The Neoclassical Growth Model (Ramsey-Cass-Koopmans)

Luiz Brotherhood
Universitat de Barcelona

Fall 2024

Contents

1 Introduction

[To be written]

Even though the most common version of the Ramsey model is in continuous time, section 2 present an extensive discussion of the model in discrete time. Why?

1. First and foremost, interpreting discrete-time equations is easier and more intuitive. Thus, the discrete-time model is interesting for students seeing the Ramsey environment for the first time.

2. Some important concepts are easier to understand in discrete time, such as present value calculations, the no-Ponzi condition, and the transversality condition.

This material might be of interest to teachers and intermediate/advanced economists because it includes some content I haven't seen in other sources, such as:

1. Heuristic argumentation on the existence and unicity of the equilibrium path using the phase diagram in discrete and continuous time (sections 2.3 and 3.3).

2. From discrete to continuous time: the "Delta-limit approach" and the no-Ponzi condition (section E.3).

3. From discrete to continuous time: the "differential equation" approach (section F).

Acknowledgement to other materials:

- "Economic Growth", by Barro and Sala-i-Martin.

- "Introduction to Modern Economic Growth", by Daron Acemoglu.

- Slides by Pontus Rendahl: `https://benjaminmoll.com/wp-content/uploads/2019/07/Pontus_Lecture1.pdf`.

- Notes by Eric Roca: `https://eric-roca.github.io/courses/mathematical_appendix/transversality_condition/` and `https://eric-roca.github.io/courses/mathematical_appendix/elasticity_of_substitution/`.

2 The model in discrete time

Time is discrete and runs forever, $t = 0, 1, \ldots$. The economy is populated by two types of agents: households and firms. Households are portrayed by an infinitely-lived representative household that enjoys utility from a consumption good. Households supply labor and can save or borrow resources across time through assets. Firms are portrayed by a representative firm that demands labor and capital to maximize profits through selling a final good. Assets are converted into capital through a one-to-one relationship (one unit of asset can be transformed into one unit of capital, and vice-versa). Each unit of the final good either becomes consumption good or capital/assets. Households and firms take as given the prices of labor (wage rate) and capital (interest rate or capital rental rate). In general equilibrium, the wage and interest rates are determined so that labor and capital markets clear. That is, the aggregate demand of labor (capital) is equal to the aggregate supply of labor (capital) in every time period. All agents have perfect foresight and know the full path of wages and interest rates that prevail in equilibrium.

There is population growth and technological growth. Population evolves exogenously according to the initial condition and law of motion

$$L_0 = 1, \quad L_{t+1} = (1+n)L_t \quad \forall t = 0, 1, \ldots, \tag{1}$$

where the population growth rate $n \geq 0$ is a parameter and the initial population size is normalized to one.

In some parts of the model, it will be useful to work with per-capita variables. Per-capita variables are denoted through lower-case variables and are obtained by dividing a given variable by the population size. For example, GDP per capita (quantity of final goods) in t is given by the total GDP, Y_t, divided by the population, L_t:

$$y_t \equiv \frac{Y_t}{L_t}.$$

The final good is produced with capital and labor through a Cobb-Douglas production function

$$Y_t = F(K_t, T_t L_t) = K_t^{\alpha}(T_t L_t)^{1-\alpha}, \tag{2}$$

where T_t is labor productivity ("technology") and $\alpha \in (0, 1)$ is a parameter measuring the importance of capital (relative to labor) for the final good production. Labor productivity evolves exogenously according to the initial condition and law of motion

$$T_0 = 1, \quad T_{t+1} = (1+x)T_t \quad \forall t = 0, 1, \ldots, \tag{3}$$

where the technological growth rate $x \geq 0$ is a parameter and the initial technology level is normalized to one.

(2) says that the total production input generated from workers is given by the number of workers, L_t, times the productivity of each worker, T_t. The term $L_t T_t$ is called "effective units of labor" or "effective labor". In some parts of the model, it will be useful to work with variables per effective units of labor. These variables are denoted by lower-case letters with the "hat" symbol, ˆ. For example, GDP per effective units of labor is

$$\hat{y}_t \equiv \frac{Y_t}{L_t T_t} = \frac{y_t}{T_t}.$$

2.1 Households

2.1.1 The household's problem

The representative household takes as given the paths of wages, interest rates, and population size, $\{w_t, r_t, L_t\}_{t=0}^{\infty}$,[1] and solves the following optimization problem in the first time period, $t = 0$:

$$\max_{\{C_t, A_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t L_t u\left(\frac{C_t}{L_t}\right) \tag{4}$$

subject to

$$C_t + A_{t+1} = L_t w_t + (1+r_t)A_t \quad \forall t = 0, 1, \ldots, \tag{5}$$

$$A_0 > 0 \text{ given,} \tag{6}$$

$$C_t \geq 0, \;\; A_{t+1} \in \mathbb{R} \quad \forall t = 0, 1, \ldots, \tag{7}$$

$$\lim_{t \to \infty} A_{t+1} \prod_{\tau=1}^{t} (1+r_\tau)^{-1} \geq 0. \tag{8}$$

Next, we describe each of the mathematical objects above in detail.

[1] The notation $\{x_t\}_{t=0}^{\infty}$ means the set of the variable x_t for time periods starting at 0 and going forward in time towards infinity. That is, $\{x_t\}_{t=0}^{\infty} = \{x_0, x_1, x_2, \ldots\}$. Therefore, $\{w_t, r_t, L_t\}_{t=0}^{\infty} = \{w_0, w_1, w_2, \ldots, r_0, r_1, r_2, \ldots, L_0, L_1, L_2, \ldots\}$.

Choice variables and the objective function There are two points worth noting in expression (4). First, the choice variables in the problem are $\{C_t, A_{t+1}\}_{t=0}^{\infty}$: in each time period t, households choose how much to consume, C_t, and how many assets to have in the beginning of the next time period, A_{t+1}. You can think of the asset stock as the balance of a checking account (current account). The household savings are then the change in the asset stock, $A_{t+1} - A_t$.

The second point of (4) is that the function the household maximizes (also known as the "objective function" of the maximization problem) is its discounted lifetime utility, $\sum_{t=0}^{\infty} \beta^t L_t u(C_t/L_t)$. $u(\cdot)$ is a utility function representing the happiness each household member enjoys from a given level of consumption within a time period. Each household member consumes the same amount, so per-capita consumption is C_t/L_t.

The utility function has the following properties:[2]

1. Increasing with respect to consumption: $u'(c) \equiv \partial u(c)/\partial c > 0$ for all $c > 0$. A household member feels happier if it consumes more.

2. Concave with respect to consumption, $u''(c) \equiv \partial u'(c)/\partial c < 0$ for all $c > 0$. As consumption grows, each additional unit of consumption generates smaller gains in happiness.

3. The marginal utility goes to infinity when c tends to zero, $\lim_{c\to 0} u'(c) = \infty$. The gain in utility from increasing consumption is extremely high if consumption is very low. As c goes to zero, this gain in utility from increasing consumption goes to infinity. This condition ensures that the household will not choose to have zero consumption in a given time period.

4. The marginal utility goes to zero when c tends to infinity, $\lim_{c\to\infty} u'(c) = 0$. A household member never gets fully satisfied with a finite amount of consumption. In the limit case when c tends to infinity this household member gets satisfied.

The total utility within the household is the utility of a given member, $u(C_t/L_t)$, times the number of household members, L_t.

[2] As an example, a commonly used utility function is the the constant intertemporal elasticity of substitution (CIES) utility function, $u(c) = (c^{1-\theta} - 1)/(1-\theta)$, where $\theta > 0$ is a parameter measuring the concavity of the utility function. In the next pages, the CIES utility function will be introduced more carefully.

