b > 0$ even when $\rho > 0$. In particular, we will show that $k_{t+1}/k_t \geq 1+a$ for some $a > 0$ and $\forall t$ is possible provided b is large enough. A high b will (intuitively) lead to a high rate of return on capital, showing that a consistently

high rate of return is a necesary condition for long-run growth. Nevertheless, we will see that no equilibrium exhibits long-run growth, demonstrating that a high rate of return is not sufficient for growth.

To maximise growth, set consumption to zero: $c_t^t = c_t^{t-1} = 0 \ \forall t$. Then goods-market clearing becomes (using Euler's formula)

$$
\begin{aligned}
k_{t+1} &= (1 - \delta)k_t + f(k_t, 2) \\
&= (1 - \delta)k_t + f_1(k_t, 2)k_t + f_2(k_t, 2)2 \\
&= (1 - \delta)k_t + b\left[\alpha\left(\frac{g(k_t, 2)}{k_t}\right)^{1-\rho} \cdot k_t + (1 - \alpha)\left(\frac{g(k_t, 2)}{2}\right)^{1-\rho} \cdot 2\right].
\end{aligned}
$$

Hence $k_{t+1}/k_t \geq 1 + a$ is equivalent to

$$
\begin{aligned}
1 + a &\leq 1 - \delta + b\left[\alpha\left(\frac{g(k_t, 2)}{k_t}\right)^{1-\rho} + (1 - \alpha)\left(\frac{g(k_t, 2)}{2}\right)^{1-\rho}\frac{2}{k_t}\right] \\
&= 1 - \delta + b\left[\alpha\left(\frac{g(k_t, 2)}{k_t}\right)^{1-\rho} + 2^\rho(1 - \alpha)\frac{g(k_t, 2)^{1-\rho}}{k_t}\right] \\
&= 1 - \delta + b\left[\alpha g(1, 2/k_t)^{1-\rho} + 2^\rho(1 - \alpha)g(1, 2/k_t)g(k_t, 2)^{-\rho}\right].
\end{aligned}
$$

Observe that as $k_t \to \infty$,

$$
g(1, 2/k_t) = \left[\alpha + 2^\rho(1 - \alpha)k_t^{-\rho}\right]^{1/\rho} \to \alpha^{1/\rho},
$$

and hence

$$
g(k_t, 2)^{-\rho} = g(1, 2/k_t)^{-\rho}k_t^{-\rho} \to 0.
$$

Therefore, we can sustain growth at all t iff

$$
1 + a \leq 1 - \delta + b\alpha\alpha^{(1-\rho)/\rho} = 1 - \delta + b\alpha^{1/\rho}.
$$

For any $a > 0$, there is a $b > 0$ such that this inequality holds, so growth is technologically feasible provided b can be chosen large enough. Observe that the right-hand side is the asymptotic rate of return on capital along a path with sustained growth. This is because that term is $(f_1(k_t, 2)k_t + f_2(k_t, 2)2)/k_t$, total income in the economy. We showed that the second term vanishes, meaning that labour's share of income converges to zero. We're therefore left with $f_1(k_t, 2)$, the marginal product of capital. So growth is technologically feasible iff the marginal product of capital can be kept sufficiently high. As we shall see, however, this does not guarantee the existence of an equilibrium with growth.

To show that no equilibria exhibit growth, we must first define what we mean by an equilibrium in this context.

Definition 8. Quantities $\{c_t^{t-1}, c_t^t, k_t, n_t\}_{t=0}^{\infty}$ and prices $\{r_t, w_t\}_{t=0}^{\infty}$ form a sequence-of-markets equilibrium iff

(1) For each $t \geq 0$, $(c_t^t, c_{t+1}^t, k_{t+1})$ solves the household's problem at prices (w_t, w_{t+1}, r_{t+1}). Moreover, $c_0^{-1} = (1 - \delta + r_0)k_0$ (solution to the problem of the household born at -1).

(2) For each $t \geq 0$, (k_t, n_t) solves the firm's problem at prices (r_t, w_t).

(3) For each $t \geq 0$, the goods market clears:

$$c_t^{t-1} + c_t^t + k_{t+1} - (1 - \delta)k_t = f(k_t, n_t).$$

(4) For each $t \geq 0$, $n_t = 2$ (labour market clears).

We now show that when $\rho > 0$, there cannot be an equilibrium with growth, regardless of how large b is. Suppose for a contradiction that there is an equilibrium with growth: $k_{t+1}/k_t \geq 1 + a$ for $a > 0$ $\forall t$.[46] Let $y_t = f(k_t, 2)$ denote GDP. Labour's share of national income is

$$\frac{2w_t}{y_t} = \frac{2(1-\alpha)b(g(k_t, 2)/2)^{1-\rho}}{bg(k_t, 2)} = \frac{2^{\rho}(1-\alpha)}{g(k_t, 2)^{\rho}} = \frac{2^{\rho}(1-\alpha)}{\alpha k_t^{\rho} + (1-\alpha)2^{\rho}}.$$

So as $k_t \to \infty$, $2w_t/y_t \to 0$. That is, GDP grows faster than labour income. This should already suggest that the young will be unable to afford to buy more capital in every period!

The budget constraint of the young implies $k_{t+1} \leq w_t - c_t^t \leq w_t$. Together with the assumption of perpetual growth, this implies that

$$1 + a \leq \frac{k_{t+1}}{k_t} \leq \frac{w_t}{k_t} = \frac{w_t}{y_t}\frac{y_t}{k_t}.$$

As $k_t \to \infty$, the final term converges to

$$\frac{y_t}{k_t} = b\left[\alpha + (1-\alpha)(2/k_t)^{\rho}\right]^{1/\rho} \to b\alpha^{1/\rho}.$$

So for all $\varepsilon > 0$, there is a $T \in \mathbf{N}$ such that for all $t \geq T$,

$$\frac{w_t}{y_t}\frac{y_t}{k_t} < \varepsilon.$$

---

[46]The proof can be extended to cover the case of non-monotonic long-run growth, but we won't bother.

So by choosing $\varepsilon < 1 + a$, we obtain

$$1 + a > \frac{w_t}{y_t} \frac{y_t}{k_t} \quad \text{for all } t \geq T,$$

contradicting long-run growth.

The intuition, which is probably getting old by now, is that since labour income grows slower than GDP, the young will eventually be unable to afford to keep the savings rate high. Note well that we required $\rho > 0$ to obtain this result. As $\rho \to$, $b\alpha^{1/\rho} \to 0$, and in fact growth can be sustained in equilibrium. This is because when $\rho = 0$ (the elasticity of income substitution is unity), there is no gradual shift of factor demand away from labour and toward capital as the capital stock grows. Hence labour's share of national income remains constant over time rather than diminishing, allowing the young to afford to continue to accumulate capital.

As an aside, until about 15 years ago, growth people were wedded to Cobb–Douglas. The rationale was that labour's share of income has in fact been roughly constant. (This was one of the Kaldor (1957) 'facts' of balanced growth.) But this hasn't been the case for the last 30 years or so, actually! Instead, the labour share has been declining, consistent with $\rho > 0$.

Even if we maintain $\rho > 0$, we can still get long-run growth if we add bequests and intergenerational altruism to the model. Under strong assumptions, Barro (1974) shows that we can in fact recover the neoclassical growth model because infinite-lived altruistic dynasties act like a single, infinite-lived representative agent! Empirically, bequests are infrequent but large, so may be important. If we want to pursue the implications of a declining labour share for inequality, adding bequests is a natural route.

The fact that there are no growth equilibria raises the question of what kinds of equilibria there are. Let's focus on steady-state equilibria. Suppose b is large; then the equilibrium rental rate r is high, so each generation has a strong incentive to save and thus chooses a steep consumption profile. To study this explicitly, write $c_t^t = c$ and $c_{t+1}^t = \lambda c \; \forall t$. The Euler equation is

$$u'(c) = \beta u'(\lambda c)(1 - \delta + r).$$

From before, the competitive zero-profit/first-order condition for capital services is

$$r = b\alpha (g(k, 2)/k)^{1-\rho}.$$

The condition for labour services, with the substitution $w = c + k$ from the young-person budget constraint, is

$$c + k = b(1 - \alpha)(g(k, 2)/2)^{1-\rho}.$$

The final equation is goods-market clearing:

$$(1 + \lambda)c + \delta k = bg(k, 2).$$

We can solve these four equations for c, k, λ and r if we assume a functional form for u. But it's clear by inspection that when b is large, r is large, and hence λ is large, i.e. the consumption profile is steep as claimed.

Postscript   We assumed as usual that the household operates a technology in its back yard for converting output goods into fresh capital. This assumption is standard and is made for assumptions of parsimony: there's no point in adding a capital-producing sector unless something interesting that will change the equilibrium allocations is going on in capital production.

Nevertheless, Larry was keen to emphasise that people usually have something like the following in mind. They think of there being a capital-goods producing firm that can convert investment (output goods) i and undepreciated old capital x into new capital according to $h(k, i) = k + i$. It buys undepreciated capital $(1 - \delta)k_t$ from dying old people and buys i_t from the output-goods firm.

From general producer theory with input prices $p \in \mathbf{R}_+^n$, production function f, a unit output price and conditional factor demand $z(p) \in \mathbf{R}_+^n$, cost-minimisation implies that $p_i = \lambda f_i(z(p))$ for each input i, where the Lagrange multiplier λ is equal to marginal cost.[47] Since marginal cost is unity (because the price of new capital is unity) and since the marginal product of each input is unity, it follows that the equilibrium price of undepreciated capital $(1 - \delta)k_t$ must be unity as well. Capital-producing firms are therefore indifferent between all input/output choices, so no profit-maximisation conditions are needed for this firm; we just impose the market-clearing condition $k_{t+1} = (1 - \delta)k_t + i_t$.

---

[47]These are just the first-order conditions. The fact that λ equals marginal cost follows directly from an envelope theorem for constrained programs.

15   Wed 11 Nov 2015

15.1   Outline of today

Today we're covering Comin and Gertler (2006), a model with both research and development. An idea is an intermediate input, just as in Romer (1987). As time passes, new ideas are discovered according to a stochastic process which is thought of as 'basic research', beyond the control of firms. Once an idea has been discovered, a firm can invest resources in trying to develop it into a marketable good, with probability of success increasing in the amount invested. Finally, firms that own developed ideas produce intermediate goods that they sell on to the final-goods producer. Growth is driven exogenously by the march of science, and endogenously by industry's development of basic scientific ideas into practical recipes. Moreover, random fluctuations in the progress of basic science give rise to low-frequency economic fluctuations ('medium-term cycles').

Today, we first study the technologies employed in the production of final and intermediate goods. The main goal is to obtain the aggregate production function, which is needed for e.g. the resource constraint. We then study the technology that turns discovered but unimplemented ideas into practical production processes; this gives rise to a law of motion for the quantity of usable intermediate goods. Finally, we scale the model such that the scaled model has a steady state that corresponds to the balanced growth path of the unscaled model.

15.2   Comin and Gertler (2006)

The central role is played by ideas, which are modelled as points in $\mathbf{R}_+$ (as in Romer (1987)). At any given time t, ideas in $[0, A_t]$ are implemented (i.e. can be used to produce an intermediate good), ideas in $(A_t, Z_t]$ are discovered but not yet implemented, and ideas in (Z_t, ∞) remain undiscovered. Idea discovery is determined by scientific progress, which is viewed as taking place outside profit-seeking firms.

In particular, Z_t is governed by the exogenous stochastic process

$$Z_{t+1} = \exp(\chi_t) Z_t \quad \text{and} \quad \chi_{t+1} = (1 - \rho)\bar{\chi} + \rho\chi_t + u_{t+1}$$

where $\rho \in (-1, 1)$ and $\{u_t\}$ is a white noise process.[48] So the growth rate $\Delta \ln(Z_t)$ follows an AR(1) process with mean $\bar{\chi}$ and persistence coefficient

---

[48]A white noise process is a stochastic process that is independently and identically distributed over time and has mean zero.

ρ. To illustrate, suppose $u_t = 1$ and $u_{t+j} = 0$ afterwards. $\Delta \ln(Z_{t+1})$ rises by 1, so the growth rate is higher this period. If $\rho = 0$ then the growth rate immediately returns to $\bar{\chi}$. If $\rho > 0$ then $\Delta \ln(Z_{t+2})$ increases by $\rho > 0$, and so on down the line. Hence the growth rate remains high for several periods, only gradually levelling off as the effect of the shock wears off.[49]

15.2.1 Production environment

In each period t, the supply side of the economy is similar to Romer (1987). The final good is produced by a competitive firm by combining intermediate goods in $[0, A_t]$ using Dixit–Stiglitz technology. Each intermediate-goods firm hires capital and labour to produce using Cobb–Douglas technology. Intermediate-goods firms are monopolists in output but act competitively in factor markets.

Let $Y : \mathbf{R}_+ \to \mathbf{R}_+$ be the function such that $Y(j)$ is the quantity of good j used.[50] The final-goods producer has the Dixit–Stiglitz production function

$$F(Y) = \left[\int_0^{A_t} Y(j)^{1/\theta} \mathrm{d}j \right]^\theta$$

that converts inputs $Y(\cdot)$ into final goods, where $\theta > 1$. Note that final-goods production does not require labour, a difference from Romer (1987). As usual, the production function is unit homogeneous. Every active intermediate-goods producer $j \in [0, A_t]$ uses the production function

$$G(K, L) = K^\alpha L^{1-\alpha}$$

to convert (K, L) units of capital and labour into intermediate good.

Now we do some aggregation. Let $K_t(j)$, $L_t(j)$ and $Y_t(j)$ denote the capital hired, labour hired and output produced of firm j at time t. We use tildes to denote the averages:

$$\widetilde{K_t} := \frac{1}{A_t} \int_0^{A_t} K_t(j) \mathrm{d}j, \quad \widetilde{L}_t := \frac{1}{A_t} \int_0^{A_t} L_t(j) \mathrm{d}j, \quad \widetilde{Y}_t := \frac{1}{A_t} \int_0^{A_t} Y_t(j) \mathrm{d}j.$$

Suppose $K_t(j)/L_t(j)$ is the same across $j \in [0, A_t]$. (This will obviously happen in equilibrium.) Then using Euler's formula and the zero-homogeneity

---

[49]Larry feels intuitively that ρ should be zero. I really didn't understand why he felt this way, however.

[50]As mentioned when discussing the Romer (1987) model (section 12.2.1), $Y(j)$ is really the intensity of usage (Radon-Nikodým derivative of quantity used) rather than the quantity of usage. But this language quickly becomes unwieldy, so we'll call it 'quantity'.

of derivatives of a unit homogeneous function,

$$
\begin{aligned}
\widetilde{Y}_t &:= \frac{1}{A_t} \int_0^{A_t} G(K_t(j), L_t(j)) \mathrm{d}j \\
&= \frac{1}{A_t} \int_0^{A_t} \left[G_1 \left(\frac{K_t(j)}{L_t(j)}, 1 \right) K_t(j) + G_2 \left(\frac{K_t(j)}{L_t(j)}, 1 \right) L_t(j) \right] \mathrm{d}j \\
&= G_1 \left(\frac{K_t(j)}{L_t(j)}, 1 \right) \frac{1}{A_t} \int_0^{A_t} K_t(j) \mathrm{d}j + G_2 \left(\frac{K_t(j)}{L_t(j)}, 1 \right) \frac{1}{A_t} \int_0^{A_t} L_t(j) \mathrm{d}j \\
&= G_1 \left(\frac{\widetilde{K}_t}{\widetilde{L}_t}, 1 \right) \widetilde{K}_t + G_2 \left(\frac{\widetilde{K}_t}{\widetilde{L}_t}, 1 \right) \widetilde{L}_t \\
&= G \left(\widetilde{K}_t, \widetilde{L}_t \right) = \widetilde{K}_t^\alpha \widetilde{L}_t^{1-\alpha}.
\end{aligned}
$$

Now suppose that each input is used equally in final-goods production, i.e. $Y_t(j) = \widetilde{Y}_t \; \forall j \in [0, A_t]$. Then the final-goods production technology yields

$$
Y_t = \left[\int_0^{A_t} \widetilde{Y}_t^{1/\theta} \mathrm{d}j \right]^\theta = \widetilde{Y}_t \left[\int_0^{A_t} \mathrm{d}j \right]^\theta = \widetilde{Y}_t A_t^\theta = \widetilde{K}_t^\alpha \widetilde{L}_t^{1-\alpha} A_t^\theta.
$$

Write K_t and L_t for aggregate capital and labour services:

$$
K_t := \int_0^{A_t} K_t(j) \mathrm{d}j = A_t \widetilde{K}_t \quad \text{and} \quad L_t := \int_0^{A_t} L_t(j) \mathrm{d}j = A_t \widetilde{L}_t.
$$

Then we can write

$$
Y_t = \left(\frac{K_t}{A_t} \right)^\alpha \left(\frac{L_t}{A_t} \right)^{1-\alpha} A_t^\theta = A_t^{\theta-1} K_t^\alpha L_t^{1-\alpha}.
$$

The right-hand side can be viewed as the aggregate production function of the economy. Define

$$
z_t := A_t^{(\theta-1)/(1-\alpha)}.
$$

By substituting for A_t to obtain $Y_t = K_t^\alpha (z_t L_t)^{1-\alpha}$, we see that z_t is labour-augmenting total factor productivity in this economy. So we'll definitely have growth in this economy if the frontier A_t of developed ideas is expanding in equilibrium.

15.2.2 Idea development

Next, we look at the environment facing firms whose ideas lie in $(A_t, Z_t]$. A firm $j \in (A_t, Z_t]$ that invests a quantity of resources $h_t(j)$ in trying to develop its idea has a probability $\lambda(h_t(j) A_t / b K_t)$ of success, where $\lambda : \mathbf{R}_+ \to [0, 1]$

is a strictly increasing, strictly concave function. (When required, we will use the functional form $\lambda(x) = \min\{x^\sigma, 1\}$ for $\sigma \in (0,1)$.) The additional terms A_t/bK_t in the argument are there to ensure the existence of a balanced growth path (on which the probability of success is constant).[51] The excuse for including A_t that is given is that development is easier when the stock of developed ideas is large, but this is a highly controversial claim. There is no excuse for including bK_t besides ensuring balanced growth. K_t takes the trend out of $h_t(j)A_t$, and b is chosen so that we can have our favourite success probability on the balanced growth path.

