

## 3 The Neoclassical Growth Model in Discrete Time

### 3.1 Setup of the Model

The neoclassical growth model is arguably the single most important workhorse in modern macroeconomics. It is widely used in growth theory, business cycle theory and quantitative applications in public finance.

Time is discrete and indexed by  $t = 0, 1, 2, \dots$ . In each period there are three goods that are traded, labor services  $n_t$ , capital services  $k_t$  and a final output good  $y_t$  that can be either consumed,  $c_t$  or invested,  $i_t$ . As usual for a complete description of the economy we have to specify technology, preferences, endowments and the information structure. Later, when looking at an equilibrium of this economy we have to specify the equilibrium concept that we intend to use.

1. Technology: The final output good is produced using inputs labor and capital services, according to the aggregate production function  $F$

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

Note that I do not allow free disposal. If I want to allow free disposal, I will specify this explicitly by defining an separate free disposal technology. Output can be consumed or invested

$$y_t = i_t + c_t$$

Investment augments the capital stock which depreciates at a constant rate  $\delta$  over time

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

We can rewrite this equation as

$$i_t = k_{t+1} - k_t + \delta k_t$$

i.e. gross investment  $i_t$  equals net investment  $k_{t+1} - k_t$  plus depreciation  $\delta k_t$ . We will require that  $k_{t+1} \geq 0$ , but not that  $i_t \geq 0$ . This assumes that the existing capital stock can be disinvested to be eaten, capital is putty-putty. Note that I have been a bit sloppy: strictly speaking the capital stock and capital services generated from this stock are different things. We will assume (once we define the ownership structure of this economy in order to define an equilibrium) that households own the capital stock and make the investment decision. They will rent out capital to the firms. We denote both the capital *stock* and the *flow* of capital services by  $k_t$ . Implicitly this assumes that there is some technology that transforms one unit of the capital stock at period  $t$  into one unit of capital services at period  $t$ . We will ignore this subtlety for the moment.

2. Preferences: There is a large number of identical, infinitely lived households. Since all households are identical and we will restrict ourselves to type-identical allocations<sup>11</sup> we can, without loss of generality assume that there is a single representative household. Preferences of each household are assumed to be representable by a time-separable utility function (Debreu's theorem discusses under which conditions preferences admit a continuous utility function representation)

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

3. Endowments: Each household has two types of endowments. At period 0 each household is born with endowments  $k_0$  of initial capital. Furthermore each household is endowed with one unit of productive time in each period, to be devoted either to leisure or to work.
4. Information: There is no uncertainty in this economy and we assume that households and firms have perfect foresight. This is a special case of rational expectations. Under our assumption the output of our model (either in the form of an equilibrium or a Pareto efficient allocation) is a degenerate probability distribution over sequences: it will have probability one on a particular sequence.
5. Equilibrium: We postpone the discussion of the equilibrium concept to a later point as we will first be concerned with an optimal growth problem, i.e. the solution of the model for Pareto optimal allocations.

### 3.2 Optimal Growth: Pareto Optimal Allocations

Consider the problem of a social planner that wants to maximize the utility of the representative agent, subject to the technological constraints of the economy. Note that, as long as we restrict our attention to type-identical allocations, an allocation that maximizes the utility of the representative agent, subject to the technology constraint is a Pareto efficient allocation and every Pareto efficient allocation solves the social planner problem below. Just as a reference we have the following definitions

**Definition 9** *An allocation  $\{c_t, k_t, n_t\}_{t=0}^{\infty}$  is feasible if for all  $t \geq 0$*

$$\begin{aligned} F(k_t, n_t) &= c_t + k_{t+1} - (1 - \delta)k_t \\ c_t &\geq 0, k_t \geq 0, 0 \leq n_t \leq 1 \\ k_0 &\leq \bar{k}_0 \end{aligned}$$

---

<sup>11</sup>Identical households receive the same allocation by assumption. In the next quarter I may come back to the issue under which conditions this is an innocuous assumption,

**Definition 10** An allocation  $\{c_t, k_t, n_t\}_{t=0}^{\infty}$  is Pareto efficient if it is feasible and there is no other feasible allocation  $\{\hat{c}_t, \hat{k}_t, \hat{n}_t\}_{t=0}^{\infty}$  such that

$$\sum_{t=0}^{\infty} \beta^t U(\hat{c}_t) > \sum_{t=0}^{\infty} \beta^t U(c_t)$$

### 3.2.1 Social Planner Problem in Sequential Formulation

The problem of the planner is

$$\begin{aligned} w(\bar{k}_0) &= \max_{\{c_t, k_t, n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t) \\ \text{s.t.} \quad F(k_t, n_t) &= c_t + k_{t+1} - (1 - \delta)k_t \\ c_t &\geq 0, k_t \geq 0, 0 \leq n_t \leq 1 \\ k_0 &\leq \bar{k}_0 \end{aligned}$$

The function  $w(\bar{k}_0)$  has the following interpretation: it gives the *total lifetime* utility of the representative household if the social planner chooses  $\{c_t, k_t, n_t\}_{t=0}^{\infty}$  *optimally* and the initial capital stock in the economy is  $\bar{k}_0$ . Under the assumptions made below the function  $w$  is strictly increasing, since a higher initial capital stock yields higher production in the initial period and hence enables more consumption or capital accumulation (or both) in the initial period.

We now make the following assumptions on preferences and technology.

**Assumption 1:**  $U$  is continuously differentiable, strictly increasing, strictly concave and bounded. It satisfies the Inada conditions  $\lim_{c \searrow 0} U'(c) = \infty$  and  $\lim_{c \rightarrow \infty} U'(c) = 0$ . The discount factor  $\beta$  satisfies  $\beta \in (0, 1)$

**Assumption 2:**  $F$  is continuously differentiable and homogenous of degree 1, strictly increasing in both arguments and strictly concave. Furthermore  $F(0, n) = F(k, 0) = 0$  for all  $k, n > 0$ . Also  $F$  satisfies the Inada conditions  $\lim_{k \searrow 0} F_k(k, 1) = \infty$  and  $\lim_{k \rightarrow \infty} F_k(k, 1) = 0$ . Also  $\delta \in [0, 1]$

From these assumptions two immediate consequences for optimal allocations are that  $n_t = 1$  for all  $t$  since households do not value leisure in their utility function. Also, since the production function is strictly increasing in capital,  $k_0 = \bar{k}_0$ . To simplify notation we define  $f(k) = F(k, 1) + (1 - \delta)k$ , for all  $k$ . The function  $f$  gives the total amount of the final good available for consumption or investment (again remember that the capital stock can be eaten). From assumption 2 the following properties of  $f$  follow more or less directly:  $f$  is continuously differentiable, strictly increasing and strictly concave,  $f(0) = 0$ ,  $f'(k) > 0$  for all  $k$ ,  $\lim_{k \searrow 0} f'(k) = \infty$  and  $\lim_{k \rightarrow \infty} f'(k) = 1 - \delta$ .

Using the implications of the assumptions, and substituting for  $c_t = f(k_t) - k_{t+1}$  we can rewrite the social planner's problem as

$$\begin{aligned} w(\bar{k}_0) &= \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1}) & (25) \\ 0 &\leq k_{t+1} \leq f(k_t) \\ k_0 &= \bar{k}_0 > 0 \text{ given} \end{aligned}$$

The only choice that the planner really faces is the choice between letting the consumer consume today versus investing in the capital stock so that the consumer can consume more tomorrow. Let an (the optimal sequence of capital stocks be denoted by  $\{k_{t+1}^*\}_{t=0}^\infty$ . The two questions that we face when looking at this problem are

1. Why do we want to solve such a hypothetical problem of an even more hypothetical social planner. The answer to this questions is that, by solving this problem, we will have solved for competitive equilibrium allocations of our model (of course we first have to define what a competitive equilibrium is). The theoretical justification underlying this result are the two welfare theorems, which hold in this model and in many others, too. We will give a loose justification of the theorems a bit later, and postpone a rigorous treatment of the two welfare theorems in infinite dimensional spaces until the next quarter.
2. How do we solve this problem?<sup>12</sup> The answer is: dynamic programming. The problem above is an infinite-dimensional optimization problem, i.e. we have to find an optimal infinite sequence  $(k_1, k_2, \dots)$  solving the problem above. The idea of dynamic programming is to find a simpler maximization problem by exploiting the stationarity of the economic environment and then to demonstrate that the solution to the simpler maximization problem solves the original maximization problem.

To make the second point more concrete, note that we can rewrite the problem above as

$$\begin{aligned}
 w(k_0) &= \max_{\substack{0 \leq k_{t+1} \leq f(k_t) \\ k_0 \text{ given}}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1}) \\
 &= \max_{\substack{0 \leq k_1 \leq f(k_0) \\ k_0 \text{ given}}} \left\{ U(f(k_0) - k_1) + \beta \left[ \max_{\substack{0 \leq k_{t+1} \leq f(k_t) \\ k_1 \text{ given}}} \sum_{t=1}^{\infty} \beta^{t-1} U(f(k_t) - k_{t+1}) \right] \right\} \\
 &= \max_{\substack{0 \leq k_1 \leq f(k_0) \\ k_0 \text{ given}}} \left\{ U(f(k_0) - k_1) + \beta \left[ \max_{\substack{0 \leq k_{t+2} \leq f(k_{t+1}) \\ k_1 \text{ given}}} \sum_{t=0}^{\infty} \beta^t U(f(k_{t+1}) - k_{t+2}) \right] \right\}
 \end{aligned}$$

Looking at the maximization problem inside the  $\square$ -brackets and comparing to the original problem (25) we see that the  $\square$ -problem is that of a social planner that, given initial capital stock  $k_1$ , maximizes lifetime utility of the representative agent from period 1 onwards. But agents don't age in our model, the technology or the utility functions doesn't change over time; this suggests that the optimal value of the problem in  $\square$ -brackets is equal to  $w(k_1)$  and hence the

<sup>12</sup>Just a caveat: infinite-dimensional maximization problems may not have a solution even if the  $u$  and  $f$  are well-behaved. So the function  $w$  may not always be well-defined. In our examples, with the assumptions that we made, everything is fine, however.

problem can be rewritten as

$$w(k_0) = \max_{\substack{0 \leq k_1 \leq f(k_0) \\ k_0 \text{ given}}} \{U(f(k_0) - k_1) + \beta w(k_1)\}$$

Again two questions arise:

- 2.1 Under which conditions is the suggestive discussion formally correct? We will come back to this in a little while
- 2.2 Is this progress? Of course, the maximization problem is much easier since, instead of maximizing over infinite sequences we maximize over just one number,  $k_1$ . But we can't really solve the maximization problem, because the function  $w(\cdot)$  appears on the right side, and we don't know this function. The next section shows ways to overcome this problem.

### 3.2.2 Recursive Formulation of Social Planner Problem

The above formulation of the social planners problem with a function on the left and right side of the maximization problem is called recursive formulation. Now we want to study this recursive formulation of the planners problem. Since the function  $w(\cdot)$  is associated with the sequential formulation, let us change notation and denote by  $v(\cdot)$  the corresponding function for the recursive formulation of the problem. Remember the interpretation of  $v(k)$ : it is the discounted lifetime utility of the representative agent from the current period onwards if the social planner is given capital stock  $k$  at the beginning of the current period and allocates consumption across time optimally for the household. This function  $v$  (the so-called value function) solves the following recursion

$$v(k) = \max_{0 \leq k' \leq f(k)} \{U(f(k) - k') + \beta v(k')\} \quad (26)$$

Note again that  $v$  and  $w$  are two very different functions;  $v$  is the value function for the recursive formulation of the planners problem and  $w$  is the corresponding function for the sequential problem. Of course below we want to establish that  $v = w$ , but this is something that we have to prove rather than something that we can assume to hold! The capital stock  $k$  that the planner brings into the current period, result of past decisions, completely determines what allocations are feasible from today onwards. Therefore it is called the “state variable”: it completely summarizes the state of the economy today (i.e. all future options that the planner has). The variable  $k'$  is decided (or controlled) today by the social planner; it is therefore called the “control variable”, because it can be controlled today by the planner.<sup>13</sup>

Equation (26) is a functional equation (the so-called Bellman equation): its solution is a function, rather than a number or a vector. Fortunately the mathematical theory of functional equation is well-developed, so we can draw on

<sup>13</sup>These terms come from control theory, a field in applied mathematics. Control theory is used in many technical applications such as astronautics.

some fairly general results. The functional equation posits that the discounted lifetime utility of the representative agent is given by the utility that this agent receives today,  $U(f(k) - k')$ , plus the discounted lifetime utility from tomorrow onwards,  $\beta v(k')$ . So this formulation makes clear the planners trade-off: consumption (and hence utility) today, versus a higher capital stock to work with (and hence higher discounted future utility) from tomorrow onwards. Hence, for a given  $k$  this maximization problem is much easier to solve than the problem of picking an infinite sequence of capital stocks  $\{k_{t+1}\}_{t=0}^{\infty}$  from before. The only problem is that we have to do this maximization for every possible capital stock  $k$ , and this posits theoretical as well as computational problems. However, it will turn out that the functional equation is much easier to solve than the sequential problem (25) (apart from some very special cases). By solving the functional equation we mean finding a value function  $v$  solving (26) and an optimal policy function  $k' = g(k)$  that describes the optimal  $k'$  for the maximization part in (26), as a function of  $k$ , i.e. for each possible value that  $k$  can take. Again we face several questions associated with equation (26):

1. Under what condition does a solution to the functional equation (26) exist and, if it exist, is unique?
2. Is there a reliable algorithm that computes the solution (by reliable we mean that it always converges to the correct solution, independent of the initial guess for  $v$ )
3. Under what conditions can we solve (26) and be sure to have solved (25), i.e. under what conditions do we have  $v = w$  and equivalence between the optimal sequential allocation  $\{k_{t+1}\}_{t=0}^{\infty}$  and allocations generated by the optimal recursive policy  $g(k)$
4. Can we say something about the qualitative features of  $v$  and  $g$ ?