$\beta \in (0, 1)$ is a time discount parameter representing the fact that households are impatient and place less value on future utilities compared to present utility. Rewrite the household's discounted lifetime utility as

$$\sum_{t=0}^{\infty} \beta^t L_t u\left(\frac{C_t}{L_t}\right) = L_0 u\left(\frac{C_0}{L_0}\right) + \beta L_1 u\left(\frac{C_1}{L_1}\right) + \beta^2 L_2 u\left(\frac{C_2}{L_2}\right) + \cdots + \beta^{10} L_{10} u\left(\frac{C_{10}}{L_{10}}\right) + \ldots . \tag{9}$$

This expression represents the household's valuation of a given consumption stream, $\{C_t\}_{t=0}^{\infty}$, measured in $t = 0$. At $t = 0$, the household values the consumption it will have in the current time period, C_0, and in each future time period, C_1, C_2, etc. Let's pick a time period: say, time period 10. $L_{10}u(C_{10}/L_{10})$ is the total utility that the household experiences in time period $t = 10$ when consuming the amount C_{10}. Although the household effectively enjoys the utility level $L_{10}u(C_{10}/L_{10})$ only when the time period 10 arrives, the household knows at time 0 it will enjoy $L_{10}u(C_{10}/L_{10})$ in $t = 10$. Therefore, in $t = 0$, the household "feels" the utility $L_{10}u(C_{10}/L_{10})$ after discounting for the fact that this utility will only be felt 10 periods ahead. This is what the term $\beta^{10}L_{10}u(C_{10}/L_{10})$ represents in (9). Since β is smaller than one, β^t falls as t grows: utility levels further in the future are given less weight. Additionally, a lower β means that future utilities are given less weight in the present.

The budget constraint Equation (5) represents the budget constraint the household faces in each time period. The right-hand side of the equation denotes the income available to be used. This income is composed of two elements: wages and assets. In a given period, each household member supplies labor to the firm and receives a wage w_t. Thus, the total labor earnings of the household are $L_t w_t$. The second component of the income in t are the assets that the household has accumulated up to that point. A_t is the amount of assets that the household chose to take from $t-1$ to t. Each unit of assets yields interest. Therefore, in the beginning of t, the asset income of the household is composed by the stock of assets, A_t, plus the return assets yielded from $t-1$ to t, $r_t A_t$. Thus, the total income coming from assets is $(1 + r_t)A_t$. The term $r_t A_t$ is called asset earnings.

The left-hand side of (15) shows the two ways the household can use its income: each euro can be used either for buying the consumption good, C_t, or for accumulating assets to the next time period, A_{t+1}.

At the beginning of a given time period t, the variables in the right-hand side of (15) are taken as given by the household. First, we already saw that the wage rate, w_t, interest rate, r_t, and population size, L_t, are taken as given by the household. Second, the variable A_t is chosen in $t-1$. Therefore, at the beginning of t, A_t was already chosen by the household and cannot be changed in t. On the other hand, the variables in the right-hand side of (15) are the choice variables that the household determines in time period t.

Bounds and Ponzi schemes Equation (6) says that the household takes as given its initial level of assets. That is, the household doesn't choose A_0.

Equation (7) says that consumption cannot be negative. However, the household can choose any real value of assets, including negative values. A negative level of assets is interpreted as borrowing (debt). Note that, if the household chooses a negative value of A_{t+1} in t, then A_{t+1} appears as a negative value in the income of the household in the next time period, $t+1$, multiplied by $1+r_{t+1}$. This shows that, if the household borrows money in a given time period, the household has to pay back this borrowing with interest rates in the next time period.

You might be wondering "Ok, the household needs to pay back the amount it borrowed plus interest rates. But then can't the household borrow even more money to pay back the original borrowing?" and you would be totally correct. Since there's an infinite number of time periods, the household can follow a strategy of this type ad infinitum and "cheat the game". This type of strategy is called a Ponzi scheme.[3]

Let's see how the household can use a Ponzi scheme in its optimization problem. Let $\{\tilde{C}_t, \tilde{A}_{t+1}\}_{t=1}^{\infty}$ be choices that satisfy the budget constraint in all time periods. Let's create new choices, $\{\hat{C}_t, \hat{A}_{t+1}\}_{t=1}^{\infty}$, based on the original choices, $\{\tilde{C}_t, \tilde{A}_{t+1}\}_{t=1}^{\infty}$, that allow the household to increase consumption through a Ponzi scheme. In the new choices, the household consumes one unit more in $t=0$, but consumption doesn't change in the other time periods:

$$\hat{C}_0 = \tilde{C}_0 + 1, \quad \hat{C}_t = \tilde{C}_t \text{ for all } t = 1, 2, \ldots. \tag{10}$$

For the choices $\{\hat{C}_t\}_{t=0}^{\infty}$ to be feasible, they need to be funded by extra borrowings in

[3] Charles Ponzi was a con artist in the early 20th century who orchestrated a notorious investment scheme that collapsed in 1920. His deceptive practices, using new investors' funds to pay off earlier ones, led to the term "Ponzi scheme".

the first time period:

$$\underbrace{\tilde{C}_0 + 1}_{\hat{C}_0} + \underbrace{\tilde{A}_1 - 1}_{\hat{A}_1} = L_0 w_0 + (1 + r_0)\tilde{A}_0.$$

This equation is true because we have assumed that $\tilde{C}_0$ and $\tilde{A}_1$ satisfy the budget constraint in $t = 0$. This equation shows that, for $\hat{C}_0$ to be feasible in the first time period, the household needs to borrow one extra euro in that time period, $\hat{A}_1 = \tilde{A}_1 - 1$.

The budget constraint in the second time period is:

$$\tilde{C}_1 + \underbrace{\tilde{A}_2 - (1 + r_1)}_{\hat{A}_2} = L_1 w_1 + (1 + r_1) \underbrace{(\tilde{A}_1 - 1)}_{\hat{A}_1}.$$

The right-hand side shows that the assets with which the household starts in $t = 1$ is $\hat{A}_1 = \tilde{A}_1 - 1$ because this is the asset level chosen in $t = 0$. Since the household needs to satisfy the budget constraint in that time period, it needs to borrow one euro plus the interest rate of time period 1, $1 + r_1$. Thus, the new asset choice is $\hat{A}_2 = \tilde{A}_2 - (1 + r_1)$.

You might be already seeing the general pattern for the next time periods. Let's see the next time period to understand the general pattern of the new choices. In time period $t = 2$, the budget constraint is

$$\tilde{C}_2 + \underbrace{\tilde{A}_3 - (1 + r_1)(1 + r_2)}_{\hat{A}_3} = L_2 w_2 + (1 + r_2) \underbrace{[\tilde{A}_2 - (1 + r_1)]}_{\hat{A}_1}.$$

Now the assets the household starts with in $t = 2$ are $\tilde{A}_2 - (1+r_1)$. Again, the household needs to borrow more money to satisfy the budget constraint. Since now the agent has to pay interest rate r_2 on top of the previous interest rate, r_1, the additional borrowing is bigger: $(1 + r_1)(1 + r_2)$. Thus, the new asset choice is $\hat{A}_3 = \tilde{A}_3 - (1 + r_1)(1 + r_2)$.

This will go on forever, with borrowings increasing every period. Therefore, the general formula for the new asset choices is[4]

$$\hat{A}_{t+1} = \tilde{A}_{t+1} - \prod_{\tau=1}^{t} (1 + r_\tau) \text{ for all } t = 0, 1, \ldots. \tag{11}$$

Let's now summarize what we have learned. We started with choices that satisfy

[4]The symbol "Π" means product. For example, $\Pi_{\tau=1}^{10}(1 + r_\tau) = (1 + r_1)(1 + r_2) \ldots (1 + r_{10})$.

constraints (15)-(17) and created new choices that allow the household to increase consumption by one unit in the first time period and still satisfy constraints (15)-(17). This shows that choosing consumption and assets to maximize $\sum_{t=0}^{\infty} \beta^t L_t u(C_t/L_t)$ subject to only constraints (15)-(17) is a mathematical problem without solution: one can always modify a given candidate for solution and obtain a higher objective function without violating any constraint.