Now suppose that $h_t(j) = h_t \; \forall j \in (A_t, Z_t]$. Then 'by a law of large numbers', A_t evolves 'deterministically'[52] according to

$$A_{t+1} = A_t + \lambda\left(\frac{h_t A_t}{bK_t}\right)(Z_t - A_t).$$

15.2.3 Scaling the model

We seek a scaling of the model such that the scaled model has a steady state. In a stochastic model like this one, a (nonstochastic) steady state is simply a stationary point of the deterministic dynamics obtained by replacing all random variables with their means.[53] In this case, we replace u_t with 0 for each t. Then $\chi_t = \bar{\chi}$ (almost surely).

To find an appropriate scaling, begin by dividing the law of motion for A_t by Z_t:

$$\frac{Z_{t+1}}{Z_t}\frac{A_{t+1}}{Z_{t+1}} = \frac{A_t}{Z_t} + \lambda\left(\frac{h_t A_t}{bK_t}\right)\left(1 - \frac{A_t}{Z_t}\right).$$

Define $a_t := A_t/Z_t$, the fraction of discovered ideas that have been developed. Also define $h_{A,t} := h_t A_t/z_t$ and $o_{z,t} := bK_t/z_t$. (Recall that $z_t :=$

---

[51] Aside: the space of balanced growth models is actually very large. The existence of a balanced growth path requires only that there exist a scaling that gets rid of change over time. For example, question 3 on homework 6 shows how secular trends can occur on a balanced growth path.

[52] It is not clear what theorem Larry had in mind here. The strong law of large numbers for continua actually fails in many measure spaces; see e.g. Judd (1985). Perhaps the most standard LLNs for continua in economics are Anderson's (1991), which relies on nonstandard analysis, and Uhlig's (1996), which uses the Pettis integral instead of Lebegue's. There are many others, e.g. those of Yeneng Sun and coauthors (Y. Sun, 2006; Y. Sun & Zhang, 2009; Duffie & Sun, 2012; He, Sun & Sun, 2016) and that of Green (1994).

[53] Apparently, this definition is motivated by the perturbation method for stochastic methods, which involves a Taylor approximation around the nonstochastic steady state.

$A_t^{(\theta-1)/(1-\alpha)}$.) Then we obtain the law of motion

$$\exp(\chi_t)a_{t+1} = a_t + \lambda\left(\frac{h_{A,t}}{o_{z,t}}\right)(1 - a_t).$$

In a nonstochastic steady state, $\chi_t = \bar{\chi}$, so the law of motion for a_{t+1} above becomes deterministic as desired. In steady state, we must therefore have

$$\exp(\bar{\chi})a = a + \lambda\left(\frac{h_A}{o_z}\right)(1 - a).$$

Having a_t constant simply means that A_t grows at rate $\bar{\chi}$, the same rate as Z_t. Since

$$h_{A,t} = \frac{h_t A_t}{z_t} = h_t A_t^{(2-\alpha-\theta)/(1-\alpha)},$$

steady state requires h_t to grow at rate $\bar{\chi}(2 - \alpha - \theta)/(1 - \alpha)$. For a plausible calibration, $\alpha \simeq 1/3$ and $\theta < 5/3$.[54] In that case, h_t will be declining on the balanced growth path. Finally, $o_{z,t} = bK_t/z_t$, so K_t must grow at the same rate as $z_t = (A_t)^{(\theta-1)/(1-\alpha)}$, namely $\bar{\chi}(\theta - 1)/(1 - \alpha)$.

Now let's look at the resource constraint of the economy (i.e. market clearing for the final good):

$$C_t + K_{t+1} - (1 - \delta)K_t + \int_{A_t}^{Z_t} h_t \mathrm{d}j = Y_t.$$

$\int_{A_t}^{Z_t} h_t \mathrm{d}j$ is the total use of resources for development. Manipulating this object a bit yields

$$\int_{A_t}^{Z_t} h_t \mathrm{d}j = h_t(Z_t - A_t) = h_t Z_t(1 - a_t) = \frac{h_t A_t}{z_t}\frac{Z_t}{A_t}(1 - a_t)z_t$$
$$= h_{A,t}(a_t^{-1} - 1)z_t.$$

So the quantity of resources devoted to development grows at the same rate as z_t on the balanced growth path, i.e. rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$. Moreover, we have already seen that K_t grows at this rate. This suggests that C_t and Y_t grow at the same rate, so that we ought to scale them by z_t. Writing $c_t := C_t/z_t$, $k_t := K_t/z_t$ and $y_t := Y_t/z_t$, the resource constraint becomes

$$c_t + \exp(\chi_t)k_{t+1} - (1 - \delta)k_t + h_{A,t}\left(a_t^{-1} - 1\right) = y_t.$$

[54]$\theta \geq 5/3$ is implausible because intermediate-goods producers will charge a markup of θ over marginal cost. A markup of $5/3$ would be very large.

In steady state, we have

$$c + (\exp(\bar{\chi}) + 1 - \delta)k + h_A \left(a^{-1} - 1\right) = y.$$

So we have a balanced growth path on which consumption, the capital stock and output all grow at rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$. From the perspective of growth theory, this will therefore be the crucial number in this model.

16 Mon 16 Nov 2015

16.1 Outline of today

Today we'll finish covering Comin and Gertler (2006). We first study the behaviour of firms. The final good is produced by a competitive representative firm that buys intermediate inputs and sells a homogeneous final good used for consumption and investment. Each developed intermediate good is sold by a monopolist. Firms that own discovered but undeveloped ideas choose how much to invest in trying to develop their idea.

We then move on to pricing stocks in each of the three kinds of firms: those with implemented production processes, those with as of yet undeveloped ideas, and those who own the rights to as of yet undiscovered ideas.

16.2 Comin and Gertler (2006) again

16.2.1 Production of the final good

Recall that the final-goods producer uses the Dixit–Stiglitz production function

$$F(Y) = \left[\int_0^{A_t} Y(j)^{1/\theta} \mathrm{d}j \right]^{\theta}$$

that converts inputs $Y(\cdot)$ into final goods, where $\theta > 1$. Let P_t denote the price of the final good, and let $P_t(j)$ be the price of the jth intermediate good. The final-goods firm is competitive in output and factor markets; it therefore solves

$$\max_{Y_t \geq 0, Y_t(\cdot) \geq 0} \left\{ P_t Y_t - \int_0^{A_t} P_t(j) Y_t(j) \mathrm{d}j : Y_t = \left[\int_0^{A_t} Y_t(j)^{1/\theta} \mathrm{d}j \right]^{\theta} \right\}.$$

The first-order condition w.r.t. $Y_t(j)$ is

$$P_t Y_t^{(\theta-1)/\theta} Y_t(j)^{(1-\theta)/\theta} = P_t(j).$$

This can be rearranged to yield the demand curve and inverse demand curve:

$$\frac{P_t(j)}{P_t} = \left(\frac{Y_t(j)}{Y_t} \right)^{-(\theta-1)/\theta} \quad \text{and} \quad Y_t(j) = Y_t \left(\frac{P_t(j)}{P_t} \right)^{-\theta/(\theta-1)}.$$

Substituting back into the constraint, we obtain

$$1 = \left[\int_0^{A_t} \left(\frac{Y_t(j)}{Y_t} \right)^{1/\theta} \mathrm{d}j \right]^{\theta} = \left[\int_0^{A_t} \left(\frac{P_t(j)}{P_t} \right)^{-1/(\theta-1)} \mathrm{d}j \right]^{\theta}.$$

119

Rearranging, we obtain the equilibrium price of final output as

$$P_t = \left[\int_0^{A_t} P_t(j)^{-1/(\theta-1)} \mathrm{d}j \right]^{-(\theta-1)}.$$

16.2.2 Production of intermediate goods

The firm producing intermediate good $j \in [0, A_t]$ is a monopolist in the output market, but acts competitively in factor markets. She therefore maximises profit subject to her demand curve and technological constraint. We'll separate this into the choice of the cost-minimising inputs $(K_t(j), L_t(j))$ to produce a given level of output $Y_t(j)$ and the profit-maximising choice of output $Y_t(j)$ and price $P_t(j)$.

Let r_t and $\widetilde{w}_t$ be factor prices denominated in units of the final good. Then the cost-minimisation problem for producing $Y_t(j)$ is

$$\min_{(K,L)\geq 0} \left\{ r_t K + \widetilde{w}_t L : K^\alpha L^{1-\alpha} = Y_t(j) \right\}.$$

Call the Lagrange multiplier on the constraint $\widetilde{s}_t \geq 0$. By the envelope theorem, $\widetilde{s}_t$ is marginal cost in units of the consumption good, so this is the object we'd like to solve for.

The first-order conditions are

$$r_t = \widetilde{s}_t \alpha K_t(j)^{\alpha-1} L_t(j)^{1-\alpha}$$
$$\widetilde{w}_t = \widetilde{s}_t (1-\alpha) K_t(j)^\alpha L_t(j)^{-\alpha}.$$

As usual, $(r_t, \widetilde{w}_t)$ must lie in the factor price frontier in order to guarantee that both equations can hold. Dividing the first-order conditions by each other,

$$\frac{r_t}{\widetilde{w}_t} = \frac{\alpha}{1-\alpha} \frac{L_t(j)}{K_t(j)}.$$

Substituting in the first-order condition for capital,

$$\frac{\alpha}{1-\alpha} \frac{L_t(j)}{K_t(j)} \widetilde{w}_t = r_t = \widetilde{s}_t \alpha K_t(j)^{\alpha-1} L_t(j)^{1-\alpha} = \widetilde{s}_t \alpha \frac{L_t(j)}{K_t(j)} \left(\frac{K_t(j)}{L_t(j)} \right)^\alpha$$

$$= \widetilde{s}_t \alpha \frac{L_t(j)}{K_t(j)} \left(\frac{\alpha}{1-\alpha} \frac{\widetilde{w}_t}{r_t} \right)^\alpha = \widetilde{s}_t \frac{L_t(j)}{K_t(j)} \frac{\alpha}{1-\alpha} \alpha^\alpha (1-\alpha)^{1-\alpha} \widetilde{w}_t^\alpha r_t^{-\alpha}.$$

Cancelling terms and rearranging gives

$$\widetilde{s}_t = \left(\frac{r_t}{\alpha} \right)^\alpha \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha}.$$

120

Note that marginal cost is constant.

The profit-maximisation problem of firm $j \in [0, A_t]$ is therefore

$$\max_{(Y,P) \geq 0} \left\{ PY - P_t \tilde{s}_t Y : Y = Y_t \left(P/P_t \right)^{-\theta/(\theta-1)} \right\}$$

Output and price are chosen to maximise profit, where costs are captured by the cost function (assuming no fixed cost), subject to the demand curve. It is important that j cannot influence P_t via her choice of $P_t(j)$; this is evident from the expression for P_t we derived at the end of the last section. This means that there are no strategic interactions between monopolists, keeping things simple.

Substituting out Y, the objective becomes

$$(P - P_t \tilde{s}_t) Y_t \left(P/P_t \right)^{-\theta/(\theta-1)} = [P - P_t \tilde{s}_t] Y_t P_t^{\theta/(\theta-1)} P^{-\theta/(\theta-1)}.$$

So the first-order condition for the optimal price $P_t(j)$ is

$$Y_t P_t^{\theta/(\theta-1)} P_t(j)^{-\theta/(\theta-1)}$$
$$- \frac{\theta}{\theta-1} [P_t(j) - P_t \tilde{s}_t] Y_t P_t^{\theta/(\theta-1)} P_t(j)^{-\theta/(\theta-1)} P_t(j)^{-1} = 0.$$

Cancelling and rearranging,

$$P_t(j) = \frac{\theta}{\theta-1} [P_t(j) - P_t \tilde{s}_t],$$

or

$$P_t(j) = \theta P_t \tilde{s}_t.$$

That is, firm j charges a constant markup θ over marginal cost $P_t \tilde{s}_t$.

Since this rule holds for every firm $j \in [0, A_t]$, the formula for P_t from the previous section gives

$$P_t = A_t^{-(\theta-1)} \theta P_t \tilde{s}_t,$$

so that

$$\theta \tilde{s}_t = A_t^{\theta-1}.$$

Since A_t will be growing in equilibrium and $\theta > 1$, this means that marginal cost grows over time.

Why is marginal cost rising? One heuristic way to see it is as follows. Once we put a household in the model, we'll have a first-order condition of the form

$$\tilde{w}_t = -\frac{u_2(C_t, L_t)}{u_1(C_t, L_t)} = -\psi L_t^\zeta C_t.$$

Since L_t must be fixed on a balanced growth path while C_t grows at a positive rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$, it's clear $\widetilde{w}_t$ will be growing (at the same rate for this utility specification). While we haven't said anything about the other factor price r_t, this at least suggests that marginal cost should be rising.

16.2.3　The stock market value of firms

So far, we haven't done much asset pricing. Instead, we've been focused on growth. It's not clear to me why we're suddenly pricing shares in firms instead of looking at e.g. the convergence properties of this model. Anyway, let's do it!

We first want the market value of active intermediate-goods firm $j \in [0, A_t]$. We begin by figuring out what their profits are in equilibrium. Rearranging the optimal price and using $\theta \widetilde{s}_t = A_t^{\theta-1}$, we have

$$\frac{P_t(j)}{P_t} = A_t^{\theta-1}.$$

We've seen that the profit of an intermediate-goods firm that charges price P when the aggregates (over which it has no control) are $(P_t, Y_t, \widetilde{s}_t)$ is

$$[P - P_t\widetilde{s}_t]\, Y_t P_t^{\theta/(\theta-1)} P^{-\theta/(\theta-1)} = \left[\frac{P}{P_t} - \widetilde{s}_t\right] P_t Y_t \left(\frac{P}{P_t}\right)^{-\theta/(\theta-1)}.$$

Let $P_t\pi(P_t, A_t)$ denote the maximised profit of an intermediate-goods firm as a function of aggregates, so that $\pi(P_t, A_t)$ is profit in units of the final good. Using $P_t(j)/P_t = A_t^{\theta-1}$ and $\widetilde{s}_t = A_t^{\theta-1}/\theta$, we obtain

$$P_t\pi(P_t, A_t) = \left[A_t^{\theta-1} - \frac{A_t^{\theta-1}}{\theta}\right] P_t Y_t \left(A_t^{\theta-1}\right)^{-\theta/(\theta-1)}$$
$$= \tfrac{\theta-1}{\theta} P_t Y_t A_t^{-1},$$

so that $\pi(P_t, A_t) = \frac{\theta-1}{\theta} Y_t A_t^{-1}$. Total profits of the intermediate-goods industry (in units of the final good) are then

$$\int_0^{A_t} \pi(P_t, A_t)\mathrm{d}j = \tfrac{\theta-1}{\theta} Y_t.$$

Now let's look at how profits behave on the balanced growth path. In terms of scaled variables,

$$\pi(P_t, A_t) = \tfrac{\theta-1}{\theta} y_t z_t A_t^{-1} = \tfrac{\theta-1}{\theta} y_t A_t^{(\alpha+\theta-2)/(1-\alpha)},$$

122

which grows at rate $\bar{\chi}(\alpha + \theta - 2)/(1 - \alpha)$ on the balanced growth path. As mentioned earlier, this is a negative rate of growth for plausible parameter values. Industry profit can be written $\frac{\theta-1}{\theta} y_t z_t$, which grows at (positive) rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$ on the balanced growth path. So total profits grow, but the profits of individual intermediate-goods manufacturers declines.

Before pricing a share in firm j, consider a simpler asset: a claim on all of firm j's profit in period $t + k$ only. If trading takes place in period t, then as usual, the Arrow–Debreu equilibrium price of this asset will be

$$\beta^k \frac{u'(C_{t+k})}{u'(C_t)} \frac{\mu(h^{t+k})}{\mu(h^t)} \pi(P_{t+k}, A_{t+k})$$

where u is the household's period utility function,[55] β is her discount factor, and μ is the probability measure on histories.[56] (See Ljungqvist and Sargent (2012, ch. 13) if it's not clear why this is the equilibrium price.)

Now consider the equilibrium price of a portfolio of securities containing claims to all of firm j's profit at every history h^{t+k} of length $t + k$, for some $k \in \mathbf{Z}_+$. It is

$$\sum_{h^{t+k}|h^t} \beta^k \frac{u'(C_{t+k})}{u'(C_t)} \frac{\mu(h^{t+k})}{\mu(h^t)} \pi(P_{t+k}, A_{t+k}) = \mathbf{E}_t \left(\beta^k \frac{u'(C_{t+k})}{u'(C_t)} \pi(P_{t+k}, A_{t+k}) \right)$$

where summation over $h^{t+k}|h^t$ means summation over all histories h^{t+k} that agree with h^t in the first t entries, and $\mathbf{E}_t$ is sloppy shorthand for the expectation conditional on history h^t.

Finally, consider a portfolio of securities containing claims to all of firm j's profit at every history h^{t+k} of length $t+k$ for every $k \in \mathbf{R}_+$. This portfolio would cost

$$\sum_{k=0}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u'(C_{t+k})}{u'(C_t)} \pi(P_{t+k}, A_{t+k}) \right).$$

Suppose now that we add a new asset to our Arrow–Debreu financial economy, viz. stocks in firm j. Owning all the stock of firm j just means getting all of firm j's profit at every history h^{t+k} of length $t + k$ for every $k \in \mathbf{R}_+$, so this asset is equivalent to the portfolio we just priced. In Arrow–Debreu

---

[55] Actually, we would have labour supply L_t as an argument in utility if we were treating the household sector properly.