The answers to these questions will be given in the next two sections: the answers to 1. and 2. will come from the Contraction Mapping Theorem, to be discussed in Section 2.2. The answer to the third question makes up what Richard Bellman called the Principle of Optimality and is discussed in Section 3.1. Finally, under more restrictive assumptions we can characterize the solution to the functional equation  $(v, g)$  more precisely. This will be done in Section 3.2. In the remaining parts of this section we will look at specific examples where we can solve the functional equation by hand. Then we will talk about competitive equilibria and the way we can construct prices so that Pareto optimal allocations, together with these prices, form a competitive equilibrium. This will be our versions of the first and second welfare theorem for the neoclassical growth model.

### 3.2.3 An Example

Consider the following example. Let the period utility function be given by  $U(c) = \ln(c)$  and the aggregate production function be given by  $F(k, n) =$

$k^\alpha n^{1-\alpha}$  and assume full depreciation, i.e.  $\delta = 1$ . Then  $f(k) = k^\alpha$  and the functional equation becomes

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

Remember that the solution to this functional equation is an entire function  $v(\cdot)$ . Now we will apply several methods to solve this functional equation.

**Guess and Verify** We will guess a particular functional form of a solution and then verify that the solution has in fact this form (note that this does not rule out that the functional equation has other solutions). This method works well for the example at hand, but not so well for most other examples that we are concerned with. Let us guess

$$v(k) = A + B \ln(k)$$

where  $A$  and  $B$  are coefficients that are to be determined. The method consists of three steps:

1. Solve the maximization problem on the right hand side, given the guess for  $v$ , i.e. solve

$$\max_{0 \leq k' \leq k^\alpha} \{\ln(k^\alpha - k') + \beta(A + B \ln(k'))\}$$

Obviously the constraints on  $k'$  never bind and the objective function is strictly concave and the constraint set is compact, for any given  $k$ . The first order condition is sufficient for the unique solution. The FOC yields

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

2. Evaluate the right hand side at the optimum  $k' = \frac{\beta B k^\alpha}{1 + \beta B}$ . this yields

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

3. In order for our guess to solve the functional equation, the left hand side of the functional equation, which we have guessed to equal  $\text{LHS} = A + B \ln(k)$  must equal the right hand side, which we just found. If we can find

coefficients  $A, B$  for which this is true, we have found a solution to the functional equation. Equating LHS and RHS yields

$$\begin{aligned} A + B \ln(k) &= -\ln(1 + \beta B) + \alpha \ln(k) + \beta A + \beta B \ln\left(\frac{\beta B}{1 + \beta B}\right) + \alpha \beta B \ln(k) \\ (B - \alpha(1 + \beta B)) \ln(k) &= -A - \ln(1 + \beta B) + \beta A + \beta B \ln\left(\frac{\beta B}{1 + \beta B}\right) \end{aligned} \quad (27)$$

But this equation has to hold for *every* capital stock  $k$ . The right hand side of (27) does not depend on  $k$  but the left hand side does. Hence the right hand side is a constant, and the only way to make the left hand side a constant is to make  $B - \alpha(1 + \beta B) = 0$ . Solving this for  $B$  yields  $B = \frac{\alpha}{1 - \alpha\beta}$ . Since the left hand side of (27) is 0, the right hand side better is, too, for  $B = \frac{\alpha}{1 - \alpha\beta}$ . Therefore the constant  $A$  has to satisfy

$$\begin{aligned} 0 &= -A - \ln(1 + \beta B) + \beta A + \beta B \ln\left(\frac{\beta B}{1 + \beta B}\right) \\ &= -A - \ln\left(\frac{1}{1 - \alpha\beta}\right) + \beta A + \frac{\alpha\beta}{1 - \alpha\beta} \ln(\alpha\beta) \end{aligned}$$

Solving this mess for  $A$  yields

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

We can also determine the optimal policy function  $k' = g(k)$  as

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

Hence our guess was correct: the function  $v^*(k) = A + B \ln(k)$ , with  $A, B$  as determined above, solves the functional equation, with associated policy function  $g(k) = \alpha \beta k^\alpha$ . Note that for this specific example the optimal policy of the social planner is to save a constant fraction  $\alpha\beta$  of total output  $k^\alpha$  as capital stock for tomorrow and consume a constant fraction  $(1 - \alpha\beta)$  of total output today. The fact that these fractions do not depend on the level of  $k$  is very unique to this example and not a property of the model in general. Also note that there may be other solutions to the functional equation; we have just constructed one (actually, for the specific example there are no others, but this needs some proving). Finally, it is straightforward to construct a sequence  $\{k_{t+1}\}_{t=0}^\infty$  from our policy function  $g$  that will turn out to solve the sequential problem (25) (of course for the specific functional forms used in the example): start from  $k_0 = \bar{k}_0$ ,  $k_1 = g(k_0) = \alpha\beta k_0^\alpha$ ,  $k_2 = g(k_1) = \alpha\beta k_1^\alpha = (\alpha\beta)^{1+\alpha} k_0^{2\alpha}$  and in general  $k_t = (\alpha\beta)^{\sum_{j=0}^{t-1} \alpha^j} k_0^{\alpha^t}$ . Obviously, since  $0 < \alpha < 1$  we have that

$$\lim_{t \rightarrow \infty} k_t = (\alpha\beta)^{\frac{1}{1-\alpha}}$$

for all initial conditions  $k_0 > 0$  (which, not surprisingly, is the unique solution to  $g(k) = k$ ).

**Value Function Iteration: Analytical Approach** In the last section we started with a clever guess, parameterized it and used the method of undetermined coefficients (guess and verify) to solve for the solution  $v^*$  of the functional equation. For just about any other than the log-utility, Cobb-Douglas production function case this method would not work; even your most ingenious guesses would fail when trying to be verified.

Consider the following iterative procedure for our previous example

1. Guess an arbitrary function  $v_0(k)$ . For concreteness let's take  $v_0(k) = 0$  for all
2. Proceed recursively by solving

$$v_1(k) = \max_{0 \leq k' \leq k^\alpha} \{\ln(k^\alpha - k') + \beta v_0(k')\}$$

Note that we can solve the maximization problem on the right hand side since we know  $v_0$  (since we have guessed) it. In particular, since  $v_0(k') = 0$  for all  $k'$  we have as optimal solution to this problem

$$k' = g_1(k) = 0 \text{ for all } k$$

Plugging this back in we get

$$v_1(k) = \ln(k^\alpha - 0) + \beta v_0(0) = \ln k^\alpha = \alpha \ln k$$

3. Now we can solve

$$v_2(k) = \max_{0 \leq k' \leq k^\alpha} \{\ln(k^\alpha - k') + \beta v_1(k')\}$$

since we know  $v_1$  and so forth.

4. By iterating on the recursion

$$v_{n+1}(k) = \max_{0 \leq k' \leq k^\alpha} \{\ln(k^\alpha - k') + \beta v_n(k')\}$$

we obtain a sequence of value functions  $\{v_n\}_{n=0}^\infty$  and policy functions  $\{g_n\}_{n=1}^\infty$ . Hopefully these sequences will converge to the solution  $v^*$  and associated policy  $g^*$  of the functional equation. In fact, below we will state and prove a very important theorem asserting exactly that (under certain conditions) this iterative procedure converges for any initial guess and converges to the correct solution, namely  $v^*$ .

In the first homework I let you carry out the first few iterations in this procedure. Note however, that, in order to find the solution  $v^*$  exactly you would have to carry out step 2. above a lot of time (in fact, infinitely many times), which is, of course, infeasible. Therefore one has to implement this procedure numerically on a computer.

**Value Function Iteration: Numerical Approach** Even a computer can carry out only a finite number of calculation and can only store finite-dimensional objects. Hence the best we can hope for is a numerical approximation of the true value function. The functional equation above is defined for all  $k \geq 0$  (in fact there is an upper bound, but let's ignore this for now). Because computer storage space is finite, we will approximate the value function for a finite number of points only.<sup>14</sup> For the sake of the argument suppose that  $k$  and  $k'$  can only take values in  $\mathcal{K} = \{0.04, 0.08, 0.12, 0.16, 0.2\}$ . Note that the value functions  $v_n$  then consists of 5 numbers,  $(v_n(0.04), v_n(0.08), v_n(0.12), v_n(0.16), v_n(0.2))$

Now let us implement the above algorithm numerically. First we have to pick concrete values for the parameters  $\alpha$  and  $\beta$ . Let us pick  $\alpha = 0.3$  and  $\beta = 0.6$ .

1. Make an initial guess  $v_0(k) = 0$  for all  $k \in \mathcal{K}$
2. Solve

$$v_1(k) = \max_{\substack{0 \leq k' \leq k^{0.3} \\ k' \in \mathcal{K}}} \{ \ln(k^{0.3} - k') + 0.6 * 0 \}$$

This obviously yields as optimal policy  $k'(k) = g_1(k) = 0.04$  for all  $k \in \mathcal{K}$  (note that since  $k' \in \mathcal{K}$  is required,  $k' = 0$  is not allowed). Plugging this back in yields

$$\begin{aligned} v_1(0.04) &= \ln(0.04^{0.3} - 0.04) = -1.077 \\ v_1(0.08) &= \ln(0.08^{0.3} - 0.04) = -0.847 \\ v_1(0.12) &= \ln(0.12^{0.3} - 0.04) = -0.715 \\ v_1(0.16) &= \ln(0.16^{0.3} - 0.04) = -0.622 \\ v_1(0.2) &= \ln(0.2^{0.3} - 0.04) = -0.55 \end{aligned}$$

3. Let's do one more step by hand

$$v_2(k) = \left\{ \max_{\substack{0 \leq k' \leq k^{0.3} \\ k' \in \mathcal{K}}} \ln(k^{0.3} - k') + 0.6v_1(k') \right\}$$

Start with  $k = 0.04$  :

$$v_2(0.04) = \max_{\substack{0 \leq k' \leq 0.04^{0.3} \\ k' \in \mathcal{K}}} \{ \ln(0.04^{0.3} - k') + 0.6v_1(k') \}$$

---

<sup>14</sup>In this introductory course I will only discuss so-called finite state-space methods, i.e. methods in which the state variable (and the control variable) can take only a finite number of values. Ken Judd, one of the world leaders in numerical methods in economics teaches an excellent second year class in computational methods, in which much more sophisticated methods for solving similar problems are discussed. I strongly encourage you to take this course at some point of your career here in Stanford.

Since  $0.04^{0.3} = 0.381$  all  $k' \in \mathcal{K}$  are possible. If the planner chooses  $k' = 0.04$ , then

$$v_2(0.04) = \ln(0.04^{0.3} - 0.04) + 0.6 * (-1.077) = -1.723$$

If he chooses  $k' = 0.08$ , then

$$v_2(0.04) = \ln(0.04^{0.3} - 0.08) + 0.6 * (-0.847) = -1.710$$

If he chooses  $k' = 0.12$ , then

$$v_2(0.04) = \ln(0.04^{0.3} - 0.12) + 0.6 * (-0.715) = -1.773$$

If  $k' = 0.16$ , then

$$v_2(0.04) = \ln(0.04^{0.3} - 0.16) + 0.6 * (-0.622) = -1.884$$

Finally, if  $k' = 0.2$ , then

$$v_2(0.04) = \ln(0.04^{0.3} - 0.2) + 0.6 * (-0.55) = -2.041$$

Hence for  $k = 0.04$  the optimal choice is  $k'(0.04) = g_2(0.04) = 0.08$  and  $v_2(0.04) = -1.710$ . This we have to do for all  $k \in \mathcal{K}$ . One can already see that this is quite tedious by hand, but also that a computer can do this quite rapidly. Table 1 below shows the value of

$$(k^{0.3} - k') + 0.6v_1(k')$$

for different values of  $k$  and  $k'$ . A \* in the column for  $k'$  that this  $k'$  is the optimal choice for capital tomorrow, for the particular capital stock  $k$  today

**Table 1**

$k$	$k'$	0.04	0.08	0.12	0.16	0.2
0.04		-1.7227	-1.7097*	-1.7731	-1.8838	-2.0407
0.08		-1.4929	-1.4530*	-1.4822	-1.5482	-1.6439
0.12		-1.3606	-1.3081*	-1.3219	-1.3689	-1.4405
0.16		-1.2676	-1.2072*	-1.2117	-1.2474	-1.3052
0.2		-1.1959	-1.1298	-1.1279*	-1.1560	-1.2045

Hence the value function  $v_2$  and policy function  $g_2$  are given by

**Table 2**

$k$	$v_2(k)$	$g_2(k)$
0.04	-1.7097	0.08
0.08	-1.4530	0.08
0.12	-1.3081	0.08
0.16	-1.2072	0.08
0.2	-1.1279	0.12

In Figure 1 we plot the true value function  $v^*$  (remember that for this example we know to find  $v^*$  analytically) and selected iterations from the numerical value function iteration procedure. In Figure 2 we have the corresponding policy functions.