The no-Ponzi condition The appropriate condition to avoid Ponzi schemes in this model is the constraint (18). This inequality is called the no-Ponzi condition.

To understand the meaning of inequality (18), we need to comprehend the concept of present value calculations. Let's see an example. If the household saves 1€ in $t = 0$, it receives 1€ $\times (1 + r_1)$ in $t = 1$ (i.e, one euro plus interests). Because of this, we say that the present value of $1 \times (1+r_1)$ euros in $t = 1$ measured in $t = 0$ is 1€. Reversely, if the household saves $(1 + r_1)^{-1}$ euros in $t = 0$, it receives $(1 + r_1)^{-1}(1 + r_1) = 1$ euro in $t = 1$. Therefore, we say that the present value of one euro in $t = 1$ measured in $t = 0$ is $(1 + r_1)^{-1}$ euros.

What is the present value of 1€ in $t = 2$ measured in $t = 0$? It is $(1 + r_1)^{-1}(1 + r_2)^{-1}$ because, if the household saves this amount from $t = 0$ until $t = 2$, it gets $(1+r_1)^{-1}(1+r_2)^{-1}(1+r_1)(1+r_2) = 1$ euro in $t = 2$. Therefore, in more generic terms, $\prod_{\tau=1}^{t}(1 + r_\tau)^{-1}$ is the present value of one euro in t measured in time zero.

We say that $\prod_{\tau=1}^{t}(1 + r_\tau)^{-1}$ is the *present-value factor* to calculate the present value in time zero of a variable that exists in time period t. For example, the present value of 20€ in $t = 10$ measured in time zero is $20 \prod_{\tau=1}^{10}(1 + r_\tau)^{-1}$ euros. Again, that's because having $20 \prod_{\tau=1}^{10}(1 + r_\tau)^{-1}$ euros in $t = 0$ and saving this amount from $t = 0$ to $t = 10$ leads to 20€ in $t = 10$.

Now, take a look at inequality (18) again. It contains the term $A_{t+1} \prod_{\tau=1}^{t}(1+r_\tau)^{-1}$. This represents the present value measured in time zero of assets chosen in time period t, A_{t+1}. Since (18) contains a limit with t going to infinity, we read it as saying that *the present value of assets in the long run (infinitely distant future) cannot be negative*. Intuitively, this condition requires the household to not have debt when time approaches infinity. The subsection "The lifetime budget constraint" in section 2.1.3 shows formally that (18) is the appropriate condition to avoid Ponzi schemes in the household's problem.

The no-Ponzi condition can be interpreted as an institutional feature of the asset market. This condition represents an environment where institutions in credit markets can keep track of the financial operations made by all individuals and regulate their behaviors to avoid Ponzi schemes.[5]

The household's problem in terms of per-capita variables Let's rewrite the problem (4)-(8) using variables in per-capita terms.

First, consider the law of motion of population size (1). It can be rewritten as

$$L_{t+1} - L_t = nL_t \quad \forall t = 0, 1, \dots, . \tag{12}$$

This is a *difference equation*: it describes the behavior of the variable L_t through its change over time (the left-hand side is the *difference* between population in $t+1$ and t). Using (1),

$$\begin{aligned} L_1 &= (1+n)L_0 = 1+n, \\ L_2 &= (1+n)L_1 = (1+n)^2, \\ L_3 &= (1+n)L_2 = (1+n)^3, \\ &\vdots \\ L_t &= (1+n)^t. \end{aligned} \tag{13}$$

Equation (13) is the *solution to the difference equation* (1) or (12). We say it is the *solution* because now the population size is not described anymore through its change over time. Equation (13) shows explicitly how population size in t depends on the population growth rate, n, and time, t (and not on population size in a different time period).

Using (13), the household's objective function becomes

$$\sum_{t=0}^{\infty} \beta^t L_t u\left(\frac{C_t}{L_t}\right) = \sum_{t=0}^{\infty} [\beta(1+n)]^t u(c_t).$$

Now let's transform the budget constraint. Divide both sides of (5) by L_t and use (1) to get

$$c_t + (1+n)a_{t+1} = w_t + (1+r_t)a_t.$$

[5] A famous case of a large Ponzi scheme was that of Bernie Madoff in the US. Madoff was sentenced for 150 years in prison in 2009.

Finally, the no-Ponzi condition can be written in per-capita terms as

$$\begin{aligned}
&\lim_{t\to\infty} A_{t+1}\prod_{\tau=1}^{t}(1+r_\tau)^{-1} \geq 0 \\
\Leftrightarrow &\lim_{t\to\infty} a_{t+1}(1+n)^{t+1}\prod_{\tau=1}^{t}(1+r_\tau)^{-1} \geq 0 \\
\Leftrightarrow &\lim_{t\to\infty} a_{t+1}(1+n)\prod_{\tau=1}^{t}\left(\frac{1+r_\tau}{1+n}\right)^{-1} \geq 0 \\
\Leftrightarrow &\lim_{t\to\infty} a_{t+1}\prod_{\tau=1}^{t}\left(\frac{1+r_\tau}{1+n}\right)^{-1} \geq 0.
\end{aligned}$$

To get to the last inequation, we divided both sides of the inequality by $1+n$.

Note that the present-value factor to be used with per-capita variables is different from the one used for variables in levels. The present value factor for per-capita variables is $\prod_{\tau=1}^{t}[(1+r_\tau)/(1+n)]^{-1}$, while the one used for non-per-capita variables is $\prod_{\tau=1}^{t}(1+r_\tau)^{-1}$.

The last thing to notice is that, since $\{L_t\}_{t=0}^{\infty}$ is taken as given by the household, choosing $\{C_t, A_{t+1}\}_{t=0}^{\infty}$ is equivalent to choosing $\{c_t, a_{t+1}\}_{t=0}^{\infty}$. Therefore, the household's problem in per-capita terms is

$$\max_{\{c_t,a_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty}[\beta(1+n)]^t u(c_t) \tag{14}$$

subject to

$$c_t + (1+n)a_{t+1} = w_t + (1+r_t)a_t \text{ for all } t = 0, 1, \ldots, \tag{15}$$

$$a_0 > 0 \text{ given,} \tag{16}$$

$$c_t \geq 0,\ a_{t+1} \in \mathbb{R} \text{ for all } t = 0, 1, \ldots, \tag{17}$$

$$\lim_{t\to\infty} a_{t+1}(1+n)\prod_{\tau=1}^{t}\left(\frac{1+r_\tau}{1+n}\right)^{-1} \geq 0. \tag{18}$$

This per-capita version of the problem shows that the household's effective time discount is given by the term $[\beta(1+n)]^t$. There are two variables affecting time discount. On the one hand, the household is impatient and weights future utilities less than present utility. This mechanism is governed by the parameter β. On the other hand, since population grows, there are more household members in the future, so the total utility of the household tends to be higher in the future because there are more people enjoying utility from consumption. This mechanism is governed by the population growth rate, n. The net effect of these two forces on time discounting is given by $\beta(1+n)$. If $\beta(1+n) < 1$, the household gives more weight to present utility compared to future utility, while the opposite happens if $\beta(1+n) > 1$.

If the household gives more weight to future utilities, the discounted lifetime utility of the household can diverge to infinity even if $\{c_t\}_{t=0}^{\infty}$ is bounded, and this leads to complications that make the model difficult to work with.[6] Therefore, we assume that $\beta(1+n) < 1$.