[56] We're not being very explicit about the uncertainty here. Formally, the state variable of the Markov chain governing uncertainty here is χ_t, so $h^t = (\chi_0, \ldots, \chi_t)$. The measure μ is defined from the law of the stochastic process $\{\chi_t\}$, which depends on the parameters $\bar{\chi}$ and ρ and the distribution of u_t. We've also assumed that $\mu(h^t) > 0$ for every history; without this, we'd have to directly use the conditional probability distribution.

equilibrium, the value V_t of firm $j \in [0, A_t]$ at time t must therefore be equal to the price of this portfolio:

$$V_t = \sum_{k=0}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u'(C_{t+k})}{u'(C_t)} \pi(P_{t+k}, A_{t+k}) \right).$$

A more formal argument for this would use a no-arbitrage condition.[57]

We can express V_t recursively as follows:

$$V_t = \pi(P_t, A_t) + \sum_{k=1}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u'(C_{t+k})}{u'(C_t)} \pi(P_{t+k}, A_{t+k}) \right)$$

$$= \pi(P_t, A_t) + \mathbf{E}_t \Big(\beta \frac{u'(C_{t+1})}{u'(C_t)}$$

$$\times \sum_{k=1}^{\infty} \beta^{k-1} \frac{u'\left(C_{(t+1)+(k-1)}\right)}{u'\left(C_{t+1}\right)} \pi\left(P_{(t+1)+(k-1)}, A_{(t+1)+(k-1)}\right) \Big)$$

$$= \pi(P_t, A_t) + \mathbf{E}_t \Big(\beta \frac{u'(C_{t+1})}{u'(C_t)}$$

$$\times \sum_{k=0}^{\infty} \beta^k \frac{u'\left(C_{(t+1)+k}\right)}{u'\left(C_{t+1}\right)} \pi\left(P_{(t+1)+k}, A_{(t+1)+k}\right) \Big)$$

$$= \pi(P_t, A_t) + \mathbf{E}_t \left(\beta \frac{u'(C_{t+1})}{u'(C_t)} V_{t+1} \right).$$

This is sometimes called the asset equation. Notice that is looks a lot like a Bellman equation, but that the discount factor is the random variable $\beta u'(C_{t+1})/u'(C_t)$ rather than (say) β alone. $\beta u'(C_{t+1})/u'(C_t)$ is therefore called the stochastic discount factor. (Or pricing kernel.)

Now let's scale the contents of the asset equation to see what happens to V_t on the balanced growth path. Define $V_{A,t} := V_t A_t / z_t$, $u'_z(C_t) := u'(C_t) z_t$

---

[57]Here's a sketch. If V_t were higher than the price of the equivalent portfolio of Arrow securities, then a trader could make unbounded profit with probability 1 by shorting the firm and buying the equivalent portfolio (over and over again). (This is called an arbitrage trade.) If the firm were undervalued, the opposite trading strategy would make unbounded profit. In either case, the trader's problem has no solution, so this cannot happen in equilibrium.

and $\pi_z(P_t, A_t) := \pi(P_t, A_t)A_t/z_t$. Then

$$
\begin{aligned}
V_{A,t} &= \pi_z(P_t, A_t) + \mathbf{E}_t \left(\beta \frac{u'(C_{t+1})}{u'(C_t)} V_{A,t+1} \frac{z_{t+1}}{A_{t+1}} \right) \frac{A_t}{z_t} \\
&= \pi_z(P_t, A_t) + \mathbf{E}_t \left(\beta \frac{u'_z(C_{t+1})}{u'_z(C_t)} V_{A,t+1} \frac{A_t}{A_{t+1}} \right) \\
&= \pi_z(P_t, A_t) + \mathbf{E}_t \left(\beta \frac{u'_z(C_{t+1})}{u'_z(C_t)} V_{A,t+1} \frac{a_t}{a_{t+1}} \frac{Z_t}{Z_{t+1}} \right) \\
&= \pi_z(P_t, A_t) + \mathbf{E}_t \left(\beta \frac{u'_z(C_{t+1})}{u'_z(C_t)} V_{A,t+1} \frac{a_t}{a_{t+1}} \exp(-\chi_t) \right).
\end{aligned}
$$

We won't prove it, but in nonstochsatic steady state (for log utility, I guess?), all objects in this expression are time-invariant as required. Hence the stock market value V_t of an individual firm $j \in [0, A_t]$ grows at the same rate as A_t/z_t on the balanced growth path. We know from before that this rate is $\bar{\chi}(\alpha + \theta - 2)/(1 - \alpha)$, which is negative for plausible parameter values. The total stock-market value of active intermediate-good firms is

$$
\int_0^{A_t} V_t \mathrm{d}j = V_t A_t,
$$

so the value of this stock market index grows at the same rate as z_t, i.e. the (positive) rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$.

Now we move on to firms in $(A_t, Z_t]$, i.e. firms that have not yet developed their ideas into practical production processes. Call the value of such a firm J_t. (Equivalently, you can think of J_t as the value of an undeveloped idea.) This time, we go straight to the asset equation:

$$
J_t = \max_{h_t \geq 0} \left\{ -h_t \right.
$$
$$
\left. + \mathbf{E}_t \left(\beta \frac{u'(C_{t+1})}{u'(C_t)} \left[\lambda \left(\frac{A_t h_t}{bK_t} \right) V_{t+1} + \left(1 - \lambda \left(\frac{A_t h_t}{bK_t} \right) \right) J_{t+1} \right] \right) \right\}.
$$

There is no immediate gain to owning an undeveloped idea, but resources h_t are expended to try to develop it. With probability $\lambda(A_t h_t / bK_t)$, development succeeds, in which case the firm will be worth V_{t+1} starting next period. With the complementary probability, development will fail and the value of the firm will be J_{t+1}.

Now let's scale to see what happens on the balanced growth path. Assume an interior optimum for h_t. The first-order condition is

$$
1 = \mathbf{E}_t \left(\beta \frac{u'(C_{t+1})}{u'(C_t)} \lambda' \left(\frac{A_t h_t}{bK_t} \right) \frac{A_t}{bK_t} (V_{t+1} - J_{t+1}) \right).
$$

Using our previous definitions and $J_{A,t} \coloneqq J_t A_t / z_t$,

$$1 = \mathbf{E}_t \left(\beta \frac{u_z'(C_{t+1})}{u_z'(C_t)} \frac{z_t}{z_{t+1}} \lambda' \left(\frac{h_{A,t}}{o_{z,t}} \right) \frac{A_t}{bK_t} (V_{A,t+1} - J_{A,t+1}) \frac{z_t}{A_t} \right)$$

$$= \mathbf{E}_t \left(\beta \frac{u_z'(C_{t+1})}{u_z'(C_t)} \frac{z_t}{z_{t+1}} \lambda' \left(\frac{h_{A,t}}{o_{z,t}} \right) \frac{1}{bK_t/z_t} (V_{A,t+1} - J_{A,t+1}) \right)$$

$$= \mathbf{E}_t \left(\beta \frac{u_z'(C_{t+1})}{u_z'(C_t)} \left(\frac{a_t}{a_{t+1}} \frac{Z_t}{Z_{t+1}} \right)^{\frac{\theta-1}{1-\alpha}} \lambda' \left(\frac{h_{A,t}}{o_{z,t}} \right) o_{z,t}^{-1} (V_{A,t+1} - J_{A,t+1}) \right)$$

$$= \mathbf{E}_t \left(\beta \frac{u_z'(C_{t+1})}{u_z'(C_t)} \left(\frac{a_t}{a_{t+1}} \right)^{\frac{\theta-1}{1-\alpha}} \exp\left(-\chi_t \frac{\theta-1}{1-\alpha} \right) \right.$$

$$\left. \times \lambda' \left(\frac{h_{A,t}}{o_{z,t}} \right) o_{z,t}^{-1} (V_{A,t+1} - J_{A,t+1}) \right).$$

On the balanced growth path, all of these objects are constant. So $J_{A,t}$ grows at the same rate as $V_{A,t}$, viz. the rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$.

Finally we look at the value of undiscovered ideas $j \in (Z_t, \infty)$. Somewhat more formally, we study the value of firms that have ownership of an idea that has not yet been discovered. This isn't as weird as it sounds: part of the value of a company like e.g. Apple is the ideas that their clever employees will come up with in the future. We'll only look at the aggregate value of undiscovered ideas. If an idea is discovered at t, it will have value J_t at t. Hence the value of all the firms that own undiscovered ideas is

$$z_t v_t^f = \sum_{k=1}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u'(C_{t+k})}{u'(C_t)} J_{t+k} (Z_{t+k} - Z_{t+k-1}) \right).$$

Scaling by z_t,

$$v_t^f = z_t^{-1} \sum_{k=1}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u_z'(C_{t+k})}{u_z'(C_t)} \frac{z_t}{z_{t+k}} J_{t+k} \frac{z_{t+k}}{A_{t+k}} \left(\frac{Z_{t+k}}{Z_{t+k-1}} - 1 \right) Z_{t+k-1} \right)$$

$$= \sum_{k=1}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u_z'(C_{t+k})}{u_z'(C_t)} J_{A,t+k} \left(\frac{Z_{t+k}}{Z_{t+k-1}} - 1 \right) \frac{Z_{t+k}}{A_{t+k}} \frac{Z_{t+k-1}}{Z_{t+k}} \right)$$

$$= \sum_{k=1}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u_z'(C_{t+k})}{u_z'(C_t)} J_{A,t+k} \left(1 - \frac{Z_{t+k-1}}{Z_{t+k}} \right) \frac{Z_{t+k}}{A_{t+k}} \right)$$

$$= \sum_{k=1}^{\infty} \mathbf{E}_t \left(\beta^k \frac{u_z'(C_{t+k})}{u_z'(C_t)} J_{A,t+k} (1 - \exp(-\chi_{t-1})) a_{t+k}^{-1} \right).$$

So on the balanced growth path, the value of all undiscovered ideas grows at rate $\bar{\chi}(\theta - 1)/(1 - \alpha)$, the same rate as z_t.

After all that exhausting work, what do we get? The trend rate of growth of all three kinds of firms is $\bar{\chi}(\theta - 1)/(1 - \alpha)$, the same as the growth rate of GDP. But suppose that new firms are only gradually admitted into the stock market, so that the stock market contains disproportionately many older active firms (low j). Since the value of active firms is declining over time, this means that we may see the stock market decline even though GDP growth and the value of all firms is growing quickly! Jovanović and Rousseau (2001) suggest that this might explain why the stock market did so badly during the Great Depression even though a great deal of scientific and industrial progress occurred during that era. Hornstein and Krusell (1996) make a similar argument, and there are apparently several other Jovanović–Rousseau and Hornstein–Krusell papers on this topic.

Postscript on asset pricing. The way we price assets here is a little weird. Since all consumers are identical, there is never any trade in equilibrium! This is in line with the 'consumption-based', general-equilibrium asset pricing theory of Lucas (1978) (the 'Lucas tree' paper.) Larry says that some people, e.g. Cochrane, think that the next big thing in finance is asset pricing in models in which trade occurs in equilibrium. This can be achieved by having idiosyncratic risk and hence idiosyncratic hedging requirements. It can also be achieved with heterogeneous beliefs, but this is much more difficult. Larry didn't go into detail, but I imagine that the problem is that traders can learn about other traders' beliefs by observing prices. This leads to rather complicated learning. Larry also said that there are two types of heterogenous-beliefs models: those with heterogeneous private information, and those in which agents have 'different models of the world'. I'm not sure what this means, but perhaps the former involves a common prior and private signals (different information sets), while the latter has heterogeneous priors and public signals?

17   Wed 18 Nov 2015

17.1   Outline of today

Today, we introduce business cycles. We'll look at some stylised facts about business cycles and examine the empirical successes and failures of basic real business cycle models.

17.2   Business cycles

Business cycles are the ups and downs in economic activity. At least for today's lecture, 'economic activity' means GDP, consumption, investment and employment. Some of the stylised facts about business cycles are as follows.

(1) Persistence. Both booms and busts last for a fairly long time.

(2) Sectoral comovement. Economic activity in most sectors, even very different ones, is highly correlated over time.

(3) Relative volatilities. Let σ_y, σ_c, σ_I and σ_n be the volatilities of log GDP, log consumption, log investment and log employment, respectively. In most developed economies, $\sigma_c \simeq \frac{1}{2}\sigma_y$, $\sigma_I \simeq 3\sigma_y$ and $\sigma_n \simeq \sigma_y$.[58]

(4) International comovement. Economic activity is highly correlated across countries. This is considered a puzzle because most models require shocks in individual countries to be highly correlated to generate international comovement.[59]

The positive goal of business cycle theory is to explain these and other empirical facts about business cycles. Business cycle theory also aims to answer normative questions. The main one is whether fluctuations are efficient, and if not, what should be done about it.

Relation to growth theory.   Historically, business cycles and growth have been viewed as distinct phenomena that should be studied independently. But there have been periods in which there was interest in unified theories of growth and cycles, e.g. the 1990. One example is Matsuyama (1999), who

---

[58]Apparently, this does not hold for the UK. Larry isn't sure why, but thinks there must be something wrong with their data.

[59]Strategic complementarities could generate comovement without correlated shocks, but the degree of strategic complementarity has to be implausibly strong to generate sizeable comovement.

modified the Romer (1987) model by replacing the fixed cost by a sunk cost. He showed that for some (very special) parameter values, this deterministic model produces cycles in equilibrium. But Larry says that at the moment, there is little interest in linking up growth and cycles.

To motivate the topic further, let's do some history of thought, beginning with the views of the American public. When it happened, the Great Depression was perceived as a period of inefficiently low output by many. The view that recessions are inefficient persisted into the 1970s. But the highly unpopular 'stagflation' (fairly high inflation and sluggish economic activity) of the 1970s and early 1980s, followed by the highly popular (!) deflationary recession triggered by Volcker's monetary policy starting in 1982, seems to have led to a widespread view that recessions needn't be a bad thing at all.

Aside on the costs of inflation. The popularity of Volcker's deflationary monetary policy suggests that the public really dislikes inflation, but some creativity is needed in order to explain this in an economic model. The standard (new Keynesian) explanation these days is that individual prices are sticky, so when fundamentals change, price changes are staggered across industries. Since prices are strategic complements (think of Bertrand), when inflation is high, a firm that cannot (or finds it too costly to) change prices frequently will sharply increase its price when it gets the change to keep up with inflation. This puts relative prices out of kilter, leading to resource misallocation. Apparently, this results in a surprisingly large welfare loss (Christiano, 2015).

So how did the views of macroeconomists evolve over the same period? Economists generally accepted that the Great Depression was inefficient, especially following the publication of Keynes (1936). This view stuck around for quite a long time, first advocated by people like Hicks and Kaldor in the UK, then by Samuelson and Solow at MIT. For some reason, this Keynesianism was usually combined with a distaste for mathematics.[60] But in the 1970s, a group of 'crazies' at Carnegie–Mellon University (Lucas, Prescott, Sargent, Sims and Wallace) who were not under the MIT spell started to build mathematical macro models using tools from general equilibrium theory.

---

[60]Keynes was quite taken with mathematical economics when he was younger (for example, he was forever badgering Frank Ramsey to write economics papers). But he seems to have soured on mathematics later in his life: he called Jan Tinbergen's pioneering work in econometrics 'black magic' and 'statistical alchemy', and the *General theory* (1936) pointedly avoids formalism. According to Larry, Samuelson and Solow basically stuck to this line, allowing for the use of simple mathematical tools but eschewing the heavy-duty machinery that Samuelson made use of in his other work.

Real business cycle theory emerged from this work; the classic papers are Long and Plosser (1983) and Prescott (1986).[61] Those who took these models seriously came to the conclusion that business cycles were efficient responses to changing fundamentals. Eventually, a 'new neoclassical synthesis' emerged that uses the tools of real business cycle theory to study models that do not have the strange welfare properties of the Carnegie–Mellon models.[62] This is the prevailing paradigm in macro today.

The Frisch–Slutsky paradigm. The mainstream macro paradigm of the last 30 years, from simple RBC models through medium-scale DSGE beasts with seven shocks and innumerable frictions (classic example: Smets and Wouters (2007)), is a special case of a more general modelling approach called the Frisch–Slutsky paradigm (Frisch, 1933; Slutsky, 1937). A model in the Frisch–Slutsky paradigm describes a system's behaviour using (1) a set of exogenous stochastic processes (impulses) and (2) a collection of response functions describing how endogenous variables are determined by the impulses (propagation mechanisms).[63] This paradigm is not unique to economics; many models in e.g. physics and biology have the same basic structure. In macro, the impulses are shocks to preferences, technology, endowments and policy (see Smets and Wouters (2007) for some examples). The response functions are given by objectives and constraints for each agent plus an equilibrium concept.

I mention this as background to something that Larry said in passing. When mainstream macroeconomists wish to explain some aspect of the behaviour of aggregate economic variables, they formulate their answer in terms of the Frisch–Slutsky paradigm. The questions to be answered are

(1) Impulse: what shock causes the phenomenon?

(2) Propagation mechanism: what kinds of objectives, constraints and equilibrium concept are required in order for the equilibrium response to this shock to replicate the phenomenon?

Some heterodox types (Larry mentioned PSE in particular) think that a Frisch–Slutsky-paradigm explanation is not much of an explanation. Loosely, their complaint is that because the impulses are always assumed to be

---

[61]People often also mention Kydland and Prescott (1982) here.

[62]The term 'new neoclassical synthesis' is from Goodfriend and King (1997). I'm not sure how widespread it is, but I think it's apt.

[63]The impulse/propagation language comes from Frisch (1933), which is, incidentally, a very insteresting read.

unobservable in macroeconomics, a mainstream macroeconomist 'explains' macroeconomic phenomena by simply assuming that the unobserved shocks behave in whatever way is required in order for her favourite model to replicate the phenomenon. These heterodox macro people are more convinced by models whose deterministic dynamics are able to replicate empirical phenomena, without recourse to unobserved shocks. This generally requires highly nonlinear dynamics, which quickly becomes intractable.

This nonlinear dynamics literature focuses both on models in which equilibrium dynamics are cylical and models in which they are chaotic.[64] Models with endogenous deterministic cycles were probably the prevailing theory in the early days of business cycle theory, e.g. Pigou (1927). There was a theoretical literature in the 1980s and early 1990s that picked up this strand, studying fairly standard deterministic macro models that exhibit chaotic and cyclical equilibrium dynamics; see e.g. Benhabib and Day (1982), Grandmont (1985), Bak, Chen, Scheinkman and Woodford (1993), Matsuyama (1999) and the mini-survey by Benhabib (2008). This line of inquiry is experiencing something of a revival, both from fairly mainstream macro types looking for analytical characterisations of volatility (La'O, 2015; Beaudry, Galizia & Portier, 2015) and from the 'agent-based modelling' crowd. The latter deal with the intractability of nonlinear dynamics (as well as rich heterogeneity) by giving up entirely on analytical results, choosing instead to simulate their models. Geanakoplos et al. (2012) gives a taste of this literature.