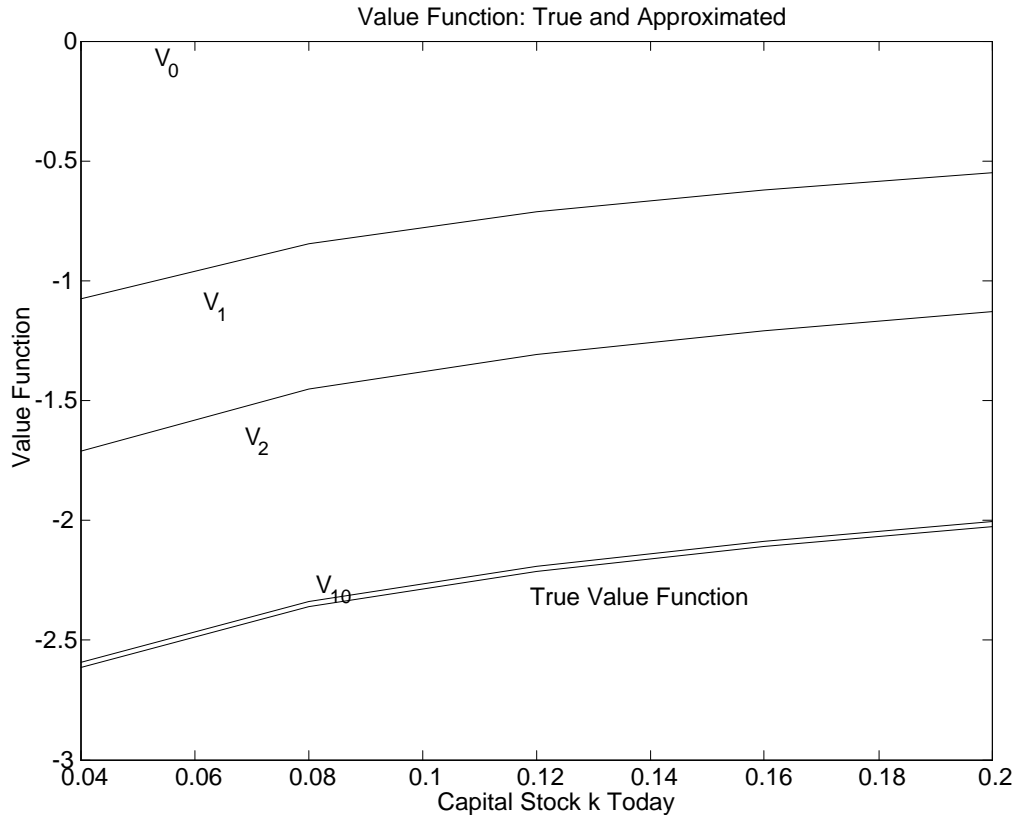


Figure 1:

We see from Figure 1 that the numerical approximations of the value function converge rapidly to the true value function. After 20 iterations the approximation and the truth are nearly indistinguishable with the naked eye. Looking at the policy functions we see from Figure 2 that the approximating policy function

do not converge to the truth (more iterations don't help). This is due to the fact that the analytically correct value function was found by allowing  $k' = g(k)$  to take any value in the real line, whereas for the approximations we restricted  $k' = g_n(k)$  to lie in  $\mathcal{K}$ . The function  $g_{10}$  approximates the true policy function as good as possible, subject to this restriction. Therefore the approximating value function will not converge exactly to the truth, either. The fact that the value function approximations come much closer is due to the fact that the utility and production function induce "curvature" into the value function, something that we may make more precise later.

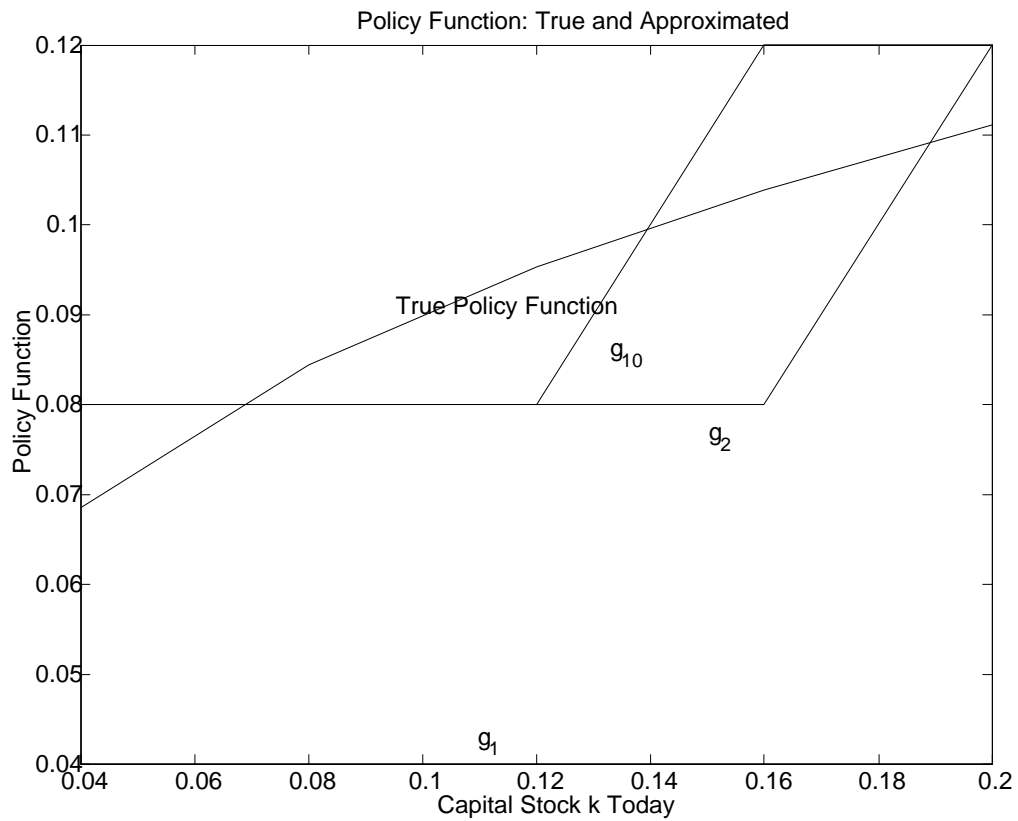


Figure 2:

### 3.2.4 The Euler Equation Approach and Transversality Conditions

We now relate our example to the traditional approach of solving optimization problems. Note that this approach also, as the guess and verify method, will only work in very simple examples, but not in general, whereas the numerical

approach works for a wide range of parameterizations of the neoclassical growth model. First let us look at a finite horizon social planners problem and then at the related infinite-dimensional problem

**The Finite Horizon Case** Let us consider the social planner problem for a situation in which the representative consumer lives for  $T < \infty$  periods, after which she dies for sure and the economy is over. The social planner problem for this case is given by

$$\begin{aligned} w^T(\bar{k}_0) &= \max_{\{k_{t+1}\}_{t=0}^T} \sum_{t=0}^T \beta^t U(f(k_t) - k_{t+1}) \\ 0 &\leq k_{t+1} \leq f(k_t) \\ k_0 &= \bar{k}_0 > 0 \text{ given} \end{aligned}$$

Obviously, since the world goes under after period  $T$ ,  $k_{T+1} = 0$ . Also, given our Inada assumptions on the utility function the constraints on  $k_{t+1}$  will never be binding and we will disregard them henceforth. The first thing we note is that, since we have a finite-dimensional maximization problem and since the set constraining the choices of  $\{k_{t+1}\}_{t=0}^T$  is closed and bounded, by the Bolzano-Weierstrass theorem a solution to the maximization problem exists, so that  $w^T(\bar{k}_0)$  is well-defined. Furthermore, since the constraint set is convex and we assumed that  $U$  is strictly concave (and the finite sum of strictly concave functions is strictly concave), the solution to the maximization problem is unique and the first order conditions are not only necessary, but also sufficient.

Forming the Lagrangian yields

$$L = U(f(k_0) - k_1) + \dots + \beta^t U(f(k_t) - k_{t+1}) + \beta^{t+1} U(f(k_{t+1}) - k_{t+2}) + \dots + \beta^T U(f(k_T) - k_{T+1})$$

and hence we can find the first order conditions as

$$\frac{\partial L}{\partial k_{t+1}} = -\beta^t U'(f(k_t) - k_{t+1}) + \beta^{t+1} U'(f(k_{t+1}) - k_{t+2}) f'(k_{t+1}) = 0 \quad \text{for all } t = 0, \dots, T-1$$

or

$$\underbrace{U'(f(k_t) - k_{t+1})}_{\substack{\text{Cost in utility} \\ \text{for saving} \\ \text{1 unit more} \\ \text{capital for } t+1}} = \underbrace{\beta U'(f(k_{t+1}) - k_{t+2})}_{\substack{\text{Discounted} \\ \text{add. utility} \\ \text{from one more} \\ \text{unit of cons.}}} \underbrace{f'(k_{t+1})}_{\substack{\text{Add. production} \\ \text{possible with} \\ \text{one more unit} \\ \text{of capital in } t+1}} \quad \text{for all } t = 0, \dots, T-1 \quad (28)$$

The interpretation of the optimality condition is easiest with a variational argument. Suppose the social planner in period  $t$  contemplates whether to save one more unit of capital for tomorrow. One more unit saved reduces consumption

by one unit, at utility cost of  $U'(f(k_t) - k_{t+1})$ . On the other hand, there is one more unit of capital to produce with tomorrow, yielding additional production  $f'(k_{t+1})$ . Each additional unit of production, when used for consumption, is worth  $U'(f(k_{t+1}) - k_{t+2})$  utiles tomorrow, and hence  $\beta U'(f(k_{t+1}) - k_{t+2})$  utiles today. At the optimum the net benefit of such a variation in allocations must be zero, and the result is the first order condition above. This first order condition some times is called an Euler equation (supposedly because it is loosely linked to optimality conditions in continuous time calculus of variations, developed by Euler). Equations (28) is second order difference equation, a system of  $T$  equations in the  $T + 1$  unknowns  $\{k_{t+1}\}_{t=0}^T$  (with  $k_0$  predetermined). However, we have the terminal condition  $k_{T+1} = 0$  and hence, under appropriate conditions, can solve for the optimal  $\{k_{t+1}\}_{t=0}^T$  uniquely. We can demonstrate this for our example from above.

Again let  $U(c) = \ln(c)$  and  $f(k) = k^\alpha$ . Then (28) becomes

$$\begin{aligned} \frac{1}{k_t^\alpha - k_{t+1}} &= \frac{\beta \alpha k_{t+1}^{\alpha-1}}{k_{t+1}^\alpha - k_{t+2}} \\ k_{t+1}^\alpha - k_{t+2} &= \alpha \beta k_{t+1}^{\alpha-1} (k_t^\alpha - k_{t+1}) \end{aligned} \quad (29)$$

with  $k_0 > 0$  given and  $k_{T+1} = 0$ . A little trick will make our life easier. Define  $z_t = \frac{k_{t+1}}{k_t^\alpha}$ . The variable  $z_t$  is the fraction of output in period  $t$  that is saved as capital for tomorrow, so we can interpret  $z_t$  as the saving rate of the social planner. Dividing both sides of (29) by  $k_{t+1}^\alpha$  we get

$$\begin{aligned} 1 - z_{t+1} &= \frac{\alpha \beta (k_t^\alpha - k_{t+1})}{k_{t+1}^\alpha} = \alpha \beta \left( \frac{1}{z_t} - 1 \right) \\ z_{t+1} &= 1 + \alpha \beta - \frac{\alpha \beta}{z_t} \end{aligned}$$

This is a first order difference equation. Since we have the boundary condition  $k_{T+1} = 0$ , this implies  $z_T = 0$ , so we can solve this equation backwards. Rewriting yields

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

We can now recursively solve backwards for the entire sequence  $\{z_t\}_{t=0}^T$ , given that we know  $z_T = 0$ . We obtain as general formula (verify this by plugging it into the first order difference equation)

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

and hence

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

One can also solve for the discounted future utility at time zero from the optimal allocation to obtain

$$\begin{aligned}
w^T(k_0) &= \alpha \ln(k_0) \sum_{j=0}^T (\alpha\beta)^j - \sum_{j=1}^T \beta^{T-j} \ln \left( \sum_{i=0}^j (\alpha\beta)^i \right) \\
&\quad + \alpha\beta \sum_{j=1}^T \beta^{T-j} \left[ \sum_{i=0}^{j-1} (\alpha\beta)^i \left\{ \ln(\alpha\beta) + \ln \left( \frac{\sum_{i=0}^{j-1} (\alpha\beta)^i}{\sum_{i=0}^j (\alpha\beta)^i} \right) \right\} \right]
\end{aligned}$$

Note that the optimal policies and the discounted future utility are functions of the time horizon that the social planner faces. Also note that for this specific example

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

and

$$\lim_{t \rightarrow \infty} w^T(k_0) = \frac{1}{1-\beta} \left[ \frac{\alpha\beta}{1-\alpha\beta} \ln(\alpha\beta) + \ln(1-\alpha\beta) \right] + \frac{\alpha}{1-\alpha\beta} \ln(k_0)$$

So is this the case that the optimal policy for the social planners problem with infinite time horizon is the limit of the optimal policies for the  $T$ -horizon planning problem (and the same is true for the value of the planning problem)? Our results from the guess and verify method seem to indicate this, and for this example this is indeed true, but a) this needs some proof and b) it is not at all true in general, but very specific to the example we considered.<sup>15</sup> We can't in general interchange maximization and limit-taking: the limit of the finite maximization problems is not necessary equal to maximization of the problem in which time goes to infinity.

In order to prepare for the discussion of the infinite horizon case let us analyze the first order difference equation

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

---

<sup>15</sup>An easy counterexample is the cake-eating problem without discounting

$$\begin{aligned}
&\max_{\{c_t\}_{t=0}^T} \sum_{t=0}^{\infty} u(c_t) \\
s.t. \quad &1 = \sum_{t=0}^T
\end{aligned}$$

with  $u$  bounded and strictly concave, whose solution for the finite time horizon is obviously  $c_t = \frac{1}{T+1}$ , for all  $t$ . The limit as  $T \rightarrow \infty$  would be  $c_t = 0$ , which obviously can't be optimal (for example  $c_t = (1-a)a^t$ , for any  $a \in (0,1)$  beats that policy.

graphically. On the y-axis of Figure 3 we draw  $z_{t+1}$  against  $z_t$  on the x-axis. Since  $k_{t+1} \geq 0$ , we have that  $z_t \geq 0$  for all  $t$ . Furthermore, as  $z_t$  approaches 0 from above,  $z_{t+1}$  approaches  $-\infty$ . As  $z_t$  approaches  $+\infty$ ,  $z_{t+1}$  approaches  $1 + \alpha\beta$  from below asymptotically. The graph intersects the x-axis at  $z^0 = \frac{\alpha\beta}{1+\alpha\beta}$ . The difference equations has two steady states where  $z_{t+1} = z_t = z$ . This can be seen by

$$\begin{aligned} z &= 1 + \alpha\beta - \frac{\alpha\beta}{z} \\ z^2 - (1 + \alpha\beta)z + \alpha\beta &= 0 \\ (z - 1)(z - \alpha\beta) &= 0 \\ z &= 1 \text{ or } z = \alpha\beta \end{aligned}$$

From Figure 3 we can also determine graphically the sequence of optimal policies  $\{z_t\}_{t=0}^T$ . We start with  $z_T = 0$  on the y-axis, go to the  $z_{t+1} = 1 + \alpha\beta - \frac{\alpha\beta}{z_t}$  curve to determine  $z_{T-1}$  and mirror it against the 45-degree line to obtain  $z_{T-1}$  on the y-axis. Repeating the argument one obtains the entire  $\{z_t\}_{t=0}^T$  sequence, and hence the entire  $\{k_{t+1}\}_{t=0}^T$  sequence. Note that going with  $t$  backwards to zero, the  $z_t$  approach  $\alpha\beta$ . Hence for large  $T$  and for small  $t$  (the optimal policies for a finite time horizon problem with long horizon, for the early periods) come close to the optimal infinite time horizon policies solved for with the guess and verify method.