2.1.2 The cake-eating problem

Before solving the full household problem (14)-(18), let's study a simplified version of it. The objective of this section is to introduce the *transversality condition*, which is a necessary optimality condition commonly featured in maximization problems with infinite time horizons.[7]

The cake-eating problem is a simplified version of the household problem, where $w_t = r_t = 0$ for all t and $n = 0$. We also assume that the utility function is the natural logarithmic function.[8] We will first see the problem assuming that the planning horizon is finite. This will allow us to draw some insights that will be extended later to the infinite planning horizon.

Finite horizon planning Time is discrete and runs from $t = 0$ to $t = T$, where T is a natural positive number. In $t = 0$, a cake eater starts with a cake with size $a_0 > 0$. In each period, it needs to choose how much to consume out of the cake and

[6] Take this simple example: each household member consumes one unit of the consumption good in every period, $c_t = 1$ for all t. If $\beta(1+n) > 1$ and $u(1) > 0$, the discounted lifetime utility of the household, $\sum_{t=0}^{\infty}[\beta(1+n)]^t u(1)$, is the infinite sum of positive numbers that get larger and larger. This sum diverges to infinity. What if $c_t = 2$ for all t? The discounted lifetime utility of the household is the same: infinity. Therefore, in principle, the model doesn't do a good job in terms of describing how a household prefers having more consumption compared to less if $\beta(1+n) > 1$.

If $u(1) < 0$, then $\sum_{t=0}^{\infty}[\beta(1+n)]^t u(1)$ diverges to minus infinity. Compare this scenario with the one where $c_t = 0.5$ for all t: discounted lifetime utility still diverges to minus infinity and we get to the the same conclusion to which we arrived in the previous example.

[7] The main insight of this section is based on teaching notes written by Eric Roca (https://eric-roca.github.io/courses/mathematical_appendix/transversality_condition/). As Roca emphasizes, the discussion in this section is not a formal mathematical proof of the transversality condition, but only an intuitive heuristic argument.

[8] Exercise 1 in section 1 asks you to prove that the utility function $u(c) = \ln(c)$ is a particular case of the CIES utility function.

how much to leave for the next period. That is, the budget constraint is

$$c_t + a_{t+1} = a_t \quad \forall t = 0, 1, \ldots, T.$$

The utility function is the natural log function and the time preference parameter is $\beta \in (0, 1)$. The discounted lifetime utility measured in $t = 0$ is

$$\sum_{t=0}^{T} \beta^t \ln(c_t).$$

The maximization problem is:

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

$$\text{subject to}$$

$$c_t + a_{t+1} = a_t \qquad \forall t = 0, 1, \ldots, T,$$

$$a_0 > 0 \text{ given},$$

$$c_t \geq 0, a_{t+1} \geq 0 \qquad \forall t = 0, 1, \ldots, T.$$

One difference between the problem above and the household problem (14) is that, in the cake-eating problem, a_t cannot be negative (no negative cake, which makes sense). Because of this, Ponzi schemes are not possible in this problem, so a no-Ponzi condition is not necessary for this problem to be well defined.

Since time ends at $t = T$, a curious feature arrises in the problem above. Note that the cake eater needs to choose a_{T+1}, but it doesn't get utility from consuming the cake in $T + 1$. There's not even a "c_{T+1}" in the problem above! Therefore, the cake eater will never leave a positive amount of cake for period $T + 1$. That is, the optimal choice needs to make $a_{T+1} = 0$.

The problem above is a maximization problem with equality and inequality constraints. We will use the Lagrangian method to solve it.[9]

One of the constraints is that a_{t+1} needs to be greater or equal to zero for all t. We saw that, at the optimal choices, the cake eater would like to make $a_{T+1} = 0$. This is called a "corner solution" because the variable a_{T+1} is in the "corner" of the interval that restricts which values the variable a_{T+1} can assume, $[0, \infty)$. A mathematically formal way to deal with optimization problems with corner solutions is to consider

[9] See Appendix B for a discussion on Lagrangians.

inequality constraints in the Lagrangian. In our case, we implement this idea by including the positivity constraint "$a_{T+1} \geq 0$" in the Lagrangian. We don't need to include the other positivity constraints $a_{t+1} \geq 0$ for $t < T-1$ because it's not optimal to choose $a_{t+1} = 0$ for some $t < T$ (otherwise there's nothing to consume, and $\log(c) \to -\infty$ when $c \to 0$), so we don't need to worry about positivity constraints for $t < T$.

We write the Lagrangian as

$$\mathcal{L} = \sum_{t=0}^{T} \beta^t \ln(c_t) + \sum_{t=0}^{T} \lambda_t \left(a_t - c_t - a_{t+1}\right) + \mu a_{T+1},$$

where λ_t are the Lagrangian multipliers associated with the budget constraint (equality constraints), and μ is the Lagrangian multiplier related to the positivity constraint $a_{T+1} \geq 0$ (inequality constraint).

Making the derivative with respect to c_t (for any given $t = 0, 1, \ldots, T$) equal to zero, we get

$$\frac{\beta^t}{c_t} = \lambda_t \quad \forall t = 0, 1, \ldots, T. \tag{19}$$

This is called the *first order condition* (FOC) for optimality with respect to c_t. We also need the FOC with respect to a_{t+1}. For a moment, let's assume that $T = 10$. The cake eater needs to choose $a_1, a_2, \ldots, a_{10}, a_{11}$. Let's say we are taking the derivative of the Lagrangian with respect to a_1. Note that a_1 appears in the first and second terms of the second summation in the Lagrangian. Therefore, the FOC for a_1 is $-\lambda_0 + \lambda_1 = 0$. Similarly, the FOC with respect to a_2 is $-\lambda_1 + \lambda_2 = 0$. So it seems that the general pattern is that $-\lambda_t + \lambda_{t+1} = 0$, right? Be careful! The FOC for a_{11} is different: it is given by $-\lambda_{10} + \mu = 0$.

Therefore, the general FOCs (for a generic T) with respect to a_{t+1} is

$$\lambda_t = \lambda_{t+1} \quad \forall t = 0, 1, \ldots, T-1, \tag{20}$$

$$\lambda_T = \mu. \tag{21}$$

Finally, there's one last optimality condition: the slackness condition associated with the inequality constraint $a_{T+1} \geq 0$.[10] The slackness condition is:

$$\mu a_{T+1} = 0. \tag{22}$$

Remember that $\mu \geq 0$.

Using (19) and (20),

$$c_{t+1} = \beta c_t \quad \forall t = 0, 1, \ldots, T-1.$$

This equation says that consumption in $t+1$ will be smaller than consumption in t, and the magnitude of this consumption decrease over time depends on how impatient the cake eater is (β). This is a simple version of the Euler equation that we will see in detail in section 2.1.3. Let's rewrite the equation above for some time periods to understand its pattern:

$$c_1 = \beta c_0,$$
$$c_2 = \beta c_1 = \beta\beta c_0 = \beta^2 c_0,$$
$$c_3 = \beta c_2 = \beta\beta^2 c_0 = \beta^3 c_0.$$

So, in general, we can write

$$c_t = \beta^t c_0 \quad \forall t = 0, 1, \ldots T. \tag{23}$$

This equation gives us c_t as a function of c_0. We can also find a_{t+1} as a function of c_0. Using the budget constraints and the (23),

$$a_1 = a_0 - c_0,$$
$$a_2 = a_1 - c_1 = a_0 - c_0 - \beta c_0,$$
$$a_3 = a_2 - c_2 = a_0 - c_0 - \beta c_0 - \beta^2 c_0.$$

Thus,

$$a_{t+1} = a_0 - \sum_{\tau=0}^{t} \beta^\tau c_0 \quad \forall t = 0, 1, \ldots T.$$

We can work this equation:

$$a_{t+1} = a_0 - \sum_{\tau=0}^{t} \beta^\tau c_0 = a_0 - c_0 \left(\frac{1 - \beta^{t+1}}{1 - \beta} \right) \quad \forall t = 0, 1, \ldots T, \tag{24}$$

[10] If the term "slackness condition" is new to you, please read Appendix B. The slackness condition is important for the argument of this section.

where we have used the formula for the sum of the first $t+1$ terms of a geometric progression with ratio β and initial value of 1.