17.3   Real business cycle theory

In light of the new neoclassical synthesis, the real business cycle will be our baseline framework for the study of business cycles. Without bells of whistles, the basic real business cycle model is just the Solow I model with a stochastic equation of motion for TFP rather than a deterministic one, plus elastic labour supply. Because the first welfare theorem holds, we look at the planner's problem (which is often substantially easier in these settings). The representative household has expected utility preferences represented by

$$\mathbf{E}_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t, n_t) \right)$$

---

[64]We know from the Boldrin and Montrucchio (1986) theorem that both types of dynamics are perfectly consistent with dynamic optimisation.

where $\mathbf{E}_0$ is sloppy notation for expectation conditional on all information available at 0. The period-by-period resource constraint is

$$c_t + k_{t+1} - (1 - \delta)k_t \leq k_t^\alpha (z_t n_t)^{1-\alpha}.$$

TFP z_t evolves stochastically according to

$$z_t = \exp(\mu t + x_t)$$
$$x_t = \rho x_{t-1} + \varepsilon_t$$

where $|\rho| < 1$ and $\{\varepsilon_t\}$ is a white noise innovation process.[65] Equivalently, $\{\ln(z_t)\}$ follows an AR(1) process:

$$\ln(z_t) = \rho \ln(z_{t-1}) + \mu + \varepsilon_t.$$

x_t can be thought of as the detrended part of $\ln(z_t)$. That's it! The economy then fluctuates randomly around the Solow I balanced growth path.

Prescott (1986) takes this model to the data. He computes TFP z_t (using the formula $y_t = k_t^\alpha n_t^{1-\alpha}$ for $\alpha \simeq 1/3$ and estimates $\rho \simeq 0.95$. So measured TFP is highly persistent. As you might expect, high persistence in the exogenous forcing process gives rise to smooth, persistent stochastic processes for the endogenous variables y_t, c_t, I_t and n_t. Importantly, though, the model has almost no endogenous propagation mechanisms: all the persistence in the dynamics of endogenous stochastic processes is driven by the persistence of the forcing process $\{z_t\}$.[66]

The big accomplishment of the model is that it's able to match the relative volatilities of GDP, consumption and investment mentioned in the introduction. For the standard calibration $\alpha = 1/3$, $\delta = 0.08/4$ (for quarterly data) and the utility function $u(c, n) = \ln(c) + \gamma \ln(1 - n_t)$ with γ chosen to match average hours worked, the relative volatilities σ_c/σ_y and σ_I/σ_y are remarkably close to 0.5 and 3 as in the data.[67] This result is driven by the concavity of u. Since people don't like fluctuations in consumption, they smooth consumption over the business cycle by keeping consumption relatively stable in the face of shocks and letting investment fluctuate instead.

---

[65]I.e. for each t, ε_t has mean zero, is independent of ε_{t-j} and x_{t-j-1} for $j \in \mathbf{N}$, and has the same marginal distribution as ε_{t-j} for $j \in \mathbf{N}$.

[66]In contrast, the Comin and Gertler (2006) displays remarkable persistence even when the forcing process $\{\chi_t\}$ has no persistence. That's why they called the paper 'medium-term cycles', I guess.

[67]There are controversies hidden here. For one, the calibration of δ is always in dispute because depreciation is very hard to measure. Another tricky question is what the labour endowment is. (8 hours? 16? We wrote it as 1 above, but that was a normalisation!) There was a big debate about this in the 1980s.

The glaring failure of the model is that σ_n/σ_y is far smaller than the figure of about 1 that we see in the data. This 'employment volatility puzzle' has been around since the pre-history history of RBC (Lucas & Rapping, 1969), and many suggested resolutions have been proposed in the intervening years. One of the earliest, which is still held in high regard by many macroeconomists, is to make labour indivisible (Rogerson, 1988). Another route is sticky wages (Hall, 2005). Then in the short run, (provided the flexible-wage equilibrium lies below the current sticky level,) we effectively have a horizontal labour supply curve, so shocks to labour demand should move employment around a lot. Christiano, Eichenbaum and Trabandt (2015) argue that all of these resolutions (and many more besides) fail, and propose their own alternative. Christiano, Trabandt and Walentin (2010) also discuss this problem.[68] In the next lecture, we'll take a closer look at some of these attemps to resolve the employment volatility puzzle.

Another problem with the basic RBC model is that it fails to generate comovement. As we wrote it, the RBC model only has one sector, but it can equivalently be written as a two-sector model in which investment and consumption goods are produced separately using the same Cobb–Douglas technology. Investment-goods-producing firms employ labour services $n_{I,t}$ and capital services $k_{I,t}$, while consumption-goods firms employ $n_{c,t}$ and $k_{c,t}$. Market clearing obviously requires that $k_{I,t} + k_{c,t} = k_t$ and $n_{I,t} + n_{c,t} = n_t$.[69] In equilibrium, factor prices must be the same in the two sectors.

The first-order condition of the final-goods sector w.r.t. labour input, imposing final-goods-market clearing, is

$$(1 - \alpha)c_t/n_{c,t} = w_t.$$

The intratemporal first-order condition of the household is

$$-u_2(c_t, n_t)/u_1(c_t, n_t) = w_t,$$

which for $u(c, n) = \ln(c) + \gamma \ln(1 - n)$ becomes we get

$$(1 - \alpha)c_t/n_{c,t} = \gamma c_t/(1 - n_t).$$

Rearranging yields

$$\frac{1 - \alpha}{n_{c,t}} = \frac{\gamma}{1 - n_{c,t} - n_{I,t}}.$$

---

[68] There was also a brief digression on the equity premium puzzle, though no suggestion was given as to how this relates to the employment volatility puzzle. Larry took the opportunity to tell us that the equity premium puzzle is actually no longer a puzzle because Boldrin, Christiano and Fisher (2001) solved it.

[69] Actually, these equations add the assumption that it's costless to move labour and capital between the two sectors. This may not be a good idealisation!

This equation implies that if $n_{c,t}$ goes down for whatever reason, $n_{I,t}$ must go up. So we get *negative* comovement in employment! This result is driven by the fact that when z_t goes up, investment jumps up more than consumption does because of the consumption-smoothing motive. Resources therefore have to be reallocated into the investment-goods industry, increasing $n_{I,t}$ and decreasing $n_{c,t}$.

While the real business cycle framework does make it hard to get positive comovement between investment- and final-goods sectors, it's perfectly possible to have comovement between sectors that produce distinct final goods. The original Long and Plosser (1983) model is like this, in fact. (Though as Larry pointed out, there is no employment volatility in that model. The comovement is in sectoral output.) Moreover, DiCecio (2009) shows that sectoral comovement arises naturally when firms have market power and markups are countercyclical.[70]

Yet another problem with RBC as a theory of cycles is that it treats TFP as exogenous. For one thing, we might think that this is an unsatisfying explanation of why cycles happen, just like it's an unsatisfying explanation of economic growth (recall the Solow I model). More troubling is that measured TFP appears not to be exogenous with respect to the endogenous economic variables; instead, TFP appears to be determined endogenously along with GDP, interest rates and so on (Evans, 1992). This problem has been recognised by proponents of RBC; for example, Prescott (1998) argues that we need models in which TFP is endogenous. There are many ways of endogenising TFP; we saw one in the Comin and Gertler (2006) model, and two more on homework 9.[71]

A final common objection to the RBC theory of business cycles is that it appears to require technological regress in order to explain recessions. Larry does not think that this is an important problem, however. The key to the RBC story is that TFP fluctuates exogenously, giving rise to business cycles. Whether or not we choose to interpret TFP fluctuations as technology shocks (as many early RBC theorists did) is inessential.

---

[70]In the data, markups are indeed countercyclical. Moreover, countercyclical markups gel with the intuitive notion that competition is fiercer in times of high demand. (Rotemberg and Saloner (1986) is a well-known formalisation of this intuition.)

[71]There is also a large and active empirical literature on the determinants of TFP. For example, Song, Storesletten and Zilibotti (2011) argue that Chinese TFP is far lower than it could be because government policy directs financing toward inefficient state-owned enterprises rather than highly productive private firms.

18 Mon 23 Nov 2015

18.1 Outline of today

Today we look at ways of resolving the employment volatility puzzle. Recall from last time that while a basic RBC model is able to get the relative volatilities of GDP, consumption and investment about right, it generates far too low employment volatility. We'll briefly say something about sticky wages, which Larry thinks is an unconvincing resolution, and search-and-matching, which Larry thinks is the right answer. We'll then study in some detail two popular resolutions that Larry is unimpressed by: efficiency wages and the Rogerson (1988) aggregation argument.

18.2 Resolving the employment volatility puzzle

As we saw in the previous lecture, sticky wages offer one way out. When the market-clearing wage is below the level $\bar{w}$ that the wage is currently 'stuck' at, equilibrium occurs at the intersection of labour demand and the horizontal line at $\bar{w}$. As labour demand shifts around in response to shocks, employment then responds vigorously as desired. Hall (2005) advocates this story, but Larry is unconvinced. He did not explain why, but apparently the reason can be found in Christiano et al. (2015).

Another natural approach would be to try to find an explanation in the search-and-matching models that have become the standard way of understanding the labour market in macro. Shimer (2005) argues (rather famously) that search-and-matching models cannot generate employment volatility of the required magnitude, but Larry thinks that the field has moved on. The models at the current cutting edge are able to generate sizeable employment volatility, he says. Although this is the resolution that Larry favours, we will say no more about it. Instead, we will examine in detail two popular resolutions that Larry thinks are unconvincing.

18.2.1 Efficiency wages

A model with efficiency wages is one in which for some reason or other, equilibrium wages are above the standard competitive-equilibrium level. Solow (1979) imported them from development economics, where the thought was that if workers are paid the competitive market wage, they will be so poor that their productivity is low. Several other microfoundations for efficiency wages have since been proposed. A famous one is Shapiro and Stiglitz (1984), where the effort of workers is imperfectly observed by their employers. In

order to induce effort in equilibrium, it must be that workers cannot expect to be immediately re-hired at the same wage if they are fired after being caught shirking. Since workers are homogeneous, this requires some degree of equilibrium unemployment, which is possible only if the equilibrium wage is above the competitive one. Another microfoundation for efficiency wages is monopoly power on the part of unions. Yet another is that workers value being paid a 'fair wage', and reduce effort if the wage is 'unfairly low' (Akerlof, 1982).

Solow (1979) pointed out that efficiency wages can dramatically change the labour-market implications of macroeconomic models. Consider a simple reduced-form model in which a representative firm has production function $f(k, z, N)$, where z a measure of technology, N is labour in efficiency units. N is equal to $ne(w)$ where n are hours worked and the efficiency $e(w)$ per hour worked depends on the wage. The firm has some market power.

To make life easier, define $x(w, n) := wn$, the wage bill. Then the firm's profit is

$$f\left(k, z, x\frac{e(w)}{w}\right) - x - rk.$$

The first-order conditions w.r.t. x and w are

$$1 = f_3\left(k, z, x\frac{e(w)}{w}\right)\frac{e(w)}{w}$$

$$0 = f_3\left(k, z, x\frac{e(w)}{w}\right)x\frac{1}{w}\left(e'(w) - \frac{e(w)}{w}\right).$$

Simplifying and eliminating x,

$$w = f_3\left(k, z, ne(w)\right)e(w)$$

$$0 = f_3\left(k, z, ne(w)\right)\left(e'(w) - \frac{e(w)}{w}\right).$$

Dividing the one by the other and rearranging, we get

$$\frac{e'(w)w}{e(w)} = 1.$$

(If you like pictures, you can draw this as a tangency.) This condition shows that the equilibrium wage w is fully pinned down the function e, which is given exogenously (though it may be microfounded by e.g. a shirking story). We can then explain employment volatility in the same way as in the sticky-wages story.

The main problem that Larry sees with this story comes from Gomme (1999). This paper is specifically about the shirking variant of efficiency wages derived from Shapiro and Stiglitz (1984). In an RBC setting, the economy is booming when z is high. Moreover, when z is high, labour demand is high, so for any given wage the labour market is relatively tight. This increases the probability that a worker who is fired after being caught shirking can quickly find a new job, weakening incentives not to shirk. The equilibrium efficiency wage must therefore be higher in a boom. But then we get procyclical wages, which dampen labour demand during booms. This substantially lowers the volatility of employment over the business cycle, and the employment volatility puzzle rears its ugly head once more.

There are ways around the Gomme (1999) critique. Alexopoulos (2004, 2006, 2007) has shown how an RBC model with shirking can generate high employment volatility. But Larry is unconvinced, though he did not say why.

18.2.2 The Rogerson (1988) aggregation argument

Perhaps the most obvious solution is to argue that the elasticity of aggregate labour supply is simply very high. Then the labour demand curve is nearly flat, so employment is highly sensitive to wages, generating highly procyclical employment. To pursue this, suppose the representative household has preferences over consumption and labour represented by

$$u(c, n) := \ln(c) - (1 + \phi)^{-1} n^{1+\phi}.$$

The intratemporal first-order condition is then

$$w_t = -\frac{u_2(c_t, n_t)}{u_1(c_t, n_t)} = c_t n_t^{\phi},$$

so that $\ln(n_t) = \phi^{-1} [\ln(w_t) - \ln(c_t)]$. Hence

$$\ln(n_{t+1}/n_t) = \phi^{-1} [\ln(w_{t+1}/w_t) - \ln(c_{t+1}/c_t)].$$

The Frisch elasticity of substitution is then

$$\frac{\mathrm{d} \ln(n_{t+1}/n_t)}{\mathrm{d} \ln(w_{t+1}/w_t)} = \phi^{-1}.$$

It turns out that the Frisch elasticity is the relevant elasticity for generating employment volatility in a business cycle model. (Ordinary Marshallian and Hicks elasticities don't make much sense in a dynamic labour supply

problem.) Formally, the Frisch elasticity is the responsiveness of the path of labour supply to changes in the path of wages when the household is compensated so that her lifetime budget constraint is unchanged. Informally, we might think that a small, temporary change in wages has little effect on the lifetime budget constraint. In that case, we can think of the Frisch elasticity as the responsiveness of the labour supply path to a temporary wage increase.

Early in the RBC literature, ϕ^{-1} was calibrated at around 3.7. This value was thought to be on the high end of what was plausible, and yet is still not large enough to obtain sizeable employment volatility. This calibration came under fire from Heckman (1984), who pointed out that the consensus in the labour literature at the time was that ϕ^{-1} was about 0! Heckman further complained that no matter how you calibrate ϕ^{-1}, the RBC model is still unable to explain unemployment fluctuations because the greater part of volatility in aggregate hours worked in the data is on the extensive margin, whereas the representative-household model has only an intensive margin.

The first point is not as insurmountable as it might appear. I don't know how firm the consensus that $\phi^{-1} \simeq 0$ was in the labour supply literature in 1984, but in the modern literature, there are several prominent figures who believe that ϕ^{-1} is larger even than 3.7. One example is Imai and Keane (2004), who estimate of a structural life-cycle model with human capital accumulation and find that $\phi^{-1} \simeq 3.8$. Another pair of examples is Oettinger (1999) and Fehr and Goette (2007), who find large values of ϕ^{-1} in field experiments.

But Rogerson (1988) saved us from having to worry about all of this. He conceded to Heckman (1984) that employment fluctuations should occur on the extensive rather than the intensive margin. But as he pointed out, this nullifies Heckman's first criticism: it doesn't matter how low an individual's Frisch elasticity is if all the action is on the extensive margin. To prove the point, he provided a simple model in which individuals have heterogeneous tastes for work, and must each make a binary choice of whether or not to work. There is no intensive margin in these consumers' problems, so each person's Frisch elasticity is zero![72] Nevertheless, Rogerson showed that consumption and labour supply behaviour in this setting are equivalent to those of a representative household with utility function u as above. Moreover, the parameter ϕ^{-1} is now a function of the distribution of tastes for work in the population, so may plausibly be high even though every individual has a

---

[72]Loosely speaking. Formally, the Frisch elasticity is not defined because labour supply is not differentiable.

Frisch elasticity of zero.

Rogerson's idea was not well-understood until recently because his formulation involves some issues of interpretation. We will therefore present the variant due to Galí (2011a, 2011b), which is formally equivalent but easier to interpret. There is a single household consisting of a unit measure of individuals. The aversion to work of person $\ell \in [0,1]$ is ℓ^ϕ. Consumption and labour supply decisions are made by diktat by the head of the household. She cares equally for all household members, i.e. she acts on the utilitarian social welfare function

$$\int_0^1 \ln(c(\ell))\mathrm{d}\ell - \int_0^1 \ell^\phi h(\ell)\mathrm{d}\ell$$

where $c(\ell)$ is person ℓ's consumption and $h(\ell) = 1$ iff person ℓ works. Perfect consumption insurance is obviously optimal: $c(\ell) = c \ \forall \ell \in [0,1]$ where c is aggregate consumption.[73] Suppose that she chooses total labour supply to be n. Then it's clearly optimal to have $h(\ell) = \mathbf{1}(\ell \in [0,n])$ (where $\mathbf{1}$ is the indicator function).[74,75] Utility can therefore be written as a function of aggregate consumption and labour supply alone:

$$\int_0^1 \ln(c)\mathrm{d}\ell - \int_0^1 \ell^\phi \mathbf{1}(\ell \in [0,n])\mathrm{d}\ell = \int_0^1 \ln(c)\mathrm{d}\ell - \int_0^n \ell^\phi \mathrm{d}\ell$$
$$= \ln(c) - (1+\phi)^{-1}n^{1+\phi}.$$

So the behaviour of consumers in this economy is observationally equivalent in terms of the aggregate variables (c,n) to the representative-agent model used before. But now the parameter ϕ has an entirely new interpretation: it controls the cross-sectional distribution of tastes for work in the population. A priori, there is no obvious reason why ϕ^{-1} can't be very large!

Nevertheless, Larry is unimpressed. He didn't say much about why besides mentioning that labour economists have once more found a problem. Chetty, Guren, Manoli and Weber (2012) argue that just as macroeconomists calibrate technological parameters and the Frisch elasticity using estimates from the

---

[73] Perfect consumption insurance will be technically convenient. But Larry emphasised that it is not a key to the result.