**The Infinite Horizon Case** Now let us turn to the infinite horizon problem and let's see how for we can get with the Euler equation approach. Remember that the problem was to

$$\begin{aligned} w(\bar{k}_0) &= \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1}) \\ 0 &\leq k_{t+1} \leq f(k_t) \\ k_0 &= \bar{k}_0 > 0 \text{ given} \end{aligned}$$

Since the period utility function is strictly concave and the constraint set is convex, the first order conditions constitute necessary conditions for an optimal sequence  $\{k_{t+1}^*\}_{t=0}^{\infty}$  (a proof of this is a formalization of the variational argument I spelled out when discussing the intuition for the Euler equation). As a reminder, the Euler equations were

$$\beta U'(f(k_{t+1}) - k_{t+2}) f'(k_{t+1}) = U'(f(k_t) - k_{t+1}) \quad \text{for all } t = 0, \dots, t, \dots$$

Again this is a second order difference equation, but now we only have an initial condition for  $k_0$ , but no terminal condition since there is no terminal time period. In a lot of applications, the transversality condition substitutes for the missing

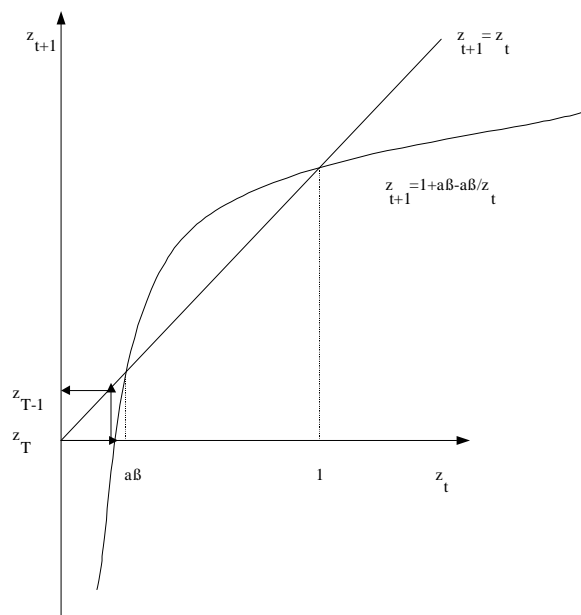


Figure 3:

terminal condition. Let us first state and then interpret the TVC<sup>16</sup>

$$\lim_{t \rightarrow \infty} \underbrace{\beta^t U'(f(k_t) - k_{t+1}) f'(k_t)}_{\substack{\text{value in discounted} \\ \text{utility terms of one} \\ \text{more unit of capital}}} \underbrace{k_t}_{\substack{\text{Total} \\ \text{Capital} \\ \text{Stock}}} = 0$$

The transversality condition states that the value of the capital stock  $k_t$ , when measured in terms of discounted utility, goes to zero as time goes to infinity. Note that this condition does not require that the capital stock itself converges to zero in the limit, only the (shadow) value of the capital stock has to converge to zero.

The transversality condition is a tricky beast, and we may spend some more time on it next quarter. For now we just state the following theorem.

**Theorem 11** *Let  $U, \beta$  and  $F$  (and hence  $f$ ) satisfy assumptions 1. and 2. Then an allocation  $\{k_{t+1}\}_{t=0}^{\infty}$  that satisfies the Euler equations and the transversality condition solves the sequential social planners problem, for a given  $k_0$ .*

This theorem states that under certain assumptions the Euler equations and the transversality condition are jointly sufficient for a solution to the social planners problem in sequential formulation. Stokey et al., p. 98-99 prove this theorem. Note that this theorem does not apply for the case in which the utility function is logarithmic; however, the proof that Stokey et al. give can easily be

---

<sup>16</sup>Often one can find an alternative statement of the TVC.

$$\lim_{t \rightarrow \infty} \lambda_t k_{t+1} = 0$$

where  $\lambda_t$  is the Lagrange multiplier on the constraint

$$c_t + k_{t+1} = f(k_t)$$

in the social planner in which consumption is not yet substituted out. From the first order condition we have

$$\begin{aligned} \beta^t U'(c_t) &= \lambda_t \\ \beta^t U'(f(k_t) - k_{t+1}) &= \lambda_t \end{aligned}$$

Hence the TVC becomes

$$\lim_{t \rightarrow \infty} \beta^t U'(f(k_t) - k_{t+1}) k_{t+1} = 0$$

This condition is equivalent to the condition given in the main text, as shown by the following argument (which uses the Euler equation)

$$\begin{aligned} 0 &= \lim_{t \rightarrow \infty} \beta^t U'(f(k_t) - k_{t+1}) k_{t+1} \\ &= \lim_{t \rightarrow \infty} \beta^{t-1} U'(f(k_{t-1}) - k_t) k_t \\ &= \lim_{t \rightarrow \infty} \beta^{t-1} \beta U'(f(k_t) - k_{t+1}) f'(k_t) k_t \\ &= \lim_{t \rightarrow \infty} \beta^t U'(f(k_t) - k_{t+1}) f'(k_t) k_t \end{aligned}$$

which is the TVC in the main text.

extended to the log-case. So although the Euler equations and the TVC may not be sufficient for every unbounded utility function, for the log-case they are.

Also note that we have said nothing about the necessity of the TVC. We have (loosely) argued that the Euler equations are necessary conditions, but is the TVC necessary, i.e. does every solution to the sequential planning problem have to satisfy the TVC? This turns out to be a hard problem, and there is not a very general result for this. However, for the log-case (with  $f$ 's satisfying our assumptions), Ekelund and Scheinkman (1985) show that the TVC is in fact a necessary condition. Refer to their paper and to the related results by Peleg and Ryder (1972) and Weitzman (1973) for further details. From now on we assert that the TVC is necessary and sufficient for optimization under the assumptions we made on  $f, U$ , but you should remember that these assertions remain to be proved.

For now we take these theoretical results for granted and proceed with our example of  $U(c) = \ln(c)$ ,  $f(k) = k^\alpha$ . For these particular functional forms, the TVC becomes

$$\begin{aligned} & \lim_{t \rightarrow \infty} \beta^t U'(f(k_t) - k_{t+1}) f'(k_t) k_t \\ &= \lim_{t \rightarrow \infty} \frac{\alpha \beta^t k_t^\alpha}{k_t^\alpha - k_{t+1}} = \lim_{t \rightarrow \infty} \frac{\alpha \beta^t}{1 - \frac{k_{t+1}}{k_t^\alpha}} \\ &= \lim_{t \rightarrow \infty} \frac{\alpha \beta^t}{1 - z_t} \end{aligned}$$

We also repeat the first order difference equation derived from the Euler equations

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

We can't solve the Euler equations form  $\{z_t\}_{t=0}^\infty$  backwards, but we can solve it forwards, conditional on guessing an initial value for  $z_0$ . We show that only one guess for  $z_0$  yields a sequence that does not violate the TVC or the nonnegativity constraint on capital or consumption.

1.  $z_0 < \alpha\beta$ . From Figure 3 we see that in finite time  $z_t < 0$ , violating the nonnegativity constraint on capital
2.  $z_0 > \alpha\beta$ . Then from Figure 3 we see that  $\lim_{t \rightarrow \infty} z_t = 1$ . (Note that, in fact, every  $z_0 > 1$  violate the nonnegativity of consumption and hence is not admissible as a starting value). We will argue that all these paths violate the TVC.
3.  $z_0 = \alpha\beta$ . Then  $z_t = \alpha\beta$  for all  $t > 0$ . For this path (which obviously satisfies the Euler equations) we have that

$$\lim_{t \rightarrow \infty} \frac{\alpha \beta^t}{1 - z_t} = \lim_{t \rightarrow \infty} \frac{\alpha \beta^t}{1 - \alpha \beta} = 0$$

and hence this sequence satisfies the TVC. From the sufficiency of the Euler equation jointly with the TVC we conclude that the sequence  $\{z_t\}_{t=0}^{\infty}$  given by  $z_t = \alpha\beta$  is an optimal solution for the sequential social planners problem. Translating into capital sequences yields as optimal policy  $k_{t+1} = \alpha\beta k_t^\alpha$ , with  $k_0$  given. But this is exactly the constant saving rate policy that we derived as optimal in the recursive problem.

Now we pick up the unfinished business from point 2. Note that we asserted above (citing Ekelund and Scheinkman) that for our particular example the TVC is a necessary condition, i.e. any sequence  $\{k_{t+1}\}_{t=0}^{\infty}$  that does not satisfy the TVC can't be an optimal solution.

Since all sequences  $\{z_t\}_{t=0}^{\infty}$  in 2. converge to 1, in the TVC both the nominator and the denominator go to zero. Let us linearly approximate  $z_{t+1}$  around the steady state  $z = 1$ . This gives

$$\begin{aligned}
z_{t+1} &= 1 + \alpha\beta - \frac{\alpha\beta}{z_t} = g(z_t) \\
z_{t+1} &\approx g(1) + (z_t - 1)g'(z_t)|_{z_t=1} \\
&= 1 + (z_t - 1) \left( \frac{\alpha\beta}{z_t^2} \right) |_{z_t=1} \\
&= 1 + \alpha\beta(z_t - 1) \\
(1 - z_{t+1}) &\approx \alpha\beta(1 - z_t) \\
&\approx \alpha\beta^{t+k-1}(1 - z_k) \quad \text{for all } k
\end{aligned}$$

Hence

$$\lim_{t \rightarrow \infty} \frac{\alpha\beta^{t+1}}{1 - z_{t+1}} \approx \lim_{t \rightarrow \infty} \frac{\beta^k}{\alpha^{t-k}(1 - z_k)} = \infty$$

as  $0 < \alpha < 1$ . Hence non of the sequences contemplated in 2. can be an optimal solution, and our solution found in 3. is indeed the unique optimal solution to the infinite-dimensional social planner problem.<sup>17</sup> Hence in this specific case the Euler equation approach, augmented by the TVC works. But as with the guess-and-verify method this is very unique to specific example at hand. Therefore for the general case we can't rely on pencil and paper, but have to resort to computational techniques. To make sure that these techniques give the desired answer, we have to study the general properties of the functional equation associated with the sequential social planner problem and the relation of its solution to the solution of the sequential problem. We will do this in Sections 2 and 3, respectively. Before this we will show that, by solving the social planners problem we have, in effect, solved for a (the) competitive equilibrium in this economy.

---

<sup>17</sup>Note that we have to (and can) pick  $k$  such that the linear approximation of  $z_{t+1}$  is accurate enough.

### 3.3 Competitive Equilibrium Growth

Suppose we have solved the social planners problem for a Pareto efficient allocation  $\{c_t^*, k_{t+1}^*\}_{t=0}^\infty$ . What we are genuinely interested in are allocations and prices that arise when firms and consumers interact in markets. These markets may be perfectly competitive, in the sense that consumers and firms act as price takers, or may entail strategic interaction between consumers and/or firms. In this section we will discuss the connection between Pareto optimal allocations and allocations arising in a competitive equilibrium. There is usually no such connection between Pareto optimal allocations and allocations arising in situations in which agents act strategically. So for the moment we leave these situations to the game theorists.

For the discussion of Pareto optimal allocations it did not matter who owns what in the economy, since the planner was allowed to freely redistribute endowments across agents. For a competitive equilibrium the question of ownership is crucial. We make the following assumption on the ownership structure of the economy: we assume that consumers own all factors of production (i.e. they own the capital stock at all times) and rent it out to the firms. We also assume that households own the firms, i.e. are claimants of the firms profits.

Now we have to specify the equilibrium concept and the market structure. We assume that the final goods market and the factor markets (for labor and capital services) are perfectly competitive, which means that households as well as firms take prices as given and beyond their control. We assume that there is a single market at time zero in which goods from all future periods are traded. After this market closes, in all future periods the agents in the economy just carry out the trades they agreed upon in period 0. We assume that all contracts are perfectly enforceable. This market is often called Arrow-Debreu market structure and the corresponding competitive equilibrium an Arrow-Debreu equilibrium.

For each period there are three goods that are traded:

1. the final output good,  $y_t$  that can be used for consumption  $c_t$  or investment. Let  $p_t$  denote the price of the period  $t$  final output good, quoted in period 0.
2. labor services  $n_t$ . Let  $w_t$  be the price of one unit of labor services delivered in period  $t$ , quoted in period 0, in terms of the period  $t$  consumption good. Hence  $w_t$  is the real wage; it tells how many units of the period  $t$  consumption goods one can buy for the receipts for one unit of labor. The nominal wage is  $p_t w_t$ .
3. capital services  $k_t$ . Let  $r_t$  be the rental price of one unit of capital services delivered in period  $t$ , quoted in period 0, in terms of the period  $t$  consumption good.  $r_t$  is the real rental rate of capital, the nominal rental rate is  $p_t r_t$ .