Note that equations (23) and (24) describe choice variables (c_t and a_{t+1}) as functions of c_0, which is also a choice variable. Therefore, c_0 is the only variable to be determined. After we find c_0 as a function of variables that the cake eater takes as given and parameters (a_0 and β), we have solved the problem.

Recall we have used all first-order conditions to obtain (23) and, on top of that, the budget constraints to obtain (24). Since these equations don't fully characterize the optimal choices only as functions of variables the cake eater takes as given (they depend on c_0), we learn that *the first-order conditions and budget constraints are not enough to fully characterize the optimal choices*.

The only optimality conditions we haven't used so far are (21) and the slackness condition (22). Remember that these two conditions are related to the optimal choice of a_{T+1}. We will use exactly this condition to find c_0. The important idea here is that, to find the choice of consumption *in the first time period*, c_0, we need to use optimality conditions associated to the cake size *in the last time period* in which the agent makes a choice, a_{T+1}. In other words, to find the *initial condition* for c, we need to use a *terminal condition* for a. This idea will show up again when we solve the household's problem in the Ramsey model.

Using (21) and (22), we find that

$$\lambda_T a_{T+1} = 0. \tag{25}$$

Equation (25) is a very important optimality condition. It is called the *transversality condition* (TVC). This condition characterizes the optimal choice of the agent in the last time period in which it makes choices. When we solve the cake-eating problem with an infinite time horizon, we will use an infinite-time-horizon version of the transversality condition (25).

Now we can finish solving the problem. Using (19) for $t = T$ in the equation above:

$$\frac{\beta^T}{c^T} a_{T+1} = 0 \Leftrightarrow a_{T+1} = 0,$$

where we have used the fact that β^T and c^T are strictly positive. Using (24) for $t = T$

in the previous equation,

$$a_{T+1} = a_0 - c_0 \left(\frac{1-\beta^{T+1}}{1-\beta} \right) = 0 \Leftrightarrow c_0 = a_0 \left(\frac{1-\beta}{1-\beta^{T+1}} \right).$$

We have finally found c_0 as a function of parameters. This completes the solution to the cake-eating problem with a finite time horizon.

Infinite planning horizon Now assume that there are infinite time periods, $t = 0, 1, \ldots$. The maximization problem is:

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

$$\text{subject to}$$

$$c_t + a_{t+1} = a_t \qquad \forall t = 0, 1, \ldots$$

$$a_0 > 0 \text{ given,}$$

$$c_t \geq 0, a_{t+1} \geq 0 \qquad \forall t = 0, 1, \ldots$$

First, write the Lagrangian:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \ln(c_t) + \sum_{t=0}^{\infty} \lambda_t \left(a_t - c_t - a_{t+1} \right).$$

We don't consider inequality constraints in the Lagrangian above because a_{t+1} needs to be strictly positive in all time periods: if the cake's size reaches zero in finite time, the agent's utility goes to minus infinity in that time period, and this can't be optimal.

The FOCs with respect to c_t and a_{t+1} are:

$$\frac{\beta^t}{c_t} = \lambda_t \quad \forall t = 0, 1, \ldots \tag{26}$$

$$\lambda_t = \lambda_{t+1} \quad \forall t = 0, 1, \ldots \tag{27}$$

After the same manipulations we did in the previous section, we get to the same equations:

$$c_t = \beta^t c_0 \quad \forall t = 0, 1, \ldots \tag{28}$$

$$a_{t+1} = a_0 - c_0 \left(\frac{1-\beta^{t+1}}{1-\beta} \right) \quad \forall t = 0, 1, \ldots \tag{29}$$

To arrive at equations (28) and (29), we have used all FOCs and budget constraints in all time periods. Similarly to the previous case, all these optimality conditions are not enough to fully solve the problem. We still need to find c_0 as a function of parameters.

Let's see this point more carefully. Some values that we choose for c_0 will give us variables outside of feasible bounds: take a look at (29). There's a negative sign there. If we choose a c_0 that is very high, we may get a negative a_{t+1} for some finite t, and this cannot happen. Note that we don't have this problem for equation (28). Any positive value of c_0 will give us positive values for c_t for all t's.

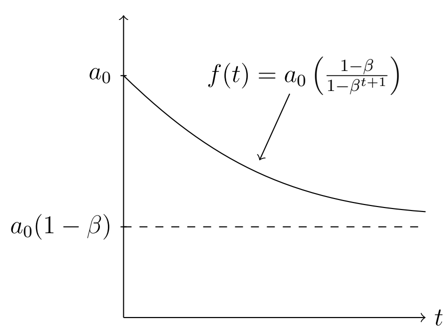
Let's see which values of c_0 will give us a_{t+1} greater than or equal to zero for all t. Working with equation (29), we have

$$a_0 - c_0 \left(\frac{1-\beta^{t+1}}{1-\beta} \right) \geq 0 \Leftrightarrow a_0 \geq c_0 \left(\frac{1-\beta^{t+1}}{1-\beta} \right) \Leftrightarrow a_0 \left(\frac{1-\beta}{1-\beta^{t+1}} \right) \geq c_0. \tag{30}$$

That is, values of c_0 that satisfy the last inequality above for all t will give us a_{t+1} greater than or equal to zero for all time periods. In other words, we have upper bounds for c_0. They are in the left-hand side of the last inequality above. Let's make clear that the left hand-side is a function of time by defining the function $f(t)$ as

$$f(t) = a_0 \left(\frac{1-\beta}{1-\beta^{t+1}} \right).$$

Note that $\lim_{t\to\infty} f(t) = (1-\beta)a_0$. We can plot the function $f(t)$:



The last inequality in (30) is saying that we must have $c_0 \leq f(t)$ for all t. The figure shows that any strictly positive c_0 smaller or equal to $(1-\beta)a_0$ will do this job. Thus, if we pick any $c_0 \in (0, (1-\beta)a_0]$ and use equations (28) and (29) and obtain $\{c_t, a_{t+1}\}_{t=1}^{\infty}$, these choices satisfy all FOCs and $a_{t+1} > 0$ for all t. Thus, there is an infinite number of solutions to the optimality and feasibility constraints that we have gotten so far, (28) and (29).

Take a step back and think about this: if we didn't know about the transversality condition, we would not be able to formally solve to cake-eating problem with an infinite planning horizon! We have used all first-order conditions and budget constraints, and this allowed us to write all endogenous variables as functions of c_0. But, without the TVC, we cannot solve for c_0 as a function of variables taken as given by the cake eater.

To find the unique and correct solution for c_0, we need the transversality condition. The transversality condition for the infinite-horizon problem is obtained by simply taking the limit of (25) with time going to infinity:

$$\lim_{t\to\infty} \lambda_t a_{t+1} = 0. \tag{31}$$

This is the transversality condition for an infinite-horizon optimization problem. This condition is extremely important in the neoclassical growth model. It is a *terminal condition* for a because it is related to the limit of a_{t+1} with t approaching infinity.