[74] Actually, the household could send a strict subset of $[0,n]$ to work and still provide measure n total labour supply. But since social welfare is invariant to changes of $h(\cdot)$ on sets of measure zero, this need not trouble us. (Much.)

[75] Christiano was keen to point out that interesting things can happen here if the matriarch doesn't perfectly observe each person's distaste for work ℓ^ϕ. Then she faces a mechanism-design problem, so will distort labour-supply and consumption choices within the household to maintain incentive-compatibility.

microeconometric literature, so they ought to calibrate the 'extensive-margin elasticity' using estimates from the labour supply literature. The authors find that when the extensive-margin elasticity is chosen in this way rather than being treated as a free parameter, employment becomes puzzlingly non-volatile again. Again!

19 Wed 25 Nov 2015

19.1 Outline of today

Today, we'll look at models in which business cycles are driven by non-fundamental shocks, i.e. shocks extrinsic to preferences, technology and endowments. We will focus on sunspots, the oldest formalisation of non-fundamental fluctuations in DSGE models. The rough idea is that when there are multiple equilibria, equilibrium selection (or 'coordination') may be based on the realisation of some random variable that is otherwise unrelated to fundamentals.

19.2 Non-fundamental theories of business cycles

Non-fundamental business cycles occupy a strange place in macroeconomics. Almost everyone, says Larry, believes in their heart that business cycles are driven at least in part by non-fundamental factors. But since the 1980s, non-fundamental business cycles have largely been dismissed by macroeconomists. I think this may be because it's proven difficult to formalise a plausible way in which extrinsic fluctuations can arise in a DSGE framework.

A prima facie plausible story about extrinsic fluctuations can be spun as follows. A hotel wants to build a new swimming pool iff it expects lots of customers. Construction workers will choose to stay at a hotel iff they have a job. So there are good and bad equilibria: one with high employment and one with low employment. Non-fundamental business cycles occur as the economy jumps between good and bad equilibria for some reason (whatever reason).

This 'coordination failure' story is as old as business cycle theory itself. The paradox of thrift, popularised by Keynes (1936), is one kind of coordination failure story (it's pretty much exactly the hotel-construction worker vignette above, actually). Many subsequent attempts to formalise 'Keynesianism' involved coordination failure.

It turned out, however, that formalising extrinsic fluctuations in the context of the DSGE models that rose to prominence in the 1980s is not so easy. The basic reason is that these models have rational expectations, so that beliefs are endogenously determined. It's not straightforward to write down a model in which endogenous 'rational' beliefs fluctuate with non-fundamental randomness. The difficulty was demonstrated formally by Cass and Shell (1983). They showed that in an infinite-horizon economy with complete markets, extrinsic uncertainty cannot affect the equilibrium allocations. Although they also showed that the same is not true for overlapping-generations

economies, most business cycle theorists like their DSGE models to have one immortal household.

Cass and Shell (1983) gave rise to a literature studying when extrinsic uncertainty can and cannot matter for equilibrium allocations. This literature focused on one particular mechanism via which an extrinsic random variable can impinge on economic activity: the sunspots mechanism. A sunspot is an extrinsic random variable that acts as a coordination device or equilibrium selection mechanism.[76] (I am speaking loosely here!) Of course, this means that sunspots can only explain volatility when there are multiple equilibria to select between.

The source of multiple equilibria in sunspots models is usually strategic complementarities.[77] These complementarities can be between e.g. firms (as in the Christiano and Harrison (1999) model below), but it can also be between the private sector and government (as in homework 8).[78] It's easy enough to show by example that strategic complementarity can give rise to equilibrium multiplicity.

Larry expressed some discomfort with this formalisation of extrinsic business cycles. The problem is that the power of the rational-expectations hypothesis to discipline equilibrium beliefs is severely diminished when there is a large multiplicity of equilibria, simply because a large number of beliefs are consistent with some equilibrium. This starts to hollow out the empirical content of the theory: to say 'it was caused by a sunspot' is to say almost nothing at all. Another concern (mine, not Larry's) with the sunspots story is that the multiple equilibria needed to support sunspot fluctuations require

---

[76]The term 'sunspot' is unfortunate in more than one way. First, Cass and Shell (1983) actually use the term to mean any extrinsic random variable, no matter what the proposed channel through which it hopes to affect equilibrium allocations. Second, the term 'sunspot' was chosen in reference to an amusing paper by Jevons (1878) in which he tried to account for the business cycle by arguing that sunspots (literally, some kind of solar activity) directly affect agricultural productivity. (An RBC theory 100 years before Prescott, and published in *Nature*!) This is potentially confusing because Jevons thought precisely that (solar) sunspots are shocks to fundamentals, i.e. he thought that (solar) sunspots were *not* 'sunspots'! Sigh.

[77]This term is often used without being defined. I think the most natural definition (which comes from the supermodular games literature) is that an interaction in which player i takes action x_i and has payoff function f_i exhibits strategic complementarity iff $f_i(x_i, x_{-i})$ is single-crossing in (x_i, x_{-i}) and $f_i(\cdot, x_{-i})$ is quasi-supermodular. The best replies are then monotone in the actions of other players by the Monotonicity Theorem of Milgrom and Shannon (1994), which is the intuition complementarity aims to capture.

[78]Christiano also mentioned Shleifer (1986), who studies a different source of complementarity: monopolists with good ideas whose patents will expire want to bring their products to market during a boom, i.e. when everyone else is going to market, in order to maximise sales before the patent expires.

bizarre economic environments with very special functional-form assumptions and knife-edge parameter values.

There has been plenty of work ouside the sunspots literature that is also concerned with coordination-failure-type theories business cycles. Bryant (1983, 1987) and Cooper and John (1988) are Keynesian coordination failure theories of business cycles that use (respectively) a DSGE framework and a game-theoretic one. The news literature arguably also falls under the heading of non-fundamental theories; see the survey by Beaudry and Portier (2014). A lot of contemporary work on non-fundamental business cycles involves informational frictions; two nice examples are Maćowiak and Wiederholt (2015) and Angeletos and La'O (2009). A third example is Angeletos and La'O (2013), who show in a very general setting how extrinsic random variables can cause aggregate fluctuations without the need for any equilibrium multiplicity!

There are also literatures using formalisations other than sunspots to study macro phenomena other than business cycles, such as bank runs, currency and debt crises, and asset bubbles. Bryant (1980) and Diamond and Dybvig (1983) study simple game-theoretic models of bank runs; Gertler and Kiyotaki (2015) is an infinite-horizon variant. Cole and Kehoe (2000) is a famous theoretical paper on sovereign default (there are many others). The global-games crowd has been pumping out lots of papers on debt and currency crises, e.g. Angeletos and Werning (2006), Angeletos, Hellwig and Pavan (2006) and Angeletos, Hellwig and Pavan (2007). There's a huge literature on asset bubbles, but I know next to nothing about it.

19.3   A sunspots model (Christiano & Harrison, 1999)

Christiano and Harrison (1999) provide a simple deterministic model in which strategic complementarities among firms (in the form of a production externality) give rise to a large multiplicity of rational-expectations equilibria. There are stable, cyclical and chaotic sunspot-independent equilibria as well as stochastic equilibria driven by a sunspot. The authors provide a partial characterisation of the equilibrium set by showing that all solutions to a tractable first-order difference equation correspond to an equilibrium.

Today, we'll introduce the model and prove the partial characterisation result for the equilibrium set. Next time, we'll study the efficient allocations, show that most equilibrium allocations are inefficient, and see how automatic stabilisers (a special form of fiscal policy under commitment) can shrink the equilibrium set until it contains only the efficient equilibrium.

19.3.1 Environment and equilibrium

The model is a variant of the neoclassical growth model, with elastic labour supply and TFP determined endogenously. We will consider only deterministic equilibria to save on notation.[79] In the usual notation, at each date t, the household takes k_t as given and solves

$$\max_{\{c_{t+j}, k_{t+1+j}, n_{t+j}\}_{j=0}^{\infty}} \sum_{j=0}^{\infty} \beta^j u(c_{t+j}, n_{t+j})$$

$$\text{s.t. } \forall j \in \mathbf{Z}_+ \quad c_{t+j} + k_{t+1+j} - (1-\delta)k_{t+j} \leq r_{t+j}k_{t+j} + w_{t+j}n_{t+j} + \pi_{t+j}$$

$$c_{t+j} \in \mathbf{R}_+, \quad k_{t+1+j} \in \mathbf{R}_+, \quad n_{t+j} \in [0,1]$$

where $u : \mathbf{R}_+ \times [0,1] \to \mathbf{R}$ is defined by

$$u(c, n) := \ln(c) + \sigma \ln(1-n) \quad \forall (c,n) \in \mathbf{R}_+ \times [0,1].$$

As usual, we say that $\{c_{t+j}, k_{t+1+j}, n_{t+j}\}_{j=0}^{\infty}$ solves the household's problem from k_t at prices $\{r_{t+j}, w_{t+j}\}_{j=0}^{\infty}$ iff it attains the maximum in this program.

The economy is sufficiently well-behaved that the intratemporal first-order condition, the Euler equation and transversality condition are sufficient for an interior sequence to attain the maximum for a given sequence of prices. Since we will be concerned only with a sufficient conditions for interior equilibria, we do not worry about necessary conditions for optima. The three conditions are

$$u_1(c_t, n_t) = \beta u_1(c_{t+1}, n_{t+1})(r_{t+1} + 1 - \delta) \qquad \forall t \in \mathbf{R}_+$$

$$u_2(c_t, n_t) = -w_t u_1(c_t, n_t) \qquad \forall t \in \mathbf{R}_+$$

$$\lim_{t \to \infty} \beta^t u_1(c_t, n_t)(r_t + 1 - \delta)k_t = 0.$$

Atomistic firms produce output y_t using capital and labour services k_t and n_t according to

$$y_t = f(z_t, k_t, n_t) = z_t k_t^\alpha n_t^{1-\alpha},$$

where z_t is TFP. TFP is determined endogenously as $z_t = Y_t^\gamma$ where Y_t is aggregate output, taken as given by each atomistic firm. We assume that $\gamma > 0$, so that there is a positive externality. This will give rise to strategic

---

[79]The development in this section is almost entirely unchanged when uncertainty is added. The Euler equation contains an expectations operator, so the sufficient condition for equilibrium given below has an expectations sign in it. See Christiano and Harrison (1999).

complementarity, as firms wish to produce more when TFP is high, which occurs when other firms are producing a lot. To keep the model neat, we assume that $\gamma = 1 - \alpha$, and set $\alpha = 1/3$ as usual.[80]

In equilibrium, $y_t = Y_t$. So the equilibrium aggregate technological constraint is

$$y_t = k_t^{\alpha/(1-\gamma)} n_t^{(1-\alpha)/(1-\gamma)} = k_t n_t^{(1-\alpha)/\alpha} = k_t n_t^2.$$

We therefore have substantial increasing returns at the aggregate level. Our assumption that $\gamma = 1 - \alpha$ will have the implication that we can scale the model to eliminate k_t (as in the Ak model). The marginal products of capital and labour are

$$\mathrm{MPK}(k_t, n_t) = n_t^2$$
$$\mathrm{MPL}(k_t, n_t) = 2k_t n_t.$$

Since the wage is no longer strictly concave in k_t and n_t, the wage will be very high in good times. This suggests that there may be equilibria in which employment fluctuates a great deal, since the household's wish to smooth leisure may be overwhelmed by the highly procyclical wage.

We say that (y_t, k_t, n_t) solves the representative firm's problem at prices (r_t, w_t) iff it satisfies the first-order conditions $r_t = \alpha y_t / k_t$ and $w_t = (1 - \alpha) y_t / n_t$. Using the equilibrium aggregate technology, the equilibrium factor prices are

$$r_t = \alpha n_t^2 = \alpha \mathrm{MPK}(k_t, n_t) = \tfrac{1}{3}\mathrm{MPK}(k_t, n_t)$$
$$w_t = (1 - \alpha) k_t n_t = \tfrac{1-\alpha}{2}\mathrm{MPL}(k_t, n_t) = \tfrac{1}{3}\mathrm{MPL}(k_t, n_t).$$

Factors of production are not paid their marginal products because firms do not take into account the production externality. This will have important consequences for efficiency.

We define equilibrium in the usual way.

Definition 9. Quantities $\{y_t, c_t, k_t, n_t\}_{t=0}^\infty$ and prices $\{r_t, w_t\}_{t=0}^\infty$ form a sequence-of-markets equilibrium iff

(1) At each date t, $\{c_{t+j}, k_{t+1+j}, n_{t+j}\}_{j=0}^\infty$ solves the household's problem from k_t at prices $\{r_{t+j}, w_{t+j}\}_{j=0}^\infty$.

(2) At each date t, (y_t, k_t, n_t) solves the representative firm's problem at prices (r_t, w_t).

---

[80]This assumption is not innocuous. Most papers in the (large) empirical literature on estimating γ find $\gamma < 2/3$.

(3) At each date t, the goods market clears:

$$c_t + k_{t+1} - (1 - \delta)k_t = y_t.$$

19.3.2 A sufficient condition for deterministic equilibrium

Now let's hunt for that sufficient condition for a sequence to be a sequence-of-markets equilibrium. With functional forms plugged in, the household's first-order conditions are

$$c_t^{-1} = \beta c_{t+1}^{-1}(r_{t+1} + 1 - \delta)$$
$$\sigma(1 - n_t)^{-1} = w_t c_t^{-1}.$$

Substituting using the equilibrium factor prices,

$$c_t^{-1} = \beta c_{t+1}^{-1} \left(\alpha n_{t+1}^2 + 1 - \delta \right)$$
$$\sigma(1 - n_t)^{-1} = (1 - \alpha)k_t n_t c_t^{-1}.$$

With the factor prices plugged in, the transversality condition is

$$\lim_{t \to \infty} \beta^t c_t^{-1} \left(\alpha n_t^2 + 1 - \delta \right) k_t = 0.$$

Finally, the resource constraint is

$$c_t + k_{t+1} - (1 - \delta)k_t = k_t n_t^2.$$

We now rescale the model to get rid of k_t. Define $\widetilde{c}_t := c_t/k_t$ and $\lambda_t := k_{t+1}/k_t$. The household's first-order conditions are then

$$\lambda_t = \beta \frac{\widetilde{c}_t}{\widetilde{c}_{t+1}} \left(\alpha n_{t+1}^2 + 1 - \delta \right)$$
$$\widetilde{c}_t = \tfrac{1-\alpha}{\sigma} n_t(1 - n_t)$$

where I also rearranged a little. Plugging the latter into the former yields

$$\lambda_t = \beta \frac{n_t(1 - n_t)}{n_{t+1}(1 - n_{t+1})} \left(\alpha n_{t+1}^2 + 1 - \delta \right).$$

The resource constraint, rearranged a little, is

$$\lambda_t = n_t^2 + 1 - \delta - \widetilde{c}_t = n_t^2 + 1 - \delta - \tfrac{1-\alpha}{\sigma} n_t(1 - n_t)$$

where I used the second FOC. Combining these expressions to eliminate λ_t then gives

$$n_t^2 + 1 - \delta - \tfrac{1-\alpha}{\sigma} n_t(1 - n_t) = \beta \frac{n_t(1 - n_t)}{n_{t+1}(1 - n_{t+1})} \left(\alpha n_{t+1}^2 + 1 - \delta \right).$$

We can rearrange this as

$$n_{t+1}(1 - n_{t+1}) = \kappa(n_t) \left(\beta \alpha n_{t+1}^2 + 1 - \delta \right)$$

where $\kappa(n_t)$ is defined by

$$\kappa(n_t)^{-1} = \frac{n_t^2 + 1 - \delta}{n_t(1 - n_t)} - \frac{1 - \alpha}{\sigma}.$$

We can rewrite this endlessly. Larry rewrote it as

$$v(n_t, n_{t+1}) = 0,$$

where $v : [0,1]^2 \to \mathbf{R}$ is defined by

$$v(n_t, n_{t+1}) := n_{t+1}(1 - n_{t+1}) - \kappa(n_t) \left(\beta \alpha n_{t+1}^2 + 1 - \delta \right).$$

It might be more helpful to write it explicitly as a quadratic:

$$[\beta \alpha \kappa(n_t) + 1] \, n_{t+1}^2 - n_{t+1} + (1 - \delta)\kappa(n_t) = 0. \tag{20}$$

We've now combined all four first-order conditions plus the resource constraint into a single first-order difference equation in n_t. But we haven't yet said anything about the transversality condition. Let's first rewrite the object whose limit must be zero:

$$\beta^t u_1(c_t, n_t)(r_t + 1 - \delta) k_t = \beta^t \frac{k_t}{c_t} \left(\alpha n_t^2 + 1 - \delta \right) = \beta^t \frac{\alpha n_t^2 + 1 - \delta}{\frac{1-\alpha}{\sigma} n_t(1 - n_t)}.$$

It follows that the transversality condition is equivalent to

$$\lim_{t \to \infty} \beta^t \frac{\alpha n_t^2 + 1 - \delta}{n_t(1 - n_t)} = 0.$$

Having rewritten everything in terms of n_t, we need to keep track of how to back out the other variables. To this end, we use the functions

$$\lambda(n) := n^2 + 1 - \delta - \frac{1-\alpha}{\sigma} n(1-n)$$

$$k\left(\{n_s\}_{s=0}^t, k_0\right) := k_0 \prod_{s=0}^{t-1} \lambda(n_s)$$

$$c\left(\{n_s\}_{s=0}^t, k_0\right) := \frac{1-\alpha}{\sigma} n(1-n) k\left(\{n_s\}_{s=0}^t, k_0\right)$$

$$y\left(c, k, k'\right) := c + k' - (1-\delta)k$$

$$r(n) := \alpha n^2$$

$$w(k, n) := (1-\alpha)kn.$$

Lemma 1 (sufficient conditions for equilibrium). Suppose that $\{n_t\}_{t=0}^{\infty}$ satisfies $n_t \in [a, b]$ $\forall t$ for some $a, b \in (0, 1)$, and that (20) holds for each t. Define $\{c_t, k_t, y_t, r_t, w_t\}_{t=0}^{\infty}$ from k_0 and $\{n_t\}_{t=0}^{\infty}$ by, for each t,

$$c_t := c\left(\{n_s\}_{s=0}^t, k_0\right)$$

$$k_t := k\left(\{n_s\}_{s=0}^t, k_0\right)$$

$$y_t := y(c_t, k_t, k_{t+1})$$

$$r_t := r(n_t)$$

$$w_t := w(k_t, n_t).$$

Then quantities $\{y_t, c_t, k_t, n_t\}_{t=0}^{\infty}$ and prices $\{r_t, w_t\}_{t=0}^{\infty}$ form a sequence-of-markets equilibrium.