Figure 4 summarizes the flows of goods and payments in the economy (note that, since all trade takes place in period 0, no payments are made after period 0)

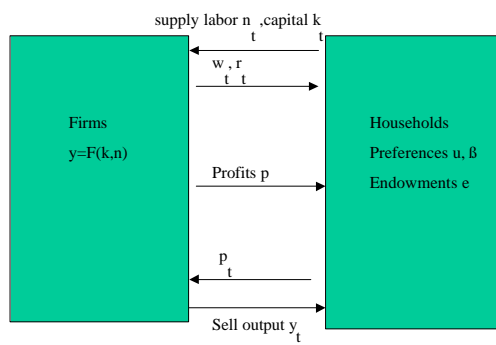


Figure 4:

### 3.3.1 Definition of Competitive Equilibrium

Now we will define a competitive equilibrium for this economy. Let us first look at firms. Without loss of generality assume that there is a single, representative firm that behaves competitively (note: when making this assumption for firms, this is a completely innocuous assumption as long as the technology features constant returns to scale. We will come back to this point). The representative firm's problem is , given a sequence of prices  $\{p_t, w_t, r_t\}_{t=0}^{\infty}$

$$\begin{aligned} \pi &= \max_{\{y_t, k_t, n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t (y_t - r_t k_t - w_t n_t) & (30) \\ \text{s.t. } y_t &= F(k_t, n_t) \text{ for all } t \geq 0 \\ y_t, k_t, n_t &\geq 0 \end{aligned}$$

Hence firms chose an infinite sequence of inputs  $\{k_t, n_t\}$  to maximize total profits  $\pi$ . Since in each period all inputs are rented (the firm does not make the capital accumulation decision), there is nothing dynamic about the firm's problem and it will separate into an infinite number of static maximization problems. More later.

Households face a fully dynamic problem in this economy. They own the capital stock and hence have to decide how much labor and capital services to supply, how much to consume and how much capital to accumulate. Taking prices  $\{p_t, w_t, r_t\}_{t=0}^{\infty}$  the representative consumer solves

$$\begin{aligned} & \max_{\{c_t, i_t, x_{t+1}, k_t, n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t) & (31) \\ \text{s.t. } \sum_{t=0}^{\infty} p_t (c_t + i_t) &\leq \sum_{t=0}^{\infty} p_t (r_t k_t + w_t n_t) + \pi \\ x_{t+1} &= (1 - \delta)x_t + i_t \quad \text{all } t \geq 0 \\ 0 &\leq n_t \leq 1, 0 \leq k_t \leq x_t \quad \text{all } t \geq 0 \\ c_t, x_{t+1} &\geq 0 \quad \text{all } t \geq 0 \\ x_0 &\text{ given} \end{aligned}$$

A few remarks are in order. First, there is only one, time zero budget constraint, the so-called Arrow-Debreu budget constraint, as markets are only open in period 0. Secondly we carefully distinguish between the capital stock  $x_t$  and capital services that households supply to the firm. Capital services are traded and hence have a price attached to them, the capital stock  $x_t$  remains in the possession of the household, is never traded and hence does not have a price attached to it.<sup>18</sup> We have implicitly assumed two things about technology: a)

<sup>18</sup>This is not quite correct: we do not require investment  $i_t$  to be positive. To the extent that  $i_t < -c_t$  is chosen by households, households in fact could transform part of the capital stock back into final output goods and sell it back to the firm. In equilibrium this will never happen of course, since it would require negative production of firms (or free disposal, which we ruled out).

the capital stock depreciates no matter whether it is rented out to the firm or not and b) there is a technology for households that transforms one unit of the capital stock at time  $t$  into one unit of capital services at time  $t$ . The constraint  $k_t \leq x_t$  then states that households cannot provide more capital services than the capital stock at their disposal produces. Also note that we only require the capital stock to be nonnegative, but not investment. In this sense the capital stock is “putty-putty”. We are now ready to define a competitive equilibrium for this economy.

**Definition 12** *A Competitive Equilibrium (Arrow-Debreu Equilibrium) consists of prices  $\{p_t, w_t, r_t\}_{t=0}^{\infty}$  and allocations for the firm  $\{k_t^d, n_t^d, y_t\}_{t=0}^{\infty}$  and the household  $\{c_t, i_t, x_{t+1}, k_t^s, n_t^s\}_{t=0}^{\infty}$  such that*

1. *Given prices  $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ , the allocation of the representative firm  $\{k_t^d, n_t^d, y_t\}_{t=0}^{\infty}$  solves (30)*
2. *Given prices  $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ , the allocation of the representative household  $\{c_t, i_t, x_{t+1}, k_t^s, n_t^s\}_{t=0}^{\infty}$  solves (31)*
3. *Markets clear*

$$\begin{aligned} y_t &= c_t + i_t \text{ (Goods Market)} \\ n_t^d &= n_t^s \text{ (Labor Market)} \\ k_t^d &= k_t^s \text{ (Capital Services Market)} \end{aligned}$$

### 3.3.2 Representation of the Economy in Arrow-Debreu Language

See section 5.1

### 3.3.3 Characterization of the Competitive Equilibrium and the Welfare Theorems

Let us start with a partial characterization of the competitive equilibrium. First of all we simplify notation and denote by  $k_t = k_t^d = k_t^s$  the equilibrium demand and supply of capital services. Similarly  $n_t = n_t^d = n_t^s$ . It is straightforward to show that in any equilibrium  $p_t > 0$  for all  $t$ , since the utility function is strictly increasing in consumption (and therefore consumption demand would be infinite at a zero price). But then, since the production function exhibits positive marginal products,  $r_t, w_t > 0$  in any competitive equilibrium because otherwise factor demands would become unbounded.

Now let us analyze the problem of the representative firm. As stated earlier, the firms do not face a dynamic decision problem as the variables chosen at period  $t$ ,  $(y_t, k_t, n_t)$  do not affect the constraints nor returns (profits) at later periods. The static profit maximization problem for the representative firm is given by

$$\max_{k_t, n_t \geq 0} p_t (F(k_t, n_t) - r_t k_t - w_t n_t)$$

Since the firm take prices as given, the usual “factor price equals marginal product” conditions arise

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

Note that this implies that the profits the firms earns in period  $t$  are equal to

$$\pi_t = p_t (F(k_t, n_t) - F_k(k_t, n_t)k_t - F_n(k_t, n_t)n_t) = 0$$

This follows from the assumption that the function  $F$  exhibits constant returns to scale (is homogeneous of degree 1)

$$F(\lambda k, \lambda n) = \lambda F(k, n) \text{ for all } \lambda > 0$$

and from Euler’s theorem<sup>19</sup> which implies that

$$F(k_t, n_t) = F_k(k_t, n_t)k_t + F_n(k_t, n_t)n_t$$

Therefore the total profits of the representative firm are equal to zero in equilibrium. This argument in fact shows that with CRTS the number of firms is indeterminate in equilibrium; it could be one firm, two firms each operating at half the scale of the one firm or 10 million firms. So this really justifies that the assumption of a single representative firm is without any loss of generality (as long as we assume that this firm acts as a price taker). It also follows that the representative household, as owner of the firm, does not receive any profits in equilibrium.

Let’s turn to that infamous representative household. Given that output and factor prices have to be positive in equilibrium it is clear that the utility maximizing choices of the household entail

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

---

<sup>19</sup>Euler’s theorem states that for any function that is homogeneous of degree  $k$  and differentiable at  $x \in \mathbf{R}^L$  we have

$$kf(x) = \sum_{i=1}^L x_i \frac{\partial f(x)}{\partial x_i}$$

**Proof.** Since  $f$  is homogeneous of degree  $k$  we have for all  $\lambda > 0$

$$f(\lambda x) = \lambda^k f(x)$$

Differentiating both sides with respect to  $\lambda$  yields

$$\sum_{i=1}^L x_i \frac{\partial f(\lambda x)}{\partial x_i} = k\lambda^{k-1} f(x)$$

Setting  $\lambda = 1$  yields

$$\sum_{i=1}^L x_i \frac{\partial f(x)}{\partial x_i} = kf(x)$$

■

From the equilibrium condition in the goods market we also obtain

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

Since utility is strictly increasing in consumption, we also can conclude that the Arrow-Debreu budget constraint holds with equality in equilibrium. Using the first results we can rewrite the household problem as

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

Again the first order conditions are necessary for a solution to the household optimization problem. Attaching  $\mu$  to the Arrow-Debreu budget constraint and ignoring the nonnegativity constraints on consumption and capital stock we get as first order conditions<sup>20</sup> with respect to  $c_t, c_{t+1}$  and  $k_{t+1}$

$$\begin{aligned} \beta^t U'(c_t) &= \mu p_t \\ \beta^{t+1} U'(c_{t+1}) &= \mu p_{t+1} \\ \mu p_t &= \mu(1 - \delta + r_{t+1})p_{t+1} \end{aligned}$$

Combining yields the Euler equation

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

Now we use the marginal pricing condition and the fact that we defined  $f(k_t) = F(k_t, 1) + (1 - \delta)k_t$

$$r_t = F_k(k_t, 1) = f'(k_t) - (1 - \delta)$$

and the market clearing condition from the goods market

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

---

<sup>20</sup>That the nonnegativity constraints on consumption do not bind follows directly from the Inada conditions. The nonnegativity constraints on capital could potentially bind if we look at the household problem in isolation. However, since from the production function  $k_t = 0$  implies  $F(k_t, 1) = 0$  and hence  $c_t = 0$  (which can never happen in equilibrium) we take the shortcut and ignore the corners with respect to capital holdings. But you should be aware of the fact that we did something here that was not very kosher, we implicitly imposed an equilibrium condition before carrying out the maximization problem of the household. This is OK here, but may lead to a lot of havoc when used in other circumstances.

in the Euler equation to obtain

$$\frac{f'(k_{t+1})\beta U'(f(k_{t+1}) - k_{t+2})}{U'(f(k_t) - k_{t+1})} = 1$$

which is exactly the same Euler equation as in the social planners problem. But as with the social planners problem the households' maximization problem is an infinite-dimensional optimization problem and the Euler equations are in general not sufficient for an optimum.

Now we assert that the Euler equation, together with the transversality condition, is a necessary and sufficient condition for optimization. This conjecture is, to the best of my knowledge, not yet proved (or disproved) for the assumptions that we made on  $U, f$ . As with the social planners problem we assert that for the assumptions we made on  $U, f$  the Euler conditions with the TVC are jointly sufficient and they are both necessary.<sup>21</sup>

The TVC for the household problem state that the value of the capital stock saved for tomorrow must converge to zero as time goes to infinity

$$\lim_{t \rightarrow \infty} p_t k_{t+1} = 0$$

But using the first order condition yields

$$\begin{aligned} \lim_{t \rightarrow \infty} p_t k_{t+1} &= \frac{1}{\mu} \lim_{t \rightarrow \infty} \beta^t U'(c_t) k_{t+1} \\ &= \frac{1}{\mu} \lim_{t \rightarrow \infty} \beta^{t-1} U'(c_{t-1}) k_t \\ &= \frac{1}{\mu} \lim_{t \rightarrow \infty} \beta^{t-1} \beta U'(c_t) (1 - \delta + r_t) k_t \\ &= \frac{1}{\mu} \lim_{t \rightarrow \infty} \beta^t U'(f(k_t) - k_{t+1}) f'(k_t) k_t \end{aligned}$$

where the Lagrange multiplier  $\mu$  on the Arrow-Debreu budget constraint is positive since the budget constraint is strictly binding. Note that this is exactly the same TVC as for the social planners problem. Hence an allocation of capital  $\{k_{t+1}\}_{t=0}^{\infty}$  satisfies the necessary and sufficient conditions for being a Pareto optimal allocations if and only if it satisfies the necessary and sufficient conditions for being part of a competitive equilibrium (always subject to the caveat about the necessity of the TVC in both problems).

This last statement is our version of the fundamental theorems of welfare economics for the particular economy that we consider. The first welfare theorem states that a competitive equilibrium allocation is Pareto efficient (under very general assumptions). The second welfare theorem states that any Pareto efficient allocation can be decentralized as a competitive equilibrium with transfers

---

<sup>21</sup>Note that Stokey et al. in Chapter 2.3, when they discuss the relation between the planning problem and the competitive equilibrium allocation use the finite horizon case, because for this case, under the assumptions made the Euler equations are both necessary and sufficient for both the planning problem and the household optimization problem, so we don't have to worry about the TVC.

(under much more restrictive assumptions), i.e. there exist prices and redistributions of initial endowments such that the prices, together with the Pareto efficient allocation is a competitive equilibrium for the economy with redistributed endowments.

In particular, when dealing with an economy with a representative agent (i.e. when restricting attention to type-identical allocations), whenever the second welfare theorem applies we can solve for Pareto efficient allocations (by solving a social planners problem) and be sure that all Pareto efficient allocations are competitive equilibrium allocations (since there is nobody to redistribute endowments to/from). If, in addition, the first welfare theorem applies we can be sure that we found *all* competitive equilibrium allocations.

Also note an important fact. The first welfare theorem is usually easy to prove, whereas the second welfare theorem is substantially harder, in particular in infinite-dimensional spaces. Suppose we have proved the first welfare theorem and we have established that there exists a unique Pareto efficient allocation (this in general requires representative agent economies and restrictions to type-identical allocations, but in these environments boils down to showing that the social planners problem has a unique solution). Then we have established that, if there is a competitive equilibrium, its allocation has to equal the Pareto efficient allocation. Of course we still need to prove existence of a competitive equilibrium, but this is not surprising given the intimate link between the second welfare theorem and the existence proof.

Back to our economy at hand. Once we have determined the equilibrium sequence of capital stocks  $\{k_{t+1}\}_{t=0}^{\infty}$  we can construct the rest of the competitive equilibrium. In particular equilibrium allocations are given by

$$\begin{aligned} c_t &= f(k_t) - k_{t+1} \\ y_t &= f(k_t) \\ i_t &= y_t - c_t \\ n_t &= 1 \end{aligned}$$

for all  $t \geq 0$ . Finally we can construct factor equilibrium prices as

$$\begin{aligned} r_t &= F_k(k_t, 1) \\ w_t &= F_n(k_t, 1) \end{aligned}$$

Finally, the prices of the final output good can be found as follows. As usual prices are determined only up to a normalization, so let us pick  $p_0 = 1$ . From the Euler equations for the household it then follows that

$$p_{t+1} = \frac{\beta U'(c_{t+1})}{U'(c_t)} p_t = \frac{\beta^{t+1} U'(c_{t+1})}{U'(c_0)}$$

and we have constructed a complete competitive equilibrium, conditional on having found  $\{k_{t+1}\}_{t=0}^{\infty}$ . The welfare theorems tell us that we can solve a social planner problem to do so, and the next sections will tell us how we can do so by using recursive methods.