Using (26), (28) and (29) in the transversality condition,

$$\lim_{t\to\infty} \frac{\beta^t}{c_t}\left[a_0 - c_0\left(\frac{1-\beta^{t+1}}{1-\beta}\right)\right] = 0 \iff \lim_{t\to\infty} \frac{\beta^t}{\beta^t c_0}\left[a_0 - c_0\left(\frac{1-\beta^{t+1}}{1-\beta}\right)\right] = 0$$
$$\iff \lim_{t\to\infty} a_0 - c_0\left(\frac{1-\beta^{t+1}}{1-\beta}\right) = 0 \iff a_0 - \frac{c_0}{1-\beta} = 0 \iff c_0 = a_0(1-\beta).$$

Finally, we have found the optimal c_0 as a function of exogenous variables! If you want, you can plug this c_0 in (28) and (29) and find the optimal paths of c_t and a_{t+1} over time.

A similar intuition to the one presented in the previous section works here. The transversality condition (31) characterizes the optimal choice "at the end of the planning horizon" (when t goes to infinity). Using that condition, we can find the consumption in the first time period. That is, by considering the *terminal condition* of the cake size, we find the *initial condition* for consumption.

We saw before that, if we ignore the transversality condition, any $c_0 \in (0, (1-\beta)a_0]$ would satisfy all FOCs and feasibility constraints. However, it is intuitive that any c_0 strictly smaller than $(1-\beta)a_0$ would not be a solution to the maximization problem. Utility is increasing with respect to consumption, so, if $c_0 < (1-\beta)a_0$, we could still increase c_0 (to some value lower than $(1-\beta)a_0$) and we would still have a cake with positive size in all time periods and, from (28), we would consume more in all time periods. In other words, if we had chosen to start with some c_0 strictly smaller than $(1-\beta)a_0$, the cake eater would have left a positive amount of cake that would never be eaten, even though the cake would be eaten in every period.

The TVC selects the initial consumption such that, starting from that consumption level and using the Euler equation (28) to obtain consumption in subsequent time periods, the cake eater consumes the whole cake during the full time horizon.

2.1.3 Solving the household's problem

Now let's go back to the Ramsey model and find the optimality conditions that characterize the household's problem (14)-(18). Using the insights from the cake-eating problem, we write the Lagrangian as

$$\mathcal{L} = \sum_{t=0}^{\infty} \left\{ [\beta(1+n)]^t u(c_t) + \lambda_t [w_t + (1+r_t)a_t - c_t - a_{t+1}(1+n)] \right\}.$$

The optimality conditions are given by the first order conditions (FOCs),

$$\frac{\partial \mathcal{L}}{\partial c_t} = 0 \quad \forall t = 0, 1, \ldots,$$

$$\frac{\partial \mathcal{L}}{\partial a_{t+1}} = 0 \quad \forall t = 0, 1, \ldots,$$

and the transversality condition

$$\lim_{t \to \infty} \lambda_t a_{t+1} = 0. \tag{32}$$

Let's first work with the FOCs:[11]

$$\frac{\partial \mathcal{L}}{\partial c_t} = 0 \Rightarrow [\beta(1+n)]^t u'(c_t) = \lambda_t, \tag{33}$$

$$\frac{\partial \mathcal{L}}{\partial a_{t+1}} = 0 \Rightarrow \lambda_t(1+n) = \lambda_{t+1}(1+r_{t+1}). \tag{34}$$

Using (33) for t and $t+1$ and substituting them in (34), we arrive at

$$u'(c_t) = \beta(1+r_{t+1})u'(c_{t+1}). \tag{35}$$

[11] If you don't understand (34) at first sight, note that $\partial\mathcal{L}/\partial a_1 = -\lambda_0(1+n) + \lambda_1(1+r_1)$ and $\partial\mathcal{L}/\partial a_2 = -\lambda_1(1+n) + \lambda_2(1+r_2)$. The general pattern is $\partial\mathcal{L}/\partial a_{t+1} = -\lambda_t(1+n) + \lambda_{t+1}(1+r_{t+1})$.

The Euler equation Equation (35) is called the Euler equation. It describes how the household optimally allocates consumption across time. The intuition behind (35) is that the household equalizes marginal utilities in two consecutive time periods, adjusting for its impatience, measured by β, and by the return on postponing consumption, $1 + r_{t+1}$.

To understand this better, let's think about the limit case where the household is not impatient, $\beta = 1$, and there's no monetary returns on savings, $r_{t+1} = 0$. In this case, the household would like to equalize its marginal utilities in two consecutive time periods, $u'(c_t) = u'(c_{t+1})$. The best way to understand this is to think what would happen if this equality doesn't hold. Let's say that $u'(c_t) > u'(c_{t+1})$. Since the marginal utility in t is higher than in $t + 1$, consuming more in t and less in $t + 1$ (through lowering savings) would increase the household's total utility because the loss in utility from a lower consumption in $t+1$ is smaller than the gain in utility from a higher consumption in t. An analogous explanation shows that $u'(c_t) < u'(c_{t+1})$ is not optimal either.

What if β is strictly less than one? In this case (still assuming $r_{t+1} = 0$), the household is impatient and would choose to have a higher c_t and lower c_{t+1} (compared to the previous paragraph), thus decreasing the marginal utility in t and increasing it in $t + 1$ so that the equality $u'(c_t) = \beta u'(c_{t+1})$ holds.

Finally, if the interest rate is positive, $r_{t+1} > 0$, then there are monetary gains from savings. Now the household wants to consume less in t and more in $t + 1$ (compared to the previous paragraph) to make the equality (35) hold.

The marginal utility function $u'(\cdot)$ shows up in the Euler equation. Let's see how the utility function affects optimal choices. Rewrite the Euler equation as

$$\frac{u'(c_t)}{u'(c_{t+1})} = \beta(1 + r_{t+1}). \tag{36}$$

If $\beta(1 + r_{t+1}) = 1$, the consumer makes $c_t = c_{t+1}$. The farther $\beta(1 + r_{t+1})$ is from one, the further apart c_t is from c_{t+1}. Now, let's see how a change in the $u'(\cdot)$ function affects consumption choices. First, remember that $u'(\cdot)$ is a decreasing function of consumption because $u(\cdot)$ is concave ($u''(c) < 0$ for all $c > 0$). Fix the right-hand side of (36) and assume that $\beta(1 + r_{t+1}) \neq 1$. The optimal pair (c_t^*, c_{t+1}^*) is one that makes the ratio of marginal utilities equal to $\beta(1 + r_{t+1})$. Since $\beta(1 + r_{t+1}) \neq 1$, $c_t^* \neq c_{t+1}^*$. What happens to the optimal choice of (c_t, c_{t+1}) if $u'(\cdot)$ is modified such that the marginal utility function falls faster with consumption (e.g., the $u''(\cdot)$ function is more negative)? Since we are keeping the right hand side fixed, now the new optimal choices (c_t^{**}, c_{t+1}^{**}) need to be closer to each other (compared to the distance between c_t^* and c_{t+1}^*) so that the ratio of marginal utilities is kept fixed and (36) holds.

A numerical example might make things clearer. Assume that, with the original utility function, optimal choices are such that $c_{t+1}^* > c_t^*$ and $u'(c_t^*) > u'(c_{t+1}^*)$. Let's say that the marginal utility in t is two times that in $t+1$, $u'(c_t^*)/u'(c_{t+1}^*) = 2$. If the $u'(\cdot)$ function is modified so that it falls faster with consumption, now the marginal utility falls more when consumption increases from c_t^* to c_{t+1}^*, implying that $u'(c_t^*)/u'(c_{t+1}^*)$ is now higher than two. To take the ratio of marginal utilities back to two, c_t^{**} and c_{t+1}^{**} get close to each other, so that the marginal utility doesn't fall as much when c_t^{**} increases to c_{t+1}^{**}.