Proof. By construction, final-goods market clearing, both of the firm's FOCs and both of the household's FOCs hold. So parts 2 and 3 of the equilibrium definition (firm optimality and final-goods market clearing) are respected. To show that part 1 of the equilibrium definition (household optimality) is satisfied, use the fact that the first-order conditions plus transversality conditions are sufficient for the optimality of an interior sequence. It is therefore enough to show that the sequence is interior and that transversality is respected.

Interiority of $\{c_t, k_t\}_{t=0}^{\infty}$ is obvious by inspection of the functions c and k. To verify transversality, note that since $\{n_t\}_{t=0}^{\infty}$ is bounded away from 0 and 1, the sequence

$$\left\{ \frac{\alpha n_t^2 + 1 - \delta}{n_t(1 - n_t)} \right\}_{t=0}^{\infty}$$

148

is bounded. Hence

$$\lim_{t \to \infty} \beta^t \frac{\alpha n_t^2 + 1 - \delta}{n_t(1 - n_t)} = 0,$$

i.e. the transversality condition is respected. ∎

What equilibria does equation (20) admit? Since it's a quadratic in n_{t+1} for each n_t, there may be two admissible transitions from a given n_t, i.e. the mapping $n_t \mapsto n_{t+1}$ may be one-to-many. (This occurs iff both roots of the quadratic lie in $(0,1)$). Moreover, the transversality condition was shown to not restrict $\{n_t\}_{t=0}^{\infty}$ beyond requiring that it must be bounded away from 0 and 1, so there is a continuum of initial conditions. So provided that roots in $(0,1)$ don't generically fail to exist, we have a very large multiplicity of equilibria!

However, Larry didn't provide any reasons to think that we can expect at least one root to lie in $(0,1)$ for any given n_t and a wide range of parameters, which rather undercuts the drama of 'so much multiplicity'. The paper doesn't seem to contain any argument like this, either. What it does do, which is arguably good enough, is explicitly construct simple example equilibria with very different properties, including stable, cyclical and chaotic ones.

Larry also drew a picture showing the graph of the mapping $n_t \mapsto n_{t+1}$ in $[0,1]^2$. It contains a lower and an upper branch, corresponding to the two roots of the quadratic in n_{t+1} that arises for each n_t. At least in the picture, both branches are everywhere within $(0,1)$, so that both roots of the quadratic in n_{t+1} at every n_t has both roots in $(0,1)$. The upper branch does not cut the diagonal, but the lower branch (corresponding to the smaller root) cuts the diagonal exactly twice, so there are two steady states.[81] Larry claimed without proof that there are generically (in (α, β, δ)) two steady states, but curiously there seems to be no argument to that effect in the paper, where the authors only look at one or two particular calibrations. And when I looked at this question in homework 10, I found (lots of) parameter values at which there are *no* steady states.

19.3.3 Sunspot equilibria

Our previous analysis was for deterministic equilibria only, and was aided by a simple sufficient condition for a sequence to be an equilibrium. Stochastic equilibria were barely covered in the lectures, but there was a substantial

---

[81] The equation is too complicated for me to understand why the picture looks the way it does, so I didn't bother drawing it. But it can be found in Figure 1 of Christiano and Harrison (1999).

problem on homework 10 on this topic. I will summarise some of the findings of Christiano and Harrison (1999).

Christiano and Harrison (1999) establish a straightforward generalisation of this partial characterisation result for the stochastic case. For the same function $v : [0,1]^2 \to \mathbf{R}$ defined above, the more general sufficient condition for rational-expectations equilibrium is $\mathbf{E}_t(v(n_t, n_{t+1})) = 0$, or more explicitly

$$\mathbf{E}\left(v(n_t, n_{t+1})|\mathcal{I}_t\right) = 0$$

where $\mathcal{I}_t$ is the time-t information set.[82]

Christiano and Harrison (1999) construct two examples of sunspot equilibria in this economy. One example is a 'conventional sunspot equilibrium' that involves randomly jumping back and forth in the vicinity of the unstable nonstochastic steady state of $v(n_t, n_{t+1}) = 0$. The other is a 'regime-switching sunspot equilibrium' which involves stochastic jumps between the upper and lower branches of $v(n_t, \cdot) = 0$. We won't go into much detail below, so see homework 10 for some of the analysis and e.g. the survey by Benhabib and Farmer (1999) for both the big picture and the technicalities.

As emphasised by Benhabib and Farmer (1999), sunspot equilibria are often randomisations over multiple deterministic equilibria. The regime-switching sunspot equilibria in Christiano and Harrison (1999) are fairly straightforward to understand at an intuitive level with this idea in mind. We won't give a more formal account than that.

The conventional sunspot equilibria have support entirely on the lower branch of $v(n_t, \cdot) = 0$, in particular in the vicinity of the stable steady state. They are called 'conventional' because they occur in a fairly wide variety of environments studied in the sunspots literature. Unlike regime-switching equilibria, their existence has nothing to do with the fact that $v(n_t, \cdot)$ has more than once branch. The fact that the conventional sunspot equilibria in the Christiano and Harrison (1999) model have support in the vicinity of a stable steady state is a general phenomenon. The key is that in the vicinity of a stable steady state, there is a continuum of deterministic rational-expectations equilibria that converge to the steady state.

In the context of the Christiano and Harrison (1999) model, we can see this as follows. By definition of stability, there exists a neighbourhood of the stable steady state such that any path along the lower branch starting in this neighbourhood converges to the steady state. Hence for any point in this

---

[82]Informally, the information set 'contains' all realisations of n_t and the other endogenous variables dated t or earlier. Formally, $\mathcal{I}_t$ is the σ-algebra generated by those random variables.

neighbourhood, there exists a deterministic rational-expectations equilibrium starting at that point and converging to the steady state. Since the lower branch is continuous, the intersection of this neighbourhood with the lower branch has the power of the continuum.

Woodford (1986) showed that there exist 'conventional' sunspot equilibria, i.e. stationary sunspot equilibria with support in the vicinity of a steady state, iff there exists a continuum of deterministic rational-expectations equilibria in the vicinity of that steady state, all of which converge to that steady state.[83,84] I do not find this result especially intuitive, and Larry didn't motivate it. But there are probably nice informal explanations somewhere in the literature!

---

[83]When I say that a rational-expectations equilibrium (a sequence) lies in the vicinity of a steady state (a point), I mean that every point in the sequence lies in the vicinity of the steady state.

[84]The term 'indeterminacy' is often used here. It seems that different authors use it in slightly different ways, but one definition is that the equilibrium dynamics are indeterminate in the vicinity of a steady state iff there exists a continuum of deterministic rational-expectations equilibria in the vicinity of that steady state. This sense of 'indeterminacy' is the same one that shows up in e.g. the Blanchard and Kahn (1980) conditions, in case that's helpful.

20 Mon 30 Nov 2015

20.1 Outline of today

Today we begin studying financial frictions. After a brief overview, we'll introduce the Bernanke, Gertler and Gilchrist (1999) model of costly state verification. Today, we'll study partial equilibrium in the credit market; next time we'll aggregate the partial-equilibrium behaviour, incorporate it into a standard DSGE model, and study the implications for business cycles.

20.2 Financial frictions

'Frictions' in DSGE models are essentially features that give rise to outcomes other than those in an RBC model. In general, they are sources of inefficiency, such a monopoly power or sticky prices. Financial frictions are simply frictions that enter in the credit market.

In the models we have studied so far, we never looked at the credit market. But as with e.g. the production of capital, it is possible to add a financial sector to these models without changing the equilibrium allocations. We could, for example, have the household put its savings in a bank, and have the bank use these funds to extend credit to entrepreneurs who produce capital goods and rent them out to goods-producing firms. Continuing on with perfect competition, entrepreneurs and banks make zero profits, so entrepreneurs pay all their profits back to banks in the form of loan repayments, and banks pay all their profits back to households in the form of a competitively determined rate of return on bank deposits.

A financial friction is a tweak to the model such that the well-functioning efficient financial market just described no longer operates efficiently. Almost invariably, this is modelled as arising from an agency problem between creditors and debtors. In the story above, banks are both creditors (to entrepreneurs, via loans) and debtors (to households, via deposits). Correspondingly, we can build in agency problems either on the assets side of banks' balance sheets (banks have to worry about the behaviour of the entrepreneurs they lend to) or on the liabilities side (households have to worry about the behaviour of the banks they deposit their savings in). Most papers do one or the other, not both (for reasons of tractability). We will focus on agency problems in the credit market, i.e. situations in which banks' lending to entrepreneurs is distorted by an asymmetric information problem.

Larry says that there are four kinds of financial frictions, all of which originated in the banking literature. I'll describe them in the context of the credit market, but all four apply equally well to the market for bank deposits.

(1) Costly state verification (CSV), called 'hidden information' or 'adverse selection' in mechanism design/contract theory. The agency problem stems from the fact that banks cannot costlessly observe the realised profitability of risky projects undertaken by the entrepreneurs that it lends to. Contracts must be written to incentivise truthful reporting in the usual way (Myerson, 1981).

(2) Hidden effort, called 'hidden action' or 'moral hazard' in mechanism design/contract theory. The agency problem is that banks cannot observe the behaviour of entrepreneurs, e.g. the effort they put into developing their business or the riskiness of their business plans. Contracts must be written, to the extent that it is possible and profitable, to incentivise doing the right thing.

(3) Adverse selection. Larry said this is different from CSV, but declined to say why. To my knowledge, there is no distinction between adverse selection and hidden information in the mechanism design literature. (Though sometimes the term 'adverse selection' is used for settings with hidden information in which no optimal mechanism is in place.)

(4) Running away. This just means that the entrepreneur can abscond with (some of) the money she borrowed from the bank. It looks to me like a special case of hidden action.

CSV is the most popular in macro these days, closely followed by hidden effort. Most people think that adverse selection is 'silly', apparently. Running away is also fairly silly, but is reasonably popular nonetheless because of how tractable it is. Christiano and Ikeda (2013) survey the all four approaches. As already advertised, we will focus on CSV.

Collateral constraints   At a later juncture in the lecture, Larry mentioned models with collateral constraints, which are typically associated with Kiyotaki and Moore (1997). The endogenous phenomenon here is that equilbrium borrowing B has to satisfy $B \leq \gamma N$ for some $\gamma \in \mathbf{R}_+$ where N is the borrower's net worth. The interpretation is that the net worth N of the borrower serves as collateral for the loan, and that γN is the salvageable value of the collateral if it is seized by the creditor following a default.

Although you may be able to microfound a collateral constraint using one of the four stories above, it seems to me that the microfoundation given in Kiyotaki and Moore (1997) is of an entirely different nature. Their story hinges on incompleteness of contracts, in particular that no enforceable

contract can be written which gives the creditor a claim to the borrower's output.[85] However, it *is* possible to write enforceable contracts for transferring ownership of the creditor's assets, e.g. land or real estate. In equilibrium, then, no loan is ever made above the resale value γN of the lender's contractible assets N, or else the lender would prefer to abscond.

Larry mentioned collateral-constraints models by way of contrast with CSV models. The problem with collateral constraints is that they are occasionally binding, giving rise to a host of computational and analytical issues. (The main problem is that the value function becomes non-differentiable, and hence methods such as perturbation are ruled out.) These difficulties can be handled in a simple model, but become prohibitive if you try to embed collateral constraints into your favourite DSGE model. A great virtue of the CSV approach is that while it also gives rise to a collateral constraint of the form $B \leq \gamma N$, this constraint always holds with equality in equilibrium, making the model nice and smooth.[86]

20.3 Costly state verification

Townsend (1979) laid the theoretical groundwork for CSV by showing that under costly state verification, the lender-optimal deterministic mechanism can be implemented by a standard debt contract.[87] The banking literature picked up on this result early in the 1980s, and macroeconomists imported the idea from there in the 1990s. The first paper to stick CSV into a DSGE model was Carlstrom and Fuerst (1997), though Bernanke et al. (1999) are often given all the credit. Macro models with CSV became reasonably popular during the Latin American debit crises of the 1990s, mainly due to Chang and Velasco (2000). But financial frictions (not just CSV) were thought by most mainstream macroeconomists to be fairly unimportant until the financial crisis of 2007–08. (This may have been because the previous financial crisis, viz. the bursting of the tech bubble, had minimal macroeconomic

---

[85]Tangentially, there's a large literature on contract incompleteness and its consequences. For example, many people think that contractual incompleteness is why firms exist! See Coase (1937), Klein, Crawford and Alchian (1978) and Williamson (1975, 1979, 1985) for one account, Grossman and Hart (1986) and Hart and Moore (1990) for a second, and Holmström and Milgrom (1991) and Holmström and Milgrom (1994) for a third.

[86]Larry also argued that CSV models have the further advantage that γ is endogenously determined in a plausible way (see below), but surely the Kiyotaki and Moore (1997) microfoundation is just as good?

[87]The qualifier 'deterministic' is important. In most cases, the lender strictly prefers to audit randomly if she is allowed to!

consequences.)[88]

We will study the framework of Bernanke et al. (1999), initially looking at partial equilibrium in the credit market. Once we stick this credit market equilibrium into an RBC framework, we'll see that the financial friction manifests itself in the form of a 'wedge' in the intertemporal Euler equation. In particular, $\beta u'(c_t)/u'(c_{t+1})$ will not be set equal to the rate of return, but rather to the rate of return times a 'wedge'. This obviously has welfare implications, but more importantly it alters the equilibrium dynamics of the model.

This gives rise to a problem that is common to many financial-frictions models. The wedge in the Euler equation varies over the business cycles in such a way that consumption and investment are pushed in opposite directions. This is a problem because in the data, both are procyclical. After spelling out this problem, we'll present a solution.

20.3.1 The credit market

The credit market is populated by a unit measure of entrepreneurs (borrowers) and a unit measure of banks (lenders). Entrepreneurs differ (only) in their net worth $N > 0$, and the distribution of entrepreneurs across values of N is described by a (probability) measure on $\mathbf{R}_{++}$ with a strictly positive density. (We can loosely interpret this to mean that there is a 'large number' of entrepreneurs at each N.) An entrepreneur's net wealth is publicly observable, so there is in effect a separate credit market for each N. Each of these markets is competitive, with a large number of entrepreneurs and a large number of banks operating in each.

Each entrepreneur has access to a project with random gross return $(1 + R^k)\omega$. The average rate of return R^k is fixed and common to all entrepreneurs. The random component ω is independent across entrepreneurs ('is idiosyncratic'). We assume that its CDF F has support on (some subset of) $[0, \infty)$ and that its mean is unity (i.e. $\int_0^\infty \omega F(\mathrm{d}\omega) = 1$). The business cycle implications of the model turn out to be highly sensitive to the precise form of F, so although we use a log-normal CDF as a baseline, it's a good idea to see what happens if F is e.g. a Pareto CDF.

Banks observe R^k, but not realisations of ω. It is possible for the bank to learn the value of ω for a given entrepreneur, but only by paying a monitoring cost (hence 'costly state verification'). Townsend (1979) showed that under these circumstances, the bank-optimal mechanism can be implemented by a

---

[88]Larry also mentioned two other papers that use CSV models: Christiano, Motto and Rostagno (2003) and Christiano et al. (2014).

standard debt contract. Such a contract is characterised by a loan amount B and a gross rate of interest Z. The entrepreneur is asked to pay back ZB (principal plus interest) if she is able. If she is unable to repay in full, the bank pays the monitoring cost to learn ω, then appropriates the entire realised return $(1 + R^k)\omega A$, where $A = N + B$ are the pre-investment assets of the 'failed' entrepreneur. This contract is obviously incentive-compatible: if the entrepreneur is able to repay in full, she gets to keep a strictly positive slice of profit, whereas she has to give everything to the bank if she goes bankrupt. We will not show that it is the optimal mechanism.[89]

The form of the contract means that all of the firm's ex-post assets $(1 + R^k)\omega A$ serve as collateral for the loan: a defaulting entrepreneur does not get to keep anything. Given the form of the contract, we can interpret the monitoring cost as the cost of bankruptcy. (It doesn't matter how bankruptcy costs are split between the entrepreneur and bank since the bank acquires ownership of the entrepreneur's assets in case of bankruptcy.) Mainly for technical convenience, we specify the cost of bankruptcy to be $\mu(1 + R^k)\omega A$ for some $\mu \in (0, 1)$, i.e. it is proportional to the (ex-post) size of the entrepreneur's balance sheet. This can be motivated with a story along the lines that the costs of bankruptcy come mainly in the form of paying lawyers and accountants by the hour to sift through the failed entrepreneur's balance sheet.

For a given contract (B, Z), there is a threshold value $\bar{\omega}$ such that the entrepreneur is unable to repay her loan in full iff $\omega < \bar{\omega}$. In particular, this threshold solves
$$\left(1 + R^k\right) \bar{\omega} A = ZB,$$
which can be rewritten as
$$\bar{\omega} = \frac{Z}{1 + R^k} \frac{L - 1}{L}$$
where $L := A/N$ is the entrepreneur's leverage. Some comparative statics: the threshold $\bar{\omega}$ is increasing in leverage L and the gross interest rate Z, both for obvious reasons. When $\omega = \bar{\omega}$, entrepreneurs are indifferent between defaulting and repaying, so it's our call. We will assume that they default in this case.

---

[89]It is not clear to me whether or not the bank is allowed to commit to this behaviour. If not then we have an additional incentive-compatibility constraint that ensures that the bank finds it worthwhile ex-post to pay the monitoring cost whenever an entrepreneur goes bankrupt.