### 3.3.4 Sequential Markets Equilibrium

[To be completed]

### 3.3.5 Recursive Competitive Equilibrium

[To be completed]

## 4 Mathematical Preliminaries

We want to study functional equations of the form

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

where  $r$  is the period return function (such as the utility function) and  $\Gamma$  is the constraint set. Note that for the neoclassical growth model  $x = k, y = k'$  and  $F(k, k') = U(f(k) - k')$  and  $\Gamma(k) = \{k' \in \mathbf{R} : 0 \leq k' \leq f(k)\}$

In order to so we define the following operator  $T$

$$(Tv)(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta v(y)\}$$

This operator  $T$  takes the function  $v$  as input and spits out a new function  $Tv$ . In this sense  $T$  is like a regular function, but it takes as inputs not scalars  $z \in \mathbf{R}$  or vectors  $\mathbf{z} \in \mathbf{R}^n$ , but functions  $v$  from some subset of possible functions. A solution to the functional equation is then a fixed point of this operator, i.e. a function  $v^*$  such that

$$v^* = Tv^*$$

We want to find out under what conditions the operator  $T$  has a fixed point (existence), under what conditions it is unique and under what conditions we can start from an arbitrary function  $v$  and converge, by applying the operator  $T$  repeatedly, to  $v^*$ . More precisely, by defining the sequence of functions  $\{v_n\}_{n=0}^{\infty}$  recursively by  $v_0 = v$  and  $v_{n+1} = Tv_n$  we want to ask under what conditions  $\lim_{n \rightarrow \infty} v_n = v^*$ .

In order to make these questions (and the answers to them) precise we have to define the domain and range of the operator  $T$  and we have to define what we mean by lim. This requires the discussion of complete metric spaces. In the next subsection I will first define what a metric space is and then what makes a metric space complete.

Then I will state and prove the contraction mapping theorem. This theorem states that an operator  $T$ , defined on a metric space, has a unique fixed point if this operator  $T$  is a contraction (I will obviously first define what a contraction is). Furthermore it assures that from any starting guess  $v$  repeated applications of the operator  $T$  will lead to its unique fixed point.

Finally I will prove a theorem, Blackwell's theorem, that provides sufficient condition for an operator to be a contraction. We will use this theorem to prove

that for the neoclassical growth model the operator  $T$  is a contraction and hence the functional equation of our interest has a unique solution.

## 4.1 Complete Metric Spaces

**Definition 13** A metric space is a set  $S$  and a function  $d : S \times S \rightarrow \mathbf{R}$  such that for all  $x, y, z \in S$

1.  $d(x, y) \geq 0$
2.  $d(x, y) = 0$  if and only if  $x = y$
3.  $d(x, y) = d(y, x)$
4.  $d(x, z) \leq d(x, y) + d(y, z)$

The function  $d$  is called a metric and is used to measure the distance between two elements in  $S$ . The second property is usually referred to as symmetry, the third as triangle inequality (because of its geometric interpretation in  $\mathbf{R}$ ). Examples of metric spaces  $(S, d)$  include<sup>22</sup>

**Example 14**  $S = \mathbf{R}$  with metric  $d(x, y) = |x - y|$

**Example 15**  $S = \mathbf{R}$  with metric  $d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases}$

**Example 16**  $S = l_\infty = \{x = \{x_t\}_{t=0}^\infty \mid x_t \in \mathbf{R}, \text{ all } t \geq 0 \text{ and } \sup_t |x_t| < \infty\}$  with metric  $d(x, y) = \sup_t |x_t - y_t|$

**Example 17** Let  $X \subseteq \mathbf{R}^l$  and  $S = C(X)$  be the set of all continuous and bounded functions  $f : X \rightarrow \mathbf{R}$ . Define the metric  $d : C(X) \times C(X) \rightarrow \mathbf{R}$  as  $d(f, g) = \sup_{x \in X} |f(x) - g(x)|$ . Then  $(S, d)$  is a metric space

A few remarks: the space  $l_\infty$  (with corresponding norm) will be important when we discuss the welfare theorems as naturally consumption allocations for models with infinitely lived consumers are infinite sequences. Why we want to require these sequences to be bounded will become clearer later.

<sup>22</sup>A function  $f : X \rightarrow \mathbf{R}$  is said to be bounded if there exists a constant  $K > 0$  such that  $|f(x)| < K$  for all  $x \in X$ .

Let  $S$  be any subset of  $\mathbf{R}$ . A number  $u \in \mathbf{R}$  is said to be an upper bound for the set  $S$  if  $s \leq u$  for all  $s \in S$ . The supremum of  $S$ ,  $\sup(S)$  is the smallest upper bound of  $S$ .

Every set in  $\mathbf{R}$  that has an upper bound has a supremum (imposed by the completeness axiom). For sets that are unbounded above some people say that the supremum does not exist, others write  $\sup(S) = \infty$ . We will follow the second convention.

Also note that  $\sup(S) = \max(S)$ , whenever the latter exists. What the sup buys us is that it always exists even when the max does not. A simple example

$$S = \left\{ -\frac{1}{n} : n \in \mathbf{N} \right\}$$

For this example  $\sup(S) = 0$  whereas  $\max(S)$  does not exist.

The space  $C(X)$  with norm  $d$  as defined above will be used immediately as we will define the domain of our operator  $T$  to be  $C(X)$ , i.e.  $T$  uses as inputs continuous and bounded functions.

Let us prove that some of the examples are indeed metric spaces. For the first example the result is trivial.

**Claim 18**  $S = \mathbf{R}$  with metric  $d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases}$  is a metric space

**Proof.** We have to show that the function  $d$  satisfies all three properties in the definition. The first three properties are obvious. For the fourth property: if  $x = z$ , the result follows immediately. So suppose  $x \neq z$ . Then  $d(x, z) = 1$ . But then either  $y \neq x$  or  $y \neq z$  (or both), so that  $d(x, y) + d(y, z) \geq 1$  ■

**Claim 19**  $l_\infty$  together with the sup-metric is a metric space

**Proof.** Take arbitrary  $x, y, z \in l_\infty$ . From the basic triangle inequality on  $\mathbf{R}$  we have that  $|x_t - y_t| \leq |x_t| + |y_t|$ . Hence, since  $\sup_t |x_t| < \infty$  and  $\sup_t |y_t| < \infty$ , we have that  $\sup_t |x_t - y_t| < \infty$ . Property 1 is obvious. If  $x = y$  (i.e.  $x_t = y_t$  for all  $t$ ), then  $|x_t - y_t| = 0$  for all  $t$ , hence  $\sup_t |x_t - y_t| = 0$ . Suppose  $x \neq y$ . Then there exists  $T$  such that  $x_T \neq y_T$ , hence  $|x_T - y_T| > 0$ , hence  $\sup_t |x_t - y_t| > 0$

Property 3 is obvious since  $|x_t - y_t| = |y_t - x_t|$ , all  $t$ . Finally for property 4. we note that for all  $t$

$$|x_t - z_t| \leq |x_t - y_t| + |y_t - z_t|$$

Since this is true for all  $t$ , we can apply the sup to both sides to obtain the result (note that the sup on both sides is finite). ■

**Claim 20**  $C(X)$  together with the sup-norm is a metric space

**Proof.** Take arbitrary  $f, g \in C(X)$ .  $f = g$  means that  $f(x) = g(x)$  for all  $x \in X$ . Since  $f, g$  are bounded,  $\sup_{x \in X} |f(x)| < \infty$  and  $\sup_{x \in X} |g(x)| < \infty$ , so  $\sup_{x \in X} |f(x) - g(x)| < \infty$ . Property 1. through 3. are obvious and for property 4. we use the same argument as before, including the fact that  $f, g \in C(X)$  implies that  $\sup_{x \in X} |f(x) - g(x)| < \infty$ . ■

## 4.2 Convergence of Sequences

The next definition will make precise the meaning of statements of the form  $\lim_{n \rightarrow \infty} v_n = v^*$ . For an arbitrary metric space  $(S, d)$  we have the following definition.

**Definition 21** A sequence  $\{x_n\}_{n=0}^\infty$  with  $x_n \in S$  for all  $n$  is said to converge to  $x \in S$ , if for every  $\varepsilon > 0$  there exists a  $N_\varepsilon \in \mathbf{N}$  such that  $d(x_n, x) < \varepsilon$  for all  $n \geq N_\varepsilon$ . In this case we write  $\lim_{n \rightarrow \infty} x_n = x$ .

This definition basically says that a sequence  $\{x_n\}_{n=0}^{\infty}$  converges to a point if we, for every distance  $\varepsilon > 0$  we can find an index  $N_\varepsilon$  so that the sequence of  $x_n$  is not more than  $\varepsilon$  away from  $x$  after the  $N_\varepsilon$  element of the sequence. Also note that, in order to verify that a sequence converges, it is usually necessary to know the  $x$  to which it converges in order to apply the definition directly.

**Example 22** Take  $S = \mathbf{R}$  with  $d(x, y) = |x - y|$ . Define  $\{x_n\}_{n=0}^{\infty}$  by  $x_n = \frac{1}{n}$ . Then  $\lim_{n \rightarrow \infty} x_n = 0$ . This is straightforward to prove, using the definition. Take any  $\varepsilon > 0$ . Then  $d(x_n, 0) = \frac{1}{n}$ . By taking  $N_\varepsilon = \frac{2}{\varepsilon}$  we have that for  $n \geq N_\varepsilon$ ,  $d(x_n, 0) = \frac{1}{n} \leq \frac{1}{N_\varepsilon} = \frac{\varepsilon}{2} < \varepsilon$  (if  $N_\varepsilon = \frac{2}{\varepsilon}$  is not an integer, take the next biggest integer).

For easy examples of sequences it is no problem to guess the limit. Note that the limit of a sequence, if it exists, is always unique (you should prove this for yourself). For not so easy examples this may not work. There is an alternative criterion of convergence, due to Cauchy.<sup>23</sup>

**Definition 23** A sequence  $\{x_n\}_{n=0}^{\infty}$  with  $x_n \in S$  for all  $n$  is said to be a Cauchy sequence if for each  $\varepsilon > 0$  there exists a  $N_\varepsilon \in \mathbf{N}$  such that  $d(x_n, x_m) < \varepsilon$  for all  $n, m \geq N_\varepsilon$ .

Hence a sequence  $\{x_n\}_{n=0}^{\infty}$  is a Cauchy sequence if for every distance  $\varepsilon > 0$  we can find an index  $N_\varepsilon$  so that the elements of the sequence do not differ by more than by  $\varepsilon$ .

**Example 24** Take  $S = \mathbf{R}$  with  $d(x, y) = |x - y|$ . Define  $\{x_n\}_{n=0}^{\infty}$  by  $x_n = \frac{1}{n}$ . This sequence is a Cauchy sequence. Again this is straightforward to prove. Fix  $\varepsilon > 0$  and take any  $n, m \in \mathbf{N}$ . Without loss of generality assume that  $m > n$ . Then  $d(x_n, x_m) = \frac{1}{n} - \frac{1}{m} < \frac{1}{n}$ . Pick  $N_\varepsilon = \frac{2}{\varepsilon}$  and we have that for  $n, m \geq N_\varepsilon$ ,  $d(x_n, 0) < \frac{1}{n} \leq \frac{1}{N_\varepsilon} = \frac{\varepsilon}{2} < \varepsilon$ . Hence the sequence is a Cauchy sequence.

So it turns out that the sequence in the last example both converges and is a Cauchy sequence. This is not an accident. In fact, one can prove the following

**Theorem 25** Suppose that  $(S, d)$  is a metric space and that the sequence  $\{x_n\}_{n=0}^{\infty}$  converges to  $x \in S$ . Then the sequence  $\{x_n\}_{n=0}^{\infty}$  is a Cauchy sequence.

**Proof.** Since  $\{x_n\}_{n=0}^{\infty}$  converges to  $x$ , there exists  $M_{\frac{\varepsilon}{2}}$  such that  $d(x_n, x) < \frac{\varepsilon}{2}$  for all  $n \geq M_{\frac{\varepsilon}{2}}$ . Therefore if  $n, m \geq N_\varepsilon$  we have that  $d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$  (by the definition of convergence and the triangle inequality). But then for any  $\varepsilon > 0$ , pick  $N_\varepsilon = M_{\frac{\varepsilon}{2}}$  and it follows that for all  $n, m \geq N_\varepsilon$  we have  $d(x_n, x_m) < \varepsilon$  ■

**Example 26** Take  $S = \mathbf{R}$  with  $d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases}$ . Define  $\{x_n\}_{n=0}^{\infty}$  by  $x_n = \frac{1}{n}$ . Obviously  $d(x_n, x_m) = 1$  for all  $n, m \in \mathbf{N}$ . Therefore the sequence is

<sup>23</sup>Augustin-Louis Cauchy (1789-1857) was the founder of modern analysis. He wrote about 800 (!) mathematical papers during his scientific life.

not a Cauchy sequence. It then follows from the preceding theorem (by taking the contrapositive) that the sequence cannot converge. This example shows that, whenever discussing a metric space, it is absolutely crucial to specify the metric.

This theorem tells us that every convergent sequence is a Cauchy sequence. The reverse does not always hold, but it is such an important property that when it holds, it is given a particular name.

**Definition 27** A metric space  $(S, d)$  is complete if every Cauchy sequence  $\{x_n\}_{n=0}^{\infty}$  with  $x_n \in S$  for all  $n$  converges to some  $x \in S$ .

Note that the definition requires that the limit  $x$  has to lie within  $S$ . We are interested in complete metric spaces since the Contraction Mapping Theorem deals with operators  $T : S \rightarrow S$ , where  $(S, d)$  is required to be a complete metric space. Also note that there are important examples of complete metric spaces, but other examples where a metric space is not complete (and for which the Contraction Mapping Theorem does not apply).