The lesson here is that, if the marginal utility function is more sensitive to changes in consumption (i.e., it falls faster with consumption increases), then the consumer will avoid having large changes in consumption over time. Equivalently, the consumer wants to *smooth consumption* more when the utility function is more concave. The economic intuition is that, if the utility function falls faster with consumption, it is as if the consumer gets satisfied quicker (full satisfaction would be reached in the limit case $\lim_{c\to\infty} u'(c) = 0$). If this agent gets satisfied more easily, it prefers not getting close to satisfaction at a given point in time because that would imply having a relatively high marginal utility in a different time period.

Note that, even though interest rates show up in the Euler equation, wage rates don't. This happens because the household can freely move resources across time periods using assets, regardless of when these real resources will actually arrive. Think about two extreme cases. In the first case, wage rates in all periods are zero, except in the first time period, when the wage rate is one million euros. The household will choose to consume a small fraction of the wage in the first time period and save the rest in order to have resources to consume in all future time periods. In the second case, wage rates in all periods are zero, except in the time period one million, when the wage is one million euros. The household will then borrow money in all periods until time period one million. In that date, the household will receive the wage, pay for all its past borrowings, and still save a fraction of the inflow to use for consumption in each future time period. Therefore, the wage in period t doesn't affect the change in consumption between t and $t+1$.

The Constant Intertemporal Elasticity of Substitution (CIES) utility function A frequently used utility function is the Constant Intertemporal Elasticity of Substitution (CIES) utility function:[12]

$$u(c) = \frac{c^{1-\theta} - 1}{1-\theta},$$

where $\theta > 0$ is a parameter. This utility function has the property that the elasticity[13] of the marginal utility with respect to consumption is constant (independent of the level of consumption) and equal to $-\theta$. That is, an infinitesimally small increase in consumption decreases the marginal utility by θ%.

We saw before that an important characteristic of the utility function that shows up in the Euler equation is how fast the marginal utility falls with consumption. A nice property of the CIES utility function is that the θ parameter captures exactly this feature. Let's see this in detail.

Since $u'(c) = c^{-\theta}$ with the CIES utility, the Euler equation (35) can be rewritten as

$$c_{t+1} = c_t[\beta(1 + r_{t+1})]^{1/\theta}. \tag{37}$$

To understand how θ affects intertemporal consumption choices, define[14] ρ through

[12]The CIES utility function is also known as Constant Relative Risk Aversion (CRRA) utility function. Since there's no uncertainty/risk in the benchmark Ramsey model, the interpretation related to risk is not interesting for our purposes.

[13]The elasticity of a real-valued univariate function $f(x)$ is formally given by $f'(x)x/f(x)$. It describes by how much $f(x)$ changes in percentage terms if x changes infinitesimally in percentage terms. Recall that the derivative captures how much $f(x)$ changes relative to the change in x (i.e., $f'(x_0) \approx [f(x_1) - f(x_0)]/(x_1 - x_0)$ if x_1 is close to x_0). The idea behind the elasticity of $f(x)$ is to describe this change in proportional terms instead of of absolute terms. Therefore, the elasticity of $f(x)$ at $x = x_0$ can be though of as

$$\frac{[f(x_1) - f(x_0)]/f(x_0)}{[x_1 - x_0]/x_0} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}\frac{x_0}{f(x_0)} \approx f'(x_0)\frac{x_0}{f(x_0)}.$$

for x_1 close to x_0.

An alternative way of understanding the elasticity of a univariate function is to recall that the derivative of the natural log of $f(x)$ with respect to the natural log of x gives the percentage change in $f(x)$ as a result of an infinitesimal change in x in percentage terms, or

$$\frac{\partial \ln(f(x))}{\partial \ln(x)} = \frac{\partial \ln(f(e^{\ln(x)}))}{\partial \ln(x)} = f'(x)\frac{x}{f(x)},$$

which is exactly equal to the elasticity of $f(x)$ with resect to x.

[14]See appendices E.2 and F.1 for a detailed discussion on ρ and β. For now, we just need to keep in mind that ρ is inversely related to β: a more impatient consumer is described by a lower β and a higher ρ.

$\beta = (1+\rho)^{-1}$ and take natural logs of both sides of the equation above to get

$$\ln(c_{t+1}) - \ln(c_t) = \frac{1}{\theta}[\ln(1+r_{t+1}) - \ln(1+\rho)]. \tag{38}$$

The equation above allows us to see how θ affects intertemporal consumption choices. If r_{t+1} is equal to ρ, the left hand side is equal to zero, implying that consumption doesn't change between $t+1$ and t. If r_{t+1} is greater (smaller) than ρ, consumption increases (decreases) from t to $t+1$. The magnitude of the change in consumption depends on $1/\theta$. The larger the θ, the closer to zero is the variation in consumption. This result is consistent with the intuition we have seen before. Since θ measures the elasticity of the marginal utility with respect to consumption, a high θ represents a household that wants to smooth consumption by a considerable degree, thereby choosing to have small variations in consumption over time. Conversely, a low θ represents a marginal utility function less responsive to consumption changes. Thus, the household equalizes marginal utilities in two different time periods (after controlling for the interest rate and time discount) through consumption levels that are distant from each other.

The no-Ponzi and the transversality conditions Let's use (37) to solve for c_t as a function of c_0. Solving recursively,[15]

$$c_t = c_0 \left[\beta^t \prod_{\tau=1}^{t} (1+r_\tau)\right]^{1/\theta}. \tag{39}$$

Note we haven't used the transversality condition (32) yet. Using (33) and (39) in the transversality condition (32),

$$\lim_{t\to\infty} [\beta(1+n)]^t c_0^{-\theta} \left[\beta^t \prod_{\tau=1}^{t} (1+r_\tau)\right]^{-1} a_{t+1} = 0.$$

Simplifying the expression above and multiplying both sides by $c_0^\theta > 0$,

$$\lim_{t\to\infty} a_{t+1} \prod_{\tau=1}^{t} \left(\frac{1+r_\tau}{1+n}\right)^{-1} = 0. \tag{40}$$

[15]First, note that $c_1 = c_0[\beta(1+r_1)]^{1/\theta}$. Second, $c_2 = c_1[\beta(1+r_2)]^{1/\theta}$. Substituting the first equation in the second equation, $c_2 = c_0[\beta^2 \prod_{\tau=1}^{2}(1+r_\tau)]^{1/\theta}$. One can see that the general pattern is (39). More formally, (39) can be proved by induction.

This is an important condition and says that the present-value of assets in the "infinitely distant future" (as t approaches ∞) needs to be zero. As seen in the cake-eating problem, this condition ensures that all resources available to the household are consumed throughout its lifetime.

A strictly positive present value of assets in the infinitely distant future would mean that the household has positive assets in the "end" of the planning horizon. Intuitively, this is as if the household is saving so much that some portion of its resources would be saved forever and never consumed.

Note that the TVC (40) is similar to the no-Ponzi condition (18). The no-Ponzi condition is an inequality, while the TVC simply substitutes the inequality in (18) for an equality. Although they are written similarly, they have very distinct "natures": the no-Ponzi condition is a characteristic of the economic environment; a feature of the credit market; a constraint that institutions impose on households to avoid unfair strategies on the asset market. Conversely, the TVC is an optimality condition that the household voluntarily chooses to follow to maximize its welfare. The no-Ponzi condition sends the following message to households: "You cannot cheat the game! The present value of your assets in the long run cannot be negative. In other words, you cannot have debt in the long run. However, the present value of your assets can be strictly positive in the long run if you want". The household then replies "Fine, I will not have debt in the long run. But I will not choose to have positive assets in the long run either. I am not stupid!".