We assume that the entrepreneur behaves as if she were risk-neutral, maximising her expected return. We will motivate this with perfect consumption insurance (more on this later); we do not require entrepreneurs to actually have risk-netural preferences. It is without loss of generality to have the entrepreneur maximise expected return normalised by the constant $N(1 + R)$, so we'll do this for convenience. Then her objective is

$$\frac{1}{N(1+R)} \int_{\bar{\omega}}^{\infty} \left[\left(1 + R^k\right) \omega A - ZB\right] F(\mathrm{d}\omega).$$

In order for this objective to make much sense, it must be that the maximised value of the objective is greater than unity, for otherwise the entrepreneur could make a bigger return by depositing her net worth N at a bank at interest rate R. (We could forbid the entrepreneur from doing this, but that doesn't seem economically sensible.)

Substituting using $ZB = (1 + R^k)\bar{\omega}A$, we can rewrite the entrepreneur's objective as

$$\frac{1 + R^k}{1 + R} L \int_{\bar{\omega}}^{\infty} (\omega - \bar{\omega}) F(\mathrm{d}\omega).$$

This expression need not actually be strictly increasing in L since $\bar{\omega}$ is increasing in L. But as L grows large $\bar{\omega}$ gets close to $Z/(1 + R^k)$, and then the entrepreneur's expected return is strictly increasing in L. This means that at a given rate of interest Z, the entrepreneur's demand for leverage L (or demand for loans B) is undefined ('infinite' in sloppy language). As a result, we will not be able to characterise credit market equilibrium as the intersection of supply and demand for credit, and that is precisely why contracts are described by both B and Z rather than just one or the other.[90]

In principle, we can determine the credit market equilibrium in B–Z space. But due to the nonmonotonicity at low B just mentioned, an entrepreneur's iso-expected-return curves are not monotonic in this space, which is rather inconvenient. We'll therefore translate contracts into L–$\bar{\omega}$ space, which is less interpretable but technically nicer. In this space, indifference curves are increasing since entrepreneurs always like higher L and lower $\bar{\omega}$. We must remember that an equilibrium in this space must lie weakly southeast of the iso-expected-return curve for expected return 1, since as argued above, entrepreneurs would otherwise like to deposit their net wealth in a bank instead of pursuing their projects.

---

[90]In the early literature on information economics, it seems that many people thought its main contribution was precisely the idea that market equilibria are characterised by both price and quality or quantity, rather than by a Marshallian cross. See e.g. Stiglitz (1987). I think that this perspective is viewed as rather archaic now.

Let's do some more exhausting rewriting of the entrepreneur's expected return. Defining $G(x) := \int_0^x \omega F(\mathrm{d}\omega)$, the integral can be rearranged as

$$\int_{\bar{\omega}}^{\infty} (\omega - \bar{\omega}) F(\mathrm{d}\omega) = \int_0^{\infty} \omega F(\mathrm{d}\omega) - \int_0^{\bar{\omega}} \omega F(\mathrm{d}\omega) - \bar{\omega} \int_{\bar{\omega}}^{\infty} F(\mathrm{d}\omega)$$

$$= 1 - G(\bar{\omega}) - \bar{\omega} \left[1 - F(\bar{\omega}) \right]$$

$$= 1 - \Gamma(\bar{\omega})$$

where $\Gamma(x) := G(x) + x\left[1 - F(x) \right]$. So we can write the entrepreneur's normalised expected return as

$$[1 - \Gamma(\bar{\omega})] \frac{1 + R^k}{1 + R} L.$$

It's easy but extremely dull to establish the following properties of Γ and G:

$$\Gamma(x) \in [0, 1], \quad \Gamma'(x) = 1 - F(x), \quad \Gamma''(x) < 0, \quad \text{and} \quad G'(x) = xF'(x)$$

for every x, and

$$\lim_{x \to 0} \Gamma(x) = 0, \quad \lim_{x \to \infty} \Gamma(x) = 1, \quad \lim_{x \to 0} G(x) = 0, \quad \text{and} \quad \lim_{x \to \infty} G(x) = 1.$$

It seems that Larry also surreptitiously assumed that $F(x) < 1 \ \forall x \in \mathbf{R}_+$, so that the distribution has unbounded support. In that case, it follows that $\Gamma'(x) > 0$.

We assume that every bank has a continuum of creditors. Then 'by a law of large numbers', each individual bank's revenue is 'known for certain'.[91] In particular, a bank's revenue (including monitoring costs) 'is certain to be'

$$\int_{\bar{\omega}}^{\infty} (ZB) F(\mathrm{d}\omega) + (1 - \mu) \int_0^{\bar{\omega}} \left(1 + R^k \right) \omega A F(\mathrm{d}\omega),$$

$$= [1 - F(\bar{\omega})] ZB + (1 - \mu) \left(1 + R^k \right) A \int_0^{\bar{\omega}} \omega F(\mathrm{d}\omega).$$

Each bank pays its depositors the competitively determined rate of return R. We assume without loss of generality that each bank lends to a unit measure of entrepreneurs, so that the total credit it extends is B (the same as the amount of credit extended to each entrepreneur). Then the bank borrows

---

[91] Once again, it is not at all clear what theorem is being invoked here (see the complaint in footnote 52 on p. 116). Hence I say 'known for certain to be c' instead of e.g. 'almost surely equal to c'.

B from households, and has to repay $(1+R)B$. Perfect competition in the deposits market the implies the zero-profit condition

$$[1 - F(\overline{\omega})]ZB + (1 - \mu)\left(1 + R^k\right) A \int_0^{\overline{\omega}} \omega F(\mathrm{d}\omega) = (1 + R)B.$$

Dividing through by B, we find that in equilibrium, the gross rate of return that households earn on deposits is

$$1 + R = B^{-1}\left\{[1 - F(\overline{\omega})]ZB + (1 - \mu)\left(1 + R^k\right) A \int_0^{\overline{\omega}} \omega F(\mathrm{d}\omega)\right\}.$$

The right-hand side is the gross return per dollar invested in entrepreneurial projects, i.e. the average gross return. But we know that efficiency requires that $1 + R$ be equal to the *marginal* gross return to investment. This is the source of the inefficient 'wedge' in the Euler equation.

21 Wed 2 Dec 2015

21.1 Outline of today

We're continuing with the CSV model from last time. We'll finish up looking at the credit market, the move on to the general equilibrium.

21.2 Costly state verification again

21.2.1 The credit market again

Recap. Banks pay depositors a rate of return R that they take as given, and which is pinned down by the zero-profit condition. They extend loans of size B to a unit measure of entrepreneurs at interest rate Z. Entrepreneurs have random return $(1 + R^k)\omega A$, and if they default the bank recovers $(1 - \mu)(1 + R^k)\omega A$.

From last time, the zero-profit condition for banks is

$$[1 - F(\overline{\omega})]ZB + (1 - \mu)\left(1 + R^k\right) A \int_0^{\overline{\omega}} \omega F(\mathrm{d}\omega) = (1 + R)B.$$

This gives rise to an inefficiency because $1 + R$ is not equal to the average marginal return $1 + R^k$.

Continuing on, we can use $ZB = (1 + R^k)\overline{\omega}A$ to rewrite the zero-profit condition as

$$[1 - F(\overline{\omega})]\left(1 + R^k\right) \overline{\omega}A + (1 - \mu)\left(1 + R^k\right) A \int_0^{\overline{\omega}} \omega F(\mathrm{d}\omega) = (1 + R)B.$$

Note that
$$\frac{B}{A} = \frac{(A - N)/N}{A/N} = \frac{L - 1}{L}.$$

Hence dividing both sides of the zero-profit condition by $(1 + R^k)A$ gives

$$[1 - F(\overline{\omega})]\overline{\omega} + (1 - \mu) \int_0^{\overline{\omega}} \omega F(\mathrm{d}\omega) = \frac{1 + R}{1 + R^k} \frac{L - 1}{L},$$

which is a locus in L–$\overline{\omega}$ space. Recalling the definitions of G and Γ, some more algebra lets us write this as

$$\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right) = \frac{1 + R}{1 + R^k} \frac{L - 1}{L},$$

which rearranges to

$$L = \left(1 - \frac{1 + R^k}{1 + R} \left[\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right) \right] \right)^{-1}.$$

At least for the log-normal distribution, the right-hand side is nonmonotonic in $\overline{\omega}$, so leverage is pinned down uniquely by $\overline{\omega}$ but not vice versa.

We suppose that banks offer the entire menu of contracts on the zero-profit locus. Entrepreneurs with net worth N then choose their favourite. Recall that the entrepreneur's (normalised) expected return is

$$\left[1 - \Gamma\left(\overline{\omega}\right) \right] \frac{1 + R^k}{1 + R} L.$$

The entrepreneur maximises this expression in $(L, \overline{\omega})$ subject to the availability (i.e. zero-profit) constraint. Substituting out L using the constraint and taking logs for convenience, the objective is

$$\left[1 - \Gamma\left(\overline{\omega}\right) \right] \frac{1 + R^k}{1 + R} \left(1 - \frac{1 + R^k}{1 + R} \left[\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right) \right] \right)^{-1}.$$

The entrepreneur maximises this expression in $\overline{\omega}$. To make our lives easier, let's take logs:

$$\ln\left(1 - \Gamma\left(\overline{\omega}\right)\right) + \ln\left(\frac{1 + R^k}{1 + R} \right) - \ln\left(1 - \frac{1 + R^k}{1 + R} \left[\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right) \right] \right).$$

Larry didn't give us any compelling reason to think that this expression is quasi-concave in $\overline{\omega}$ in general. The most he did was show us a quasi-concave-looking picture for a particular parameterisation. Even so, let's use the first-order condition and hope for the best:

$$\frac{\Gamma'\left(\overline{\omega}\right)}{1 - \Gamma\left(\overline{\omega}\right)} = \frac{\frac{1 + R^k}{1 + R} \left[\Gamma'\left(\overline{\omega}\right) - \mu G'\left(\overline{\omega}\right) \right]}{1 - \frac{1 + R^k}{1 + R} \left[\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right) \right]}.$$

Recalling that $\Gamma'(x) = 1 - F(x)$ and $G'(x) = xF'(x)$,

$$\frac{1 - F\left(\overline{\omega}\right)}{1 - \Gamma\left(\overline{\omega}\right)} = \frac{\frac{1 + R^k}{1 + R} \left[1 - F\left(\overline{\omega}\right) - \mu\overline{\omega}F'\left(\overline{\omega}\right) \right]}{1 - \frac{1 + R^k}{1 + R} \left[\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right) \right]}.$$

Given a functional form for F and values for R^k, R and μ, this unsightly equation can be solved numerically for the equilibrium value of $\overline{\omega}$. We can

161

then back out leverage from the zero-profit condition as

$$L = \left(1 - \frac{1+R^k}{1+R}\left[\Gamma\left(\overline{\omega}\right) - \mu G\left(\overline{\omega}\right)\right]\right)^{-1}.$$

When we come to aggregating credit market behaviour, it will be very important that equilibrium leverage L does not vary with net worth N. With values of $\overline{\omega}$ and L, we can solve for anything else that might interest us. The 'risk spread' is

$$\frac{Z}{1+R} = \frac{1+R^k}{1+R}\overline{\omega}\frac{L}{L-1}.$$

Observe that the risk spread is the same across entrepreneurs with different net wealth. This is handy! The loan amount is $B = (L-1)N$, proportional to net worth. This looks like the collateral constraint mentioned earlier, but it always holds with equality. Apparently Levin, Natalucci and Zakrajšek (2004) find that loan contracts in the data are very similar to what this model predicts.

To do some comparative statics, let's use the log-normal distribution: $\ln(\omega) \sim \mathcal{N}\left(\mu, \sigma^2\right)$. A basic property of this distribution is that

$$\ln \mathbf{E}(\omega) = \mu + \frac{1}{2}\sigma^2.$$

Hence our restriction that $\mathbf{E}(\omega) = 1$ imposes $\mu = -\frac{1}{2}\sigma^2$. This leaves us with one free parameter, σ, which we call 'risk'. When σ is increased, the new equilibrium debt contract involves a higher Z (and hence risk spread) and lower leverage. But this finding is not robust; there are other distributions for which we get perverse comparative statics.[92]

There's currently a lot of interest in how cross-sectional risk σ interacts with the business cycle. Bloom (2009) famously showed that σ seems to be countercyclical, so that recessions are times of high 'risk'. The big question in this literature is whether business cycles are driven in part by exogenously varying risk (as in Bloom), or whether cross-sectional risk depends with aggregate conditions, or whether both vary with some other exogenous factor.

A possible problem with this model is that entrepreneurs have access to no form of funding other than bank loans. In particular, they cannot issue equity. Larry defended this with reference to Cole and Ohanian (1999), who

---

[92]Larry couldn't remember which distributions have this feature, but thought that the Pareto distribution might be one of them.

showed that during the Great Depression, many firms sold physical capital to keep their dividend payments flowing. If firms typically behave like this (which I'm pretty sure they don't) then equity financing is not so different from debt.

Another problem is that the only reason why the risk spread is above unity is that bankruptcy costs must be paid. In the data, the risk spread moves around a lot, probably due to e.g. risk aversion and liquidity. An adequate model would incorporate these factors as well.

Yet another concern is that banks cannot default in this model. We could reinterpret the model to have some entrepreneurs be financial institutions, but this is not ideal. If we do this, we have to change the calibration of the model. The calibration used in the slides imples leverage of around 2, which is about right for non-financial firms in the US. But financial firms have leverage of around 15–20!

For the policy-minded, another issue is that capital regulations (restricting how much leverage borrwers can have) cannot do any good in this model. In equilibrium, lending is inefficiently *low*!

21.2.2   General equilibrium

The rest of the economy is a standard neoclassical model. Period utility is $u(c) = \ln(c)$. The budget constraint is

$$c_t + B_{t+1} + K_{t+1} \le (1 + R_{t-1})B_t + (1 - \delta)K_t + r_t K_t + w_t \ell_t.$$

where B_t are bank deposits. This gives rise to two asset Euler equations as usual. Each household has 'certain' income 'by a law of large numbers', so there's perfect consumption insurance. This 'large family' assumption is essential for tractability purposes, but is not at all innocuous!

The basic interaction structure is as follows. Entrepreneurs can turn the services of 'raw' capital into 'effective' capital; raw capital is K_t, 'effective' capital is ωK_t. Firms that go bankrupt disappear forever (about 50% in the data); successful firms survive. The net worth of an entrepreneur is the income she earns from renting the services of her stock of effective capital to final-goods producing firms and from selling her effective capital on to capital-goods producing firms, less loan repayments. Final goods firms also sell to capital producers if they need to produce more capital than entrepreneurs can sell them.

Households deposit savings in banks, which lend to entrepreneurs. They also save by buying capital from the capital goods firm. Entrepreneurs buy raw capital from the capital firm, convert it into effective capital, rent it out

to final-goods firms and sell it back to capital producers. (It should be clear that we could move capital-goods production into the household without changing the equilibrium allocations.)

To analyse the general equilibrium, we need to derive aggregate entrepreneurial behaviour. Let $f_t(N)$ be the density at N of entrepreneurial net worth. One of these entrepreneurs gets a loan B_{t+1}^N and buys capital $K_{t+1}^N = N + B_{t+1}^N$. She gets out ωK_{t+1}^N, which she rents out at rate r_{t+1}, and then sells the remainder $(1-\delta)\omega K_{t+1}^N$ back to capital producers at price one. So the rate of return for the entrepreneur is $\omega(1 + R_{t+1}^k)$ in the previous notation, where now $1 + R_{t+1}^K = r_{t+1} + 1 - \delta$. Leverage K_{t+1}^N/N has to satisfy the zero-profit condition as before. The zero-profit condition depends on $\overline{\omega}_{t+1}$, which solves the unsightly equation we derived earlier. Total net worth is $N_{t+1} := \int_{\mathbf{R}_+} N f_t(N) \mathrm{d}N$. Total effective capital is

$$K_{t+1} = \int_{\mathbf{R}_+} \int_{\mathbf{R}_+} K_{t+1}^N \omega F(\mathrm{d}\omega) f_t(N) \mathrm{d}N.$$

Since leverage K_{t+1}^N/N is the same for all entrepreneurs,

$$K_{t+1} = \frac{K_{t+1}^N}{N} \int_{\mathbf{R}_+} \left(\int_{\mathbf{R}_+} \omega F(\mathrm{d}\omega) \right) N f_t(N) \mathrm{d}N = \frac{K_{t+1}^N}{N} \int_{\mathbf{R}_+} N f_t(N) \mathrm{d}N.$$

This depends only on the mean of the cross-sectional distribution of net worth. Without this feature, we would have to keep track of the entire distribution of net worth at each point in time, which would not be especially convenient for computational purposes.

Aggregate net wealth of entrepreneurs is $[1 - \Gamma(\overline{\omega}_{t+1})](1 + R_{t+1}^k)K_{t+1}$. A fraction $1 - \gamma$ of this is transferred to the household as a lump-sum (this behaviour looks a lot like the data). The household transfers resources W_{t+1}^e to the entrepreneur. There's a law of motion for entrepreneurial net wealth.

There's another couple of equations (goods market clearing, deposit market clearing...). We have seven variables $c_t, I_t, \overline{\omega}_t, R_t^k, K_t, N_t, R_t$ and seven equilibrium conditions. The first four equilibrium conditions look the same as in the neoclassical model, except that $R_t \neq R_t^k$.

The asset Euler equation for capital is

$$u'(c_t) = \beta u'(c_{t+1}) \left(1 + R_{t+1}^k \right) (1 - \tau_{t+1})$$

where

$$1 - \tau_{t+1} = \frac{1 + R_{t+1}}{1 + R_{t+1}^k}.$$

This looks like a time-varying tax on saving. We know from the neoclassical growth model that the Euler equation associated with the efficient allocations has $\tau_{t+1} = 0$ at all t. In that case, the household's saving decisions are made on the basis of the average marginal rate of return R_{t+1}^k. The 'wedge' $1 - \tau_{t+1}$ is therefore a source of inefficiency. Larry said that 'the steady-state welfare cost of the financial friction is 8.9% for a reasonable parameterisation', but I cannot for the life of me understand what a 8.9% drop in utility means. Utility is ordinal!