**Example 28** Let  $S$  be the set of all continuous, strictly decreasing functions on  $[1, 2]$  and let the metric on  $S$  be defined as  $d(f, g) = \sup_{x \in [1, 2]} |f(x) - g(x)|$ . I claim that  $(S, d)$  is not a complete metric space. This can be proved by an example of a sequence of functions  $\{f_n\}_{n=0}^{\infty}$  that is a Cauchy sequence, but does not converge within  $S$ . Define  $f_n : [0, 1] \rightarrow \mathbf{R}$  by  $f_n(x) = \frac{1}{nx}$ . Obviously all  $f_n$  are continuous and strictly decreasing on  $[1, 2]$ , hence  $f_n \in S$  for all  $n$ . Let us first prove that this sequence is a Cauchy sequence. Fix  $\varepsilon > 0$  and take  $N_\varepsilon = \frac{2}{\varepsilon}$ . Suppose that  $m, n \geq N_\varepsilon$  and without loss of generality assume that  $m > n$ . Then

$$\begin{aligned} d(f_n, f_m) &= \sup_{x \in [1, 2]} \left| \frac{1}{nx} - \frac{1}{mx} \right| \\ &= \sup_{x \in [1, 2]} \frac{1}{nx} - \frac{1}{mx} \\ &= \sup_{x \in [1, 2]} \frac{m - n}{mnx} \\ &= \frac{m - n}{mn} = \frac{1 - \frac{n}{m}}{n} \\ &\leq \frac{1}{n} \leq \frac{1}{N_\varepsilon} = \frac{\varepsilon}{2} < \varepsilon \end{aligned}$$

Hence the sequence is a Cauchy sequence. But since for all  $x \in [1, 2]$ ,  $\lim_{n \rightarrow \infty} f_n(x) = 0$ , the sequence converges to the function  $f$ , defined as  $f(x) = 0$ , for all  $x \in [1, 2]$ . But obviously, since  $f$  is not strictly decreasing,  $f \notin S$ . Hence  $(S, d)$  is not a complete metric space. Note that if we choose  $S$  to be the set of all continuous and decreasing (or increasing) functions on  $\mathbf{R}$ , then  $S$ , together with the sup-norm, is a complete metric space.

**Example 29** Let  $S = \mathbf{R}^L$  and  $d(x, y) = \sqrt[L]{\sum_{i=1}^L |x_i - y_i|^L}$ .  $(S, d)$  is a complete metric space. This is easily proved by proving the following three lemmata (which is left to the reader).

1. Every Cauchy sequence  $\{x_n\}_{n=0}^\infty$  in  $\mathbf{R}^L$  is bounded
2. Every bounded sequence  $\{x_n\}_{n=0}^\infty$  in  $\mathbf{R}^L$  has a subsequence  $\{x_{n_i}\}_{i=0}^\infty$  converging to some  $x \in \mathbf{R}^L$  (Bolzano-Weierstrass Theorem)
3. For every Cauchy sequence  $\{x_n\}_{n=0}^\infty$  in  $\mathbf{R}^L$ , if a subsequence  $\{x_{n_i}\}_{i=0}^\infty$  converges to  $x \in \mathbf{R}^L$ , then the entire sequence  $\{x_n\}_{n=0}^\infty$  converges to  $x \in \mathbf{R}^L$ .

**Example 30** This last example is very important for the applications we are interested in. Let  $X \subseteq \mathbf{R}^L$  and  $C(X)$  be the set of all bounded continuous functions  $f : X \rightarrow \mathbf{R}$  with  $d$  being the sup-norm. Then  $(C(X), d)$  is a complete metric space.

**Proof.** (This follows SLP, pp. 48) We already proved that  $(C(X), d)$  is a metric space. Now we want to prove that this space is complete. Let  $\{f_n\}_{n=0}^\infty$  be an arbitrary sequence of functions in  $C(X)$  which is Cauchy. We need to establish the existence of a function  $f \in C(X)$  such that for all  $\varepsilon > 0$  there exists  $N_\varepsilon$  satisfying  $\sup_{x \in X} |f_n(x) - f(x)| < \varepsilon$  for all  $n \geq N_\varepsilon$ .

We will proceed in three steps: a) find a candidate for  $f$ , b) establish that the sequence  $\{f_n\}_{n=0}^\infty$  converges to  $f$  in the sup-norm and c) show that  $f \in C(X)$ .

1. Since  $\{f_n\}_{n=0}^\infty$  is Cauchy, for each  $\varepsilon > 0$  there exists  $M_\varepsilon$  such that  $\sup_{x \in X} |f_n(x) - f_m(x)| < \varepsilon$  for all  $n, m \geq M_\varepsilon$ . Now fix a particular  $x \in X$ . Then  $\{f_n(x)\}_{n=0}^\infty$  is just a sequence of numbers. Now

$$|f_n(x) - f_m(x)| \leq \sup_{y \in X} |f_n(y) - f_m(y)| < \varepsilon$$

Hence the sequence of numbers  $\{f_n(x)\}_{n=0}^\infty$  is a Cauchy sequence in  $\mathbf{R}$ . Since  $\mathbf{R}$  is a complete metric space,  $\{f_n(x)\}_{n=0}^\infty$  converges to some number, call it  $f(x)$ . By repeating this argument for all  $x \in X$  we derive our candidate function  $f$ ; it is the pointwise limit of the sequence of functions  $\{f_n\}_{n=0}^\infty$ .

2. Now we want to show that  $\{f_n\}_{n=0}^\infty$  converges to  $f$  as constructed above. Hence we want to argue that  $d(f_n, f)$  goes to zero as  $n$  goes to infinity. Fix  $\varepsilon > 0$ . Since  $\{f_n\}_{n=0}^\infty$  is Cauchy, it follows that there exists  $N_\varepsilon$  such that  $d(f_n, f_m) < \varepsilon$  for all  $n, m \geq N_\varepsilon$ . Now fix  $x \in X$ . For any  $m \geq n \geq N_\varepsilon$  we have (remember that the norm is the sup-norm)

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| \\ &\leq d(f_n, f_m) + |f_m(x) - f(x)| \\ &\leq \frac{\varepsilon}{2} + |f_m(x) - f(x)| \end{aligned}$$

But since  $\{f_n\}_{n=0}^\infty$  converges to  $f$  pointwise, we have that  $|f_m(x) - f(x)| < \frac{\varepsilon}{2}$  for all  $m \geq N_\varepsilon(x)$ , where  $N_\varepsilon(x)$  is a number that may (and in general does) depend on  $x$ . But then, since  $x \in X$  was arbitrary,  $|f_n(x) - f(x)| < \varepsilon$

for all  $n \geq N_\varepsilon$  (the key is that this  $N_\varepsilon$  does *not* depend on  $x$ ). Therefore  $\sup_{x \in X} |f_n(x) - f(x)| = d(f_n, f) \leq \varepsilon$  and hence the sequence  $\{f_n\}_{n=0}^\infty$  converges to  $f$ .

3. Finally we want to show that  $f \in C(X)$ , i.e. that  $f$  is bounded and continuous. Since  $\{f_n\}_{n=0}^\infty$  lies in  $C(X)$ , all  $f_n$  are bounded, i.e. there is a sequence of numbers  $\{K_n\}_{n=0}^\infty$  such that  $\sup_{x \in X} |f_n(x)| \leq K_n$ . But by the triangle inequality, for arbitrary  $n$

$$\begin{aligned} \sup_{x \in X} |f(x)| &= \sup_{x \in X} |f(x) - f_n(x) + f_n(x)| \\ &\leq \sup_{x \in X} |f(x) - f_n(x)| + \sup_{x \in X} |f_n(x)| \\ &\leq \sup_{x \in X} |f(x) - f_n(x)| + K_n \end{aligned}$$

But since  $\{f_n\}_{n=0}^\infty$  converges to  $f$ , there exists  $N_\varepsilon$  such that  $\sup_{x \in X} |f(x) - f_n(x)| < \varepsilon$  for all  $n \geq N_\varepsilon$ . Fix an  $\varepsilon$  and take  $K = K_{N_\varepsilon} + 2\varepsilon$ . It is obvious that  $\sup_{x \in X} |f(x)| \leq K$ . Hence  $f$  is bounded. Finally we prove continuity of  $f$ . Let the Euclidean metric on  $\mathbf{R}^L$  be denoted by  $\|x - y\| = \sqrt[2]{\sum_{l=1}^L |x_l - y_l|^2}$ . We need to show that for every  $\varepsilon > 0$  and every  $x \in X$  there exists a  $\delta(\varepsilon, x) > 0$  such that if  $\|x - y\| < \delta(\varepsilon, x)$  then  $|f(x) - f(y)| < \varepsilon$ , for all  $x, y \in X$ . Fix  $\varepsilon$  and  $x$ . Pick a  $k$  large enough so that  $d(f_k, f) < \frac{\varepsilon}{3}$  (which is possible as  $\{f_n\}_{n=0}^\infty$  converges to  $f$ ). Choose  $\delta(\varepsilon, x) > 0$  such that  $\|x - y\| < \delta(\varepsilon, x)$  implies  $|f_k(x) - f_k(y)| < \frac{\varepsilon}{3}$ . Since all  $f_n \in C(X)$ ,  $f_k$  is continuous and hence such a  $\delta(\varepsilon, x) > 0$  exists. Now

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\ &\leq d(f, f_k) + |f_k(x) - f_k(y)| + d(f_k, f) \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

■

### 4.3 The Contraction Mapping Theorem

Now we are ready to state the theorem that will give us the existence and uniqueness of a fixed point of the operator  $T$ , i.e. existence and uniqueness of a function  $v^*$  satisfying  $v^* = Tv^*$ . Let  $(S, d)$  be a metric space. Just to clarify, an operator  $T$  (or a mapping) is just a function that maps elements of  $S$  into some other space. The operator that we are interested in maps functions into functions, but the results in this section apply to any metric space. We start with a definition of what a contraction mapping is.

**Definition 31** *Let  $(S, d)$  be a metric space and  $T : S \rightarrow S$  be a function mapping  $S$  into itself. The function  $T$  is a contraction mapping if there exists a number  $\beta \in (0, 1)$  satisfying*

$$d(Tx, Ty) \leq \beta d(x, y) \text{ for all } x, y \in S$$

The number  $\beta$  is called the modulus of the contraction mapping

A geometric example of a contraction mapping for  $S = [0, 1]$ ,  $d(x, y) = |x - y|$  is contained in SLP, p. 50. Note that a function that is a contraction mapping is automatically a continuous function, as the next lemma shows

**Lemma 32** *Let  $(S, d)$  be a metric space and  $T : S \rightarrow S$  be a function mapping  $S$  into itself. If  $T$  is a contraction mapping, then  $T$  is continuous.*

**Proof.** Remember from the definition of continuity we have to show that for all  $s_0 \in S$  and all  $\varepsilon > 0$  there exists a  $\delta(\varepsilon, s_0)$  such that whenever  $s \in S$ ,  $d(s, s_0) < \delta(\varepsilon, s_0)$ , then  $d(Ts, Ts_0) < \varepsilon$ . Fix an arbitrary  $s_0 \in S$  and  $\varepsilon > 0$  and pick  $\delta(\varepsilon, s_0) = \varepsilon$ . Then

$$d(Ts, Ts_0) \leq \beta d(s, s_0) \leq \beta \delta(\varepsilon, s_0) = \beta \varepsilon < \varepsilon$$

■

We now can state and prove the contraction mapping theorem. Let by  $v_n = T^n v_0 \in S$  denote the element in  $S$  that is obtained by applying the operator  $T$   $n$ -times to  $v_0$ , i.e. the  $n$ -th element in the sequence starting with an arbitrary  $v_0$  and defined recursively by  $v_n = T v_{n-1} = T(T v_{n-2}) = \dots = T^n v_0$ . Then we have

**Theorem 33** *Let  $(S, d)$  be a complete metric space and suppose that  $T : S \rightarrow S$  is a contraction mapping with modulus  $\beta$ . Then a) the operator  $T$  has exactly one fixed point  $v^* \in S$  and b) for any  $v_0 \in S$ , and any  $n \in \mathbf{N}$  we have*

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

A few remarks before the proof. Part a) of the theorem tells us that there is a  $v^* \in S$  satisfying  $v^* = T v^*$  and that there is only one such  $v^* \in S$ . Part b) asserts that from any starting guess  $v_0$ , the sequence  $\{v_n\}_{n=0}^\infty$  as defined recursively above converges to  $v^*$  at a geometric rate of  $\beta$ . This last part is important for computational purposes as it makes sure that we, by repeatedly applying  $T$  to any (as crazy as can be) initial guess  $v_0 \in S$ , will eventually converge to the unique fixed point and it gives us a lower bound on the speed of convergence. But now to the proof.

**Proof.** First we prove part a) Start with an arbitrary  $v_0$ . As our candidate for a fixed point we take  $v^* = \lim_{n \rightarrow \infty} v_n$ . We first have to establish that the sequence  $\{v_n\}_{n=0}^\infty$  in fact converges to a function  $v^*$ . We then have to show that this  $v^*$  satisfies  $v^* = T v^*$  and we then have to show that there is no other  $\hat{v}$  that also satisfies  $\hat{v} = T \hat{v}$

Since by assumption  $T$  is a contraction

$$\begin{aligned} d(v_{n+1}, v_n) &= d(T v_n, T v_{n-1}) \leq \beta d(v_n, v_{n-1}) \\ &= \beta d(T v_{n-1}, T v_{n-2}) \leq \beta^2 d(v_{n-1}, v_{n-2}) \\ &= \dots = \beta^n d(v_1, v_0) \end{aligned}$$

where we used the way the sequence  $\{v_n\}_{n=0}^\infty$  was constructed, i.e. the fact that  $v_{n+1} = Tv_n$ . For any  $m > n$  it then follows from the triangle inequality that

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

By making  $n$  large we can make  $d(v_m, v_n)$  as small as we want. Hence the sequence  $\{v_n\}_{n=0}^\infty$  is a Cauchy sequence. Since  $(S, d)$  is a complete metric space, the sequence converges in  $S$  and therefore  $v^* = \lim_{n \rightarrow \infty} v_n$  is well-defined.