Moreover, when σ follows an exogenous stochastic process, the wedge follows an endogenously determined stochastic process. The equilibrium dynamics turn out to be quite different from those of an RBC models. In particular, for reasons Larry doesn't quite understand, the financial friction gives rise to a great deal of sluggishnes in the responses of the economy to shocks. This matches the data well, but it's a little hard to puzzle out why it happens!

There is a potential problem for the model here. When the wedge jumps, investment drops and consumption goes up because saving is effectively being taxed. But both consumption and investment are procyclical in the data! We could interpret this as a problem with our model of the financial sector, but we could also treat it as a rejection of the RBC model. And in fact, when we insert CSV into a new Keynesian model, this problem goes away. The reason is that when a risk shock hits so that τ_{t+1} shifts, the central bank's Taylor rule prevents R from moving very much. Consumption and investment decisions therefore do not move in opposite directions. (This is what Christiano et al. (2014) do.)

References

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics*, *97*(4), 543–569.

Alexopoulos, M. (2004). Unemployment and the business cycle. *Journal of Monetary Economics*, *51*(2), 277–298.

Alexopoulos, M. (2006). Shirking in a monetary business cycle model. *Canadian Journal of Economics*, *39*(3), 689–718.

Alexopoulos, M. (2007). A monetary business cycle model with unemployment. *Journal of Economic Dynamics and Control*, *31*(12), 3904–3940.

Anderson, R. M. (1991). Non-standard analysis with applications to economics. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of mathematical economics* (Chap. 39, Vol. 4, pp. 2145–2208). Amsterdam: North-Holland.

Angeletos, G.-M., Hellwig, C., & Pavan, A. (2006). Signaling in a global game: Coordination and policy traps. *Journal of Political Economy*, *114*(3), 452–484.

Angeletos, G.-M., Hellwig, C., & Pavan, A. (2007). Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. *Econometrica*, *75*(3), 711–756.

Angeletos, G.-M. & La'O, J. (2009). Noisy business cycles. In D. Acemoğlu, K. Rogoff, & M. Woodford (Eds.), *NBER macroeconomics annual* (Chap. 5, Vol. 24, pp. 319–378). Cambridge, MA: MIT Press.

Angeletos, G.-M. & La'O, J. (2013). Sentiments. *Econometrica*, *81*(2), 739–779.

Angeletos, G.-M. & Werning, I. (2006). Crises and prices: Information aggregation, multiplicity, and volatility. *American Economic Review*, *96*(5), 1720–1736.

Bak, P., Chen, K., Scheinkman, J., & Woodford, M. (1993). Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche*, *47*(1), 3–30.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, *82*(6), 1095–1117.

Basu, S. & Fernald, J. G. (1997). Returns to scale in U.S. production: Estimates and implications. *Journal of Political Economy*, *105*(2), 249–283.

Beaudry, P., Galizia, D., & Portier, F. (2015). *Reviving the limit cycle view of macroeconomic fluctuations*. NBER working paper 20101.

Beaudry, P. & Portier, F. (2014). News-driven business cycles: Insights and challenges. *Journal of Economic Literature*, *52*(4), 993–1074.

Benhabib, J. (2008). Chaotic dynamics in economics. In S. N. Durlauf & L. E. Blume (Eds.), *The new Palgrave dictionary of economics* (2nd). London: Palgrave Macmillan. Retrieved from http://www.dictionaryofeconomics. com/article?id=pde2008_C000104

Benhabib, J. & Day, R. H. (1982). A characterization of erratic dynamics in the overlapping generations model. *Journal of Economic Dynamics and Control*, *4*(November), 75–114.

Benhabib, J. & Farmer, R. E. A. (1999). Indeterminacy and sunspots in macroeconomics. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (Chap. 6, Vol. 1, pp. 387–448). Amsterdam: North-Holland.

Benveniste, L. M. & Scheinkman, J. A. (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica*, *47*(3), 727–732.

Bernanke, B., Gertler, M., & Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (Chap. 21, Vol. 1, pp. 1341–1393). Amsterdam: North-Holland.

Blanchard, O. J. & Kahn, C. M. (1980). The solution of linear difference models under rational expectations. *Econometrica*, *48*(5), 1305–1311.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, *77*(3), 623–685.

Boldrin, M., Christiano, L. J., & Fisher, J. D. M. (2001). Habit persistence, asset returns, and the business cycle. *American Economic Review*, *91*(1), 149–166.

Boldrin, M. & Montrucchio, L. (1986). On the indeterminacy of capital accumulation paths. *Journal of Economic Theory*, *40*(1), 26–39.

Bryant, J. (1980). A model of reserves, bank runs, and deposit insurance. *Journal of Banking & Finance*, *4*(4), 335–344.

Bryant, J. (1983). A simple rational expectations Keynes-type model. *Quarterly Journal of Economics*, *98*(3), 525–528.

Bryant, J. (1987). The paradox of thrift, liquidity preference and animal spirits. *Econometrica*, *55*(5), 1231–1235.

Carlstrom, C. T. & Fuerst, T. S. (1997). Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review*, *87*(5), 893–910.

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies*, *32*(3), 233–240.

Cass, D. & Shell, K. (1983). Do sunspots matter? *Journal of Political Economy, 91*(2), 193–227.

Chang, R. & Velasco, A. (2000). Liquidity crises in emerging markets: Theory and policy. In B. S. Bernanke & J. J. Rotemberg (Eds.), *NBER macroeconomics annual* (Chap. 1, Vol. 14, pp. 11–78). Cambridge, MA: MIT Press.

Cheremukhin, A., Golosov, M., Guriev, S., & Tsyvinski, A. (2013). *Was Stalin necessary for Russia's economic development?* NBER working paper 19425.

Chetty, R., Guren, A., Manoli, D., & Weber, A. (2012). Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities. In D. Acemoğlu, J. Parker, & M. Woodford (Eds.), *NBER macroeconomics annual* (Chap. 1, Vol. 27, pp. 1–56). Cambridge, MA: MIT Press.

Christiano, L. J. (1989). Understanding Japan's saving rate: The reconstruction hypothesis. *Quarterly Review of the Federal Reserve Bank of Minneapolis, Spring 1989*, 10–25.

Christiano, L. J. (2015). *Discussion of 'Networks and the macroeconomy: An empirical exploration' by Daron Acemoglu, Ufuk Akcigit and William Kerr*. NBER working paper 19425.

Christiano, L. J., Eichenbaum, M. S., & Trabandt, M. (2015). *Unemployment and business cycles*. Working paper, 15 Apr 2015.

Christiano, L. J. & Harrison, S. G. (1999). Chaos, sunspots and automatic stabilizers. *Journal of Monetary Economics, 44*(1), 3–31.

Christiano, L. J. & Ikeda, D. (2013). Government policy, credit markets, and economic activity. In M. D. Bordo & W. Roberds (Eds.), *The origins, history, and future of the Federal Reserve: A return to Jekyll Island* (Chap. 5, pp. 226–331). Cambridge: Cambridge University Press.

Christiano, L. J., Motto, R., & Rostagno, M. (2003). The Great Depression and the Friedman-Schwarz hypothesis. *Journal of Money, Credit and Banking, 35*(6), 1119–1197.

Christiano, L. J., Motto, R., & Rostagno, M. (2014). Risk shocks. *American Economic Review, 104*(1), 27–65.

Christiano, L. J., Trabandt, M., & Walentin, K. (2010). DSGE models for monetary policy analysis. In B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics* (Chap. 7, Vol. 3, pp. 285–367). Amsterdam: North-Holland.

Coase, R. H. (1937). The nature of the firm. *Economica, 4*(16), 386–405.

Cole, H. L. & Kehoe, T. J. (2000). Self-fulfilling debt crises. *Review of Economic Studies, 67*(1), 91–116.

Cole, H. L. & Ohanian, L. E. (1999). The Great Depression in the United States from a neoclassical perspective. *Quarterly Review of the Federal Reserve Bank of Minneapolis, Winter 1999, 2*–24.

Comin, D. & Gertler, M. (2006). Medium-term business cycles. *American Economic Review, 96*(3), 523–551.

Cooper, R. & John, A. (1988). Coordinating coordination failures in Keynesian models. *Quarterly Journal of Economics, 103*(3), 441–463.

Debreu, G. (1959). *Theory of value*. New Haven, CT: Yale University Press.

Diamond, D. W. & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy, 91*(3), 401–419.

DiCecio, R. (2009). Sticky wages and sectoral labor comovement. *Journal of Economic Dynamics and Control, 33*(3), 538–553.

Dixit, A. K. & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review, 67*(3), 297–308.

Duffie, D. & Sun, Y. (2012). The exact law of large numbers for independent random matching. *Journal of Economic Theory, 147*(1), 1105–1139.

Evans, C. L. (1992). Productivity shocks and real business cycles. *Journal of Monetary Economics, 29*(2), 191–208.

Fehr, E. & Goette, L. (2007). Do workers work more if wages are high? Evidence from a randomized field experiment. *American Economic Review, 97*(1), 298–317.

Frisch, R. (1933). Propagation and impulse problems in dynamic economics. In *Economic essays in honor of Gustav Cassel* (pp. 171–205). London: George Allen & Unwin.

Galí, J. (2011a). The return of the wage Phillips curve. *Journal of the European Economic Association, 9*(3), 436–461.

Galí, J. (2011b). *Unemployment fluctuations and stabilization policies: A new Keynesian perspective*. Cambridge, MA: MIT Press.

Geanakoplos, J., Axtell, R., Farmer, J. D., Howitt, P., Conlee, B., Goldstein, J., . . . Yang, C.-Y. (2012). Getting at systemic risk via an agent-based model of the housing market. *American Economic Review, 102*(3), 53–58.

Gertler, M. & Kiyotaki, N. (2015). Banking, liquidity, and bank runs in an infinite horizon economy. *American Economic Review, 105*(7), 2011–2043.

Gomme, P. (1999). Shirking, unemployment and aggregate fluctuations. *International Economic Review, 40*(1), 3–21.

Goodfriend, M. & King, R. (1997). The new neoclassical synthesis and the role of monetary policy. In B. S. Bernanke & J. Rotemberg (Eds.), *NBER*

macroeconomics annual (Chap. 5, Vol. 12, pp. 231–296). Cambridge, MA: MIT Press.

Gordon, R. J. (1990). *The measurement of durable goods prices*. Chicago, IL: University of Chicago Press.

Grandmont, J.-M. (1985). On endogenous competitive business cycles. *Econometrica*, *53*(5), 995–1045.

Green, E. J. (1994). *Individual-level randomness in a nonatomic population*. Working paper, Feb 1994.

Greenwood, J., Hercowitz, Z., & Huffman, G. W. (1988). Investment, capacity utilization, and the real business cycle. *American Economic Review*, *78*(3), 402–417.

Greenwood, J., Hercowitz, Z., & Krusell, P. (1997). Long-run implications of investment-specific technological change. *American Economic Review*, *87*(3), 342–362.

Grossman, S. J. & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, *94*(4), 691–719.

Hall, R. E. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, *95*(1), 50–65.

Hart, O. D. & Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy*, *98*(6), 1119–1158.

He, W., Sun, X., & Sun, Y. (2016). Modeling infinitely many agents. *Theoretical Economics*, *forthcoming*.

Heckman, J. (1984). Comments on the Ashenfelter and Kydland papers. *Carnegie-Rochester Conference Series on Public Policy*, *21*, 209–224.

Holmström, B. & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, *7*(special issue), 24–52.

Holmström, B. & Milgrom, P. (1994). The firm as an incentive system. *American Economic Review*, *84*(4), 972–991.

Hornstein, A. & Krusell, P. (1996). Can technology improvements cause productivity slowdowns? In B. S. Bernanke & J. Rotemberg (Eds.), *NBER macroeconomics annual* (Chap. 4, Vol. 11, pp. 209–276). Cambridge, MA: MIT Press.

Imai, S. & Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review*, *45*(2), 601–641.

Jevons, W. S. (1878). Commercial crises and sun-spots. *Nature*, *19*(472), 33–37.

Jones, L. E. & Manuelli, R. E. (1997). The sources of growth. *Journal of Economic Dynamics and Control*, *21*(1), 75–114.

Jovanović, B. & Rousseau, P. L. (2001). *Stock markets in the new economy*. Working paper, 14 Aug 2001.

Judd, K. L. (1985). The law of large numbers with a continuum of IID random variables. *Journal of Economic Theory*, *35*(1), 19–25.

Kalai, E. & Lehrer, E. (1993). Rational learning leads to Nash equilibrium. *Econometrica*, *61*(5), 1019–1045.

Kaldor, N. (1957). A model of economic growth. *The Economic Journal*, *67*(268), 591–624.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. London: Palgrave Macmillan.

King, R. G., Plosser, C. I., & Rebelo, S. T. (1988). Production, growth and business cycles I: The basic neoclassical model. *Journal of Monetary Economics*, *21*(2–3), 195–232.

Kiyotaki, N. & Moore, J. (1997). Credit cycles. *Journal of Political Economy*, *105*(2), 211–248.

Klein, B., Crawford, R. G., & Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law & Economics*, *21*(2), 297–326.

Koopmans, T. C. (1965). On the concept of optimal economic growth. In *The econometric approach to development planning* (pp. 225–300). Amsterdam: North-Holland.

Krusell, P. & Ríos-Rull, J.-V. (1996). Vested interests in a postitive theory of stagnation and growth. *Review of Economic Studies*, *63*(2), 301–329.

Kydland, F. E. & Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, *50*(6), 1345–1370.

La'O, J. (2015). *A traffic jam theory of recessions*. Working paper, 15 Sep 2015.

Levin, A. T., Natalucci, F. M., & Zakrajšek, E. (2004). *The magnitude and cyclical behavior of financial market frictions*. Working paper.

Ljungqvist, L. & Sargent, T. J. (2012). *Recursive macroeconomic theory* (3rd). Cambridge, MA: MIT Press.

Long, J. B. & Plosser, C. I. (1983). Real business cycles. *Journal of Political Economy*, *91*(1), 39–69.

Lucas, R. E. (1978). Asset pricing in an exchange economy. *Econometrica*, *46*(6), 1429–1445.

Lucas, R. E. & Rapping, L. A. (1969). Real wages, employment, and inflation. *Journal of Political Economy*, *77*(5), 721–754.

Maćowiak, B. & Wiederholt, M. (2015). Business cycle dynamics under rational inattention. *Review of Economic Studies*, *82*(4), 1502–1532.

Matsuyama, K. (1999). Growing through cycles. *Econometrica*, *67*(2), 335–347.

Milgrom, P. & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, *62*(1), 157–180.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, *6*(1), 58–73.

Oettinger, G. S. (1999). An empirical analysis of the daily labor supply of stadium vendors. *Journal of Political Economy*, *107*(2), 360–392.

Parente, S. L. & Prescott, E. C. (2000). *Barriers to riches*. Cambridge, MA: MIT Press.

Pigou, A. C. (1927). *Industrial fluctuations*. London: Macmillan.

Prescott, E. C. (1986). Theory ahead of business cycle measurement. *Quarterly Review of the Federal Reserve Bank of Minneapolis*, *Fall 1986*, 9–22.

Prescott, E. C. (1998). Needed: A theory of total factor productivity. *International Economic Review*, *39*(3), 525–551.

Prescott, E. C. & Mehra, R. (1980). Recursive competitive equilibrium: The case of homogeneous households. *Econometrica*, *48*(6), 1365–1379.

Ramsey, F. P. (1928). A mathematical theory of saving. *The Economic Journal*, *38*(152), 543–559.

Rogerson, R. (1988). Indivisible labor, lotteries and equilibrium. *Journal of Monetary Economics*, *21*(1), 3–16.

Romer, P. M. (1987). Growth based on increasing returns due to specialization. *American Economic Review*, *77*(2), 56–62.

Rotemberg, J. J. & Saloner, G. (1986). A supergame-theoretic model of price wars during booms. *American Economic Review*, *77*(3), 390–407.

Santos, M. S. (1991). Smoothness of the policy function in discrete time economic models. *Econometrica*, *59*(5), 1365–1382.

Shapiro, C. & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, *74*(3), 433–444.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, *95*(1), 25–49.

Shleifer, A. (1986). Implementation cycles. *Journal of Political Economy*, *94*(6), 1163–1190.

Slutsky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica*, *5*(2), 105–146.

Smets, F. & Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, *97*(3), 586–606.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics, 70*(1), 65–94.

Solow, R. M. (1960). Investment and technical progress. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical models in the social sciences, 1959: Proceedings of the first Stanford symposium* (Chap. 7, pp. 89–104). Stanford, CA: Stanford University Press.

Solow, R. M. (1979). Another possible source of wage stickiness. *Journal of Macroeconomics, 1*(1), 79–82.

Song, Z., Storesletten, K., & Zilibotti, F. (2011). Growing like China. *American Economic Review, 101*(1), 196–233.

Stiglitz, J. E. (1987). The causes and consequences of the dependence of quality on price. *Journal of Economic Literature, 25*(1), 1–48.

Stokey, N. L., Lucas, R. E., & Prescott, E. C. (1989). *Recursive methods in economic dynamics.* Cambridge, MA: Harvard University Press.

Sun, Y. (2006). The exact law of large numbers via Fubini extension and characterization of insurable risks. *Journal of Economic Theory, 126*(1), 31–69.

Sun, Y. & Zhang, Y. (2009). Individual risk and Lebesgue extension without aggregate uncertainty. *Journal of Economic Theory, 144*(1), 432–443.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory, 21*(2), 265–293.

Uhlig, H. (1996). A law of large numbers for large economies. *Economic Theory, 8*, 41–50.

Williamson, O. E. (1975). *Markets and hierarchies: Analysis and antitrust implications.* New York, NY: Free Press.

Williamson, O. E. (1979). Transaction-cost economics: The governance of contractual relations. *Journal of Law & Economics, 22*(2), 233–261.

Williamson, O. E. (1985). *The economic institutions of capitalism.* New York, NY: Free Press.

Woodford, M. (1986). *Stationary sunspot equilibria: The case of small fluctuations around a deterministic steady state.* Working paper, Sep 1986.

Young, A. (1992). A tale of two cities: Factor accumulation and technological change in Hong Kong and Singapore. In O. J. Blanchard & S. Fischer (Eds.), *NBER macroeconomics annual* (Chap. 1, Vol. 7, pp. 13–64). Cambridge, MA: MIT Press.