Now we establish that  $v^*$  is a fixed point of  $T$ , i.e. we need to show that  $Tv^* = v^*$ . But

$$Tv^* = T\left(\lim_{n \rightarrow \infty} v_n\right) = \lim_{n \rightarrow \infty} T(v_n) = \lim_{n \rightarrow \infty} v_{n+1} = v^*$$

Note that the fact that  $T(\lim_{n \rightarrow \infty} v_n) = \lim_{n \rightarrow \infty} T(v_n)$  follows from the continuity of  $T$ .<sup>24</sup>

Now we want to prove that the fixed point of  $T$  is unique. Suppose there exists another  $\hat{v} \in S$  such that  $\hat{v} = T\hat{v}$  and  $\hat{v} \neq v^*$ . Then there exists  $c > 0$  such that  $d(\hat{v}, v^*) = a$ . But

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

a contradiction. Here the second equality follows from the fact that we assumed that both  $\hat{v}, v^*$  are fixed points of  $T$  and the inequality follows from the fact that  $T$  is a contraction.

We prove part b) by induction. For  $n = 0$  (using the convention that  $T^0 v = v$ ) the claim automatically holds. Now suppose that

$$d(T^k v_0, v^*) \leq \beta^k d(v_0, v^*)$$

We want to prove that

$$d(T^{k+1} v_0, v^*) \leq \beta^{k+1} d(v_0, v^*)$$

But

$$d(T^{k+1} v_0, v^*) = d(T(T^k v_0), Tv^*) \leq \beta d(T^k v_0, v^*) \leq \beta^{k+1} d(v_0, v^*)$$

where the first inequality follows from the fact that  $T$  is a contraction and the second follows from the induction hypothesis. ■

---

<sup>24</sup>Almost by definition. Since  $T$  is continuous for every  $\varepsilon > 0$  there exists a  $\delta(\varepsilon)$  such that  $d(v_n - v^*) < \delta(\varepsilon)$  implies  $d(T(v_n) - T(v^*)) < \varepsilon$ . Hence the sequence  $\{T(v_n)\}_{n=0}^\infty$  converges and  $\lim_{n \rightarrow \infty} T(v_n)$  is well-defined. We showed that  $\lim_{n \rightarrow \infty} v_n = v^*$ . Hence both  $\lim_{n \rightarrow \infty} T(v_n)$  and  $\lim_{n \rightarrow \infty} v_n$  are well-defined. Then obviously  $\lim_{n \rightarrow \infty} T(v_n) = T(v^*) = T(\lim_{n \rightarrow \infty} v_n)$ .

This theorem is extremely useful in order to establish that our functional equation of interest has a unique fixed point. It is, however, not very operational as long as we don't know how to determine whether a given operator is a contraction mapping. There is some good news, however. Blackwell, in 1965 provided sufficient conditions for an operator to be a contraction mapping. It turns out that these conditions can be easily checked in a lot of applications. Since they are only sufficient however, failure of these conditions does not imply that the operator is not a contraction. In these cases we just have to look somewhere else. Here is Blackwell's theorem.

**Theorem 34** *Let  $X \subseteq \mathbf{R}^L$  and  $B(X)$  be the space of bounded functions  $f : X \rightarrow \mathbf{R}$  with the  $d$  being the sup-norm. Let  $T : B(X) \rightarrow B(X)$  be an operator satisfying*

1. *Monotonicity: If  $f, g \in B(X)$  are such that  $f(x) \leq g(x)$  for all  $x \in X$ , then  $(Tf)(x) \leq (Tg)(x)$  for all  $x \in X$ .*
2. *Discounting: Let the function  $f + a$ , for  $f \in B(X)$  and  $a \in \mathbf{R}_+$  be defined by  $(f + a)(x) = f(x) + a$  (i.e. for all  $x$  the number  $a$  is added to  $f(x)$ ). There exists  $\beta \in (0, 1)$  such that for all  $f \in B(X)$ ,  $a \geq 0$  and all  $x \in X$*

$$[T(f + a)](x) \leq [Tf](x) + \beta a$$

*If these two conditions are satisfied, then the operator  $T$  is a contraction with modulus  $\beta$ .*

**Proof.** In terms of notation, if  $f, g \in B(X)$  are such that  $f(x) \leq g(x)$  for all  $x \in X$ , then we write  $f \leq g$ . We want to show that if the operator  $T$  satisfies conditions 1. and 2. then there exists  $\beta \in (0, 1)$  such that for all  $f, g \in B(X)$  we have that  $d(Tf, Tg) \leq \beta d(f, g)$ .

Fix  $x \in X$ . Then  $f(x) - g(x) \leq \sup_{y \in X} |f(y) - g(y)|$ . But this is true for all  $x \in X$ . So using our notation we have that  $f \leq g + d(f, g)$  (which means that for any value of  $x \in X$ , adding the constant  $d(f, g)$  to  $g(x)$  gives something bigger than  $f(x)$ ).

But from  $f \leq g + d(f, g)$  it follows by monotonicity that

$$\begin{aligned} Tf &\leq T[g + d(f, g)] \\ &\leq Tg + \beta d(f, g) \end{aligned}$$

where the last inequality comes from discounting. Hence we have

$$Tf - Tg \leq \beta d(f, g)$$

Switching the roles of  $f$  and  $g$  around we get

$$-(Tf - Tg) \leq \beta d(g, f) = \beta d(f, g)$$

(by symmetry of the metric). Combining yields

$$\begin{aligned}(Tf)(x) - (Tg)(x) &\leq \beta d(f, g) \text{ for all } x \in X \\ (Tg)(x) - (Tf)(x) &\leq \beta d(f, g) \text{ for all } x \in X\end{aligned}$$

Therefore

$$\sup_{x \in X} |(Tf)(x) - (Tg)(x)| = d(Tf, Tg) \leq \beta d(f, g)$$

and  $T$  is a contraction mapping with modulus  $\beta$ . ■

Note that do not require the functions in  $B(X)$  to be continuous. It is straightforward to prove that  $(B(X), d)$  is a complete metric space once we proved that  $(B(X), d)$  is a complete metric space. Also note that we could restrict ourselves to continuous and bounded functions and Blackwell's theorem obviously applies. Note however that Blackwells theorem requires the metric space to be a space of functions, so we lose generality as compared to the Contraction mapping theorem (which is valid for any complete metric space). But for our purposes it is key that, once Blackwell's conditions are verified we can invoke the CMT to argue that our functional equation of interest has a unique solution that can be obtained by repeated iterations on the operator  $T$ .

We can state an alternative version of Blackwell's theorem

**Theorem 35** *Let  $X \subseteq \mathbf{R}^L$  and  $B(X)$  be the space of bounded functions  $f : X \rightarrow \mathbf{R}$  with the  $d$  being the sup-norm. Let  $T : B(X) \rightarrow B(X)$  be an operator satisfying*

1. *Monotonicity: If  $f, g \in B(X)$  are such that  $f(x) \leq g(x)$  for all  $x \in X$ , then  $(Tf)(x) \geq (Tg)(x)$  for all  $x \in X$ .*
2. *Discounting: Let the function  $f + a$ , for  $f \in B(X)$  and  $a \in \mathbf{R}_+$  be defined by  $(f + a)(x) = f(x) + a$  (i.e. for all  $x$  the number  $a$  is added to  $f(x)$ ). There exists  $\beta \in (0, 1)$  such that for all  $f \in B(X), a \geq 0$  and all  $x \in X$*

$$[T(f - a)](x) \leq [Tf](x) + \beta a$$

*If these two conditions are satisfied, then the operator  $T$  is a contraction with modulus  $\beta$ .*

The proof is identical to the first theorem and hence omitted.

As an application of the mathematical structure we developed let us look back at the neoclassical growth model. The operator  $T$  corresponding to our functional equation was

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

Define as our metric space  $(B(0, \infty), d)$  the space of bounded functions on  $(0, \infty)$  with  $d$  being the sup-norm. We want to argue that this operator has a unique fixed point and we want to apply Blackwell's theorem and the CMT. So let us verify that all the hypotheses for Blackwell's theorem are satisfied.

1. First we have to verify that the operator  $T$  maps  $B(0, \infty)$  into itself (this is very often forgotten). So if we take  $v$  to be bounded, since we assumed that  $U$  is bounded, then  $Tv$  is bounded. Note that you may be in big trouble here if  $U$  is not bounded.<sup>25</sup>
2. How about monotonicity. It is obvious that this is satisfied. Suppose  $v \leq w$ . Let by  $g_v(k)$  denote an optimal policy (need not be unique) corresponding to  $v$ . Then for all  $k \in (0, \infty)$

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

Even by applying the policy  $g_v(k)$  (which need not be optimal for the situation in which the value function is  $w$ ) gives higher  $Tw(k)$  than  $Tv(k)$ . Choosing the policy for  $w$  optimally does only improve the value  $(Tv)(k)$ .

3. Discounting. This also straightforward

$$\begin{aligned}
T(v + a)(k) &= \max_{0 \leq k' \leq f(k)} \{U(f(k) - k') + \beta(v(k') + a)\} \\
&= \max_{0 \leq k' \leq f(k)} \{U(f(k) - k') + \beta v(k')\} + \beta a \\
&= Tv(k) + \beta a
\end{aligned}$$

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

One can also prove some theoretical properties of the Howard improvement algorithm using the Contraction Mapping Theorem and Blackwell's conditions. Even though we could state the results in much generality, we will confine our discussion to the neoclassical growth model. Remember that the Howard improvement algorithm iterates on feasible policies [To be completed]

---

<sup>25</sup>Somewhat surprisingly, in many applications the problem is that  $u$  is not bounded below; unboundedness from above is sometimes easy to deal with.

We made the assumption that  $f \in C^2$ ,  $f' > 0$ ,  $f'' < 0$ ,  $\lim_{k \rightarrow 0} f'(k) = \infty$  and  $\lim_{k \rightarrow \infty} f'(k) = 1 - \delta$ . Hence there exists a unique  $\hat{k}$  such that  $f(\hat{k}) = \hat{k}$ . Hence for all  $k_t > \hat{k}$  we have  $k_{t+1} \leq f(k_t) < k_t$ . Therefore we can effectively restrict ourselves to capital stocks in the set  $[0, \max(k_0, \hat{k})]$ . Hence, even if  $u$  is not bounded above we have that for all feasible paths policies  $u(f(k) - k') \leq u(f(\max(k_0, \hat{k})) - k') < \infty$ , and hence by sticking a function  $v$  into the operator that is bounded above, we get a  $Tv$  that is bounded above. Lack of boundedness from below is a much harder problem in general.

## 4.4 The Theorem of the Maximum

An important theorem is the theorem of the maximum. It will help us to establish that, if we stick a continuous function  $f$  into our operator  $T$ , the resulting function  $Tf$  will also be continuous and the optimal policy function will be continuous in an appropriate sense.

We are interested in problems of the form

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

The function  $h$  gives the value of the maximization problem, conditional on the state  $x$ . We define

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

Hence  $G$  is the set of all choices  $y$  that attain the maximum of  $f$ , given the state  $x$ , i.e.  $G(x)$  is the set of argmax'es. Note that  $G(x)$  need not be single-valued.

In the example that we study the function  $f$  will consist of the sum of the current return function  $r$  and the continuation value  $v$  and the constraint set describes the resource constraint. The theorem of the maximum is also widely used in microeconomics. There, most frequently  $x$  consists of prices and income,  $f$  is the (static) utility function, the function  $h$  is the indirect utility function,  $\Gamma$  is the budget set and  $G$  is the set of consumption bundles that maximize utility at  $x = (p, m)$ .

Before stating the theorem we need a few definitions. Let  $X, Y$  be arbitrary sets (in what follows we will be mostly concerned with the situations in which  $X$  and  $Y$  are subsets of Euclidean spaces). A correspondence  $\Gamma : X \Rightarrow Y$  maps each element  $x \in X$  into a subset  $\Gamma(x)$  of  $Y$ . Hence the image of the point  $x$  under  $\Gamma$  may consist of more than one point (in contrast to a function, in which the image of  $x$  always consists of a singleton).

**Definition 36** *A compact-valued correspondence  $\Gamma : X \Rightarrow Y$  is upper-hemicontinuous at a point  $x$  if  $\Gamma(x) \neq \emptyset$  and if for all sequences  $\{x_n\}$  in  $X$  converging to some  $x \in X$  and all sequences  $\{y_n\}$  in  $Y$  such that  $y_n \in \Gamma(x_n)$  for all  $n$ , there exists a convergent subsequence of  $\{y_n\}$  that converges to some  $y \in \Gamma(x)$ . A correspondence is upper-hemicontinuous if it is upper-hemicontinuous at all  $x \in X$ .*

A few remarks: by talking about convergence we have implicitly assumed that  $X$  and  $Y$  (together with corresponding metrics) are metric spaces. Also, a correspondence is compact-valued, if for all  $x \in X, \Gamma(x)$  is a compact set. Also this definition requires  $\Gamma$  to be compact-valued. With this additional requirement the definition of upper hemicontinuity actually corresponds to the definition of a correspondence having a closed graph. See, e.g. Mas-Colell et al. p. 949-950 for details.

**Definition 37** *A correspondence  $\Gamma : X \Rightarrow Y$  is lower-hemicontinuous at a point  $x$  if  $\Gamma(x) \neq \emptyset$  and if for every  $y \in \Gamma(x)$  and every sequence  $\{x_n\}$  in  $X$*

converging to  $x \in X$  there exists  $N \geq 1$  and a sequence  $\{y_n\}$  in  $Y$  converging to  $y$  such that  $y_n \in \Gamma(x_n)$  for all  $n \geq N$ . A correspondence is lower-hemicontinuous if it is lower-hemicontinuous at all  $x \in X$ .

**Definition 38** A correspondence  $\Gamma : X \rightrightarrows Y$  is continuous if it is both upper-hemicontinuous and lower-hemicontinuous.

Note that a single-valued correspondence (i.e. a function) that is upper-hemicontinuous is continuous. Now we can state the theorem of the maximum.

**Theorem 39** Let  $X \subseteq \mathbf{R}^L$  and  $Y \subseteq \mathbf{R}^M$ , let  $f : X \times Y \rightarrow \mathbf{R}$  be a continuous function, and let  $\Gamma : X \rightrightarrows Y$  be a compact-valued and continuous correspondence. Then  $h : X \rightarrow \mathbf{R}$  is continuous and  $G : X \rightarrow Y$  is nonempty, compact-valued and upper-hemicontinuous.

The proof is somewhat tedious and omitted here (you probably have done it in micro anyway).