The lifetime budget constraint In exercise X, you are asked to show that the budget constraints from time period 0 to $T > 0$ imply that

$$\left[\sum_{t=0}^{T} c_t D_t\right] + (1+n)a_{T+1}D_t = \left[\sum_{t=0}^{T} w_t D_t\right] + (1+r_0)a_0, \tag{41}$$

where D_t is the present-value factor for per-capita variables,

$$D_t = \begin{cases} \prod_{\tau=1}^{t} \left(\frac{1+r_\tau}{1+n}\right)^{-1} & \text{if } t > 0 \\ 1 & \text{if } t = 0. \end{cases}$$

Equation (41) is the budget constraint aggregated from time period 0 to T and it has an interesting economic interpretation. The right-hand side contains the present value of all wage payments from 0 to T and the initial asset income, $(1+r_0)a_0$. These two elements constitute the real resources that the household has available to use in the time interval between 0 and T. The left-hand side shows how the household uses these resources in this time interval: each unit of resource can either be used for consumption or for accumulating assets to when the time interval ends (i.e., in $t = T + 1$).

Why don't assets in time periods $t = 1, 2, \ldots, T$ show up in (41)? This happens because these assets are not real resource inflows: they come from initial assets and wages received between 0 and T. In other words, these assets are a mechanism to transfer real resources over time. Conversely, assets in time zero, a_0, *are* real resources in the sense that, when the model starts in the first time period, $(1+r_0)a_0$ is an inflow of resources into the current account of the household that don't come from previous wage payments.

Since equation (41) holds for any $T = 0, 1, 2, \ldots$, we can take the limit of both sides with $T \to \infty$,

$$\left[\sum_{t=0}^{\infty} c_t D_t\right] + (1+n) \underbrace{\lim_{T\to\infty} a_{T+1} D_t}_{\geq 0 \text{ (no-Ponzi)}} = \left[\sum_{t=0}^{\infty} w_t D_t\right] + (1+r_0)a_0.$$

An important insight can be drawn here and shows that the no-Ponzi condition (18) is the "appropriate" condition to avoid Ponzi schemes in this model. The right-hand side of this equation contains all resources that the household has available to use throughout its lifetime. It is given by all wage payments and initial assets. The left-hand side shows that these resources are used either as consumption throughout life or as assets accumulated "in the long run" (when t tends to ∞). The no-Ponzi condition requires the present value of these assets in the long run to be non-negative. If the household could choose to have negative assets, the equation above makes clear that the household would be able to consume more than its lifetime resources.

This shows that the no-Ponzi condition (18) is not "too strong" nor "too weak". It constraints the household by the exact amount so that it cannot use more resources than those available to be used. Note that one simple alternative way to avoid Ponzi schemes in this model is to prevent borrowings ($a_{t+1} \geq 0$ for all t). This constraint, however, is too strong. At a given point in time, there are wage payments to be made to the household in the future. The household should, therefore, be allowed to borrow some amount to anticipate future wage payments. The no-Ponzi condition (18) does exactly that.

Note that, to get to the equation above, we haven't used any optimality conditions. We have used only the budget constraint equations. Therefore, one doesn't need to consider the household's optimal choices to arrive at the mathematical expression for the no-Ponzi condition.

However, if now we consider the optimality conditions for the household choices, we can use the transversality condition in the equation above:

$$\left[\sum_{t=0}^{\infty} c_t D_t\right] + (1+n) \underbrace{\lim_{T\to\infty} a_{T+1} D_t}_{=0 \text{ from TVC (40)}} = \left[\sum_{t=0}^{\infty} w_t D_t\right] + (1+r_0)a_0. \tag{42}$$

The transversality condition implies that the household will not choose to have strictly positive assets in the long-run. Therefore, the second term of the left-hand side disappears and we get to the result that the present value of all consumption made throughout the household's lifetime must be equal to the real resources available for the household over its planning horizon. If the limit in the equation above were strictly positive, this would imply that the household consumes less resources than the total amount available for consumption throughout its life.

The initial consumption Previously, we have found the optimal c_t as a function of c_0 (equation (39)). Let's now find the optimal c_0 as a function of variables that the household takes as given. Substitute the c_t of equation (39) into the lifetime budget constraint (42) to get

$$c_0 = \frac{1}{\mu}[(1+r_0)a_0 + \tilde{w}], \tag{43}$$

where $1/\mu$ is the propensity to consume out of lifetime resources in the first time period, and μ is given by

$$\mu \equiv \sum_{t=0}^{\infty} e^{[\ln(\beta(1+n)) + \frac{1-\theta}{\theta}(\ln(\beta)+\bar{r}_t)]t}, \tag{44}$$

$\bar{r}_t$ is the average (of an increasing function of the) interest rate between 0 and t,

$$\bar{r}_t \equiv \frac{1}{t}\sum_{\tau=1}^{t} \ln(1+r_\tau). \tag{45}$$

and $\tilde{w} \equiv \sum_{t=0}^{\infty} w_t D_t$ is the present value of all wage payments.

An increase in average interest rates, $\bar{r}_t$, for given lifetime resources, has two effects on the propensity to consume in the first period. First, higher interest rates increase the cost of current consumption relative to future consumption, an *intertemporal-substitution effect* that motivates households to shift consumption from the present to the future. Second, higher interest rates have an *income effect* that tends to raise consumption at all dates. The net effect of an increase in $\bar{r}_t$ on $1/\mu$ depends on which of the two forces dominates. If $\theta < 1$, $1/\mu$ declines with $\bar{r}(t)$ because the substitution effect dominates. The intuition is that, when θ is low, households care relatively little about consumption smoothing, and the intertemporal-substitution effect is large. Conversely, if $\theta > 1$, $1/\mu$ rises with $\bar{r}(t)$ because the substitution effect is relatively weak. Finally, if $\theta = 1$ (log utility), the two effects exactly cancel. The effects of $\bar{r}(t)$ on $1/\mu$ carry over to effects on c_0 if we hold constant the lifetime resources term, $(1 + r_0)a_0 + \tilde{w}$. In fact, however, $\tilde{w}$ falls with $\bar{r}_t$ for a given path of w_t. This third effect reinforces the substitution effect that we mentioned before.[16]

A final technical insight is noting that we needed to use use the TVC in (42) to obtain the solution for c_0. This is similar to the mechanics of the cake-eating problem, where the terminal condition for assets determine the initial consumption.

2.2 Firms

Firms take as given the paths of the wage rates, capital rental rates, and labor productivity, $\{w_t, R_t, T_t\}_{t=0}^{\infty}$, and choose capital and labor to maximize profits. For each t, their maximization problem is

$$\max_{K_t^d, L_t^d} (K_t^d)^\alpha (T_t L_t^d)^{1-\alpha} - R_t K_t^d - w_t L_t^d,$$

where we have used the supperscript "d" to denote "demand". This notation will be useful to distinguish labor demand, L_t^d, from labor supply (or population size), L_t, in the next section.

The first order conditions for optimality are

$$\underbrace{\frac{\partial\{(K_t^d)^\alpha (T_t L_t^d)^{1-\alpha}\}}{\partial K_t^d}}_{\text{Marginal productivity of capital}} = \underbrace{R_t}_{\text{Price of capital}},$$

[16]This paragraph was copied and adapted from "Economic Growth" by Barro and Sala-i-Martin.

$$\underbrace{\frac{\partial\{(K_t^d)^\alpha (T_t L_t^d)^{1-\alpha}\}}{\partial L_t^d}}_{\text{Marginal productivity of labor}} = \underbrace{w_t}_{\text{Price of labor}}.$$

These equations say that the firm equalizes the marginal productivity of a given input to this input's price. Taking the derivatives explicitly,

$$\alpha(K_t^d)^{\alpha-1}(T_t L_t^d)^{1-\alpha} = R_t,$$

$$(1-\alpha)(K_t^d)^{\alpha} T_t^{1-\alpha}(L_t^d)^{-\alpha} = w_t.$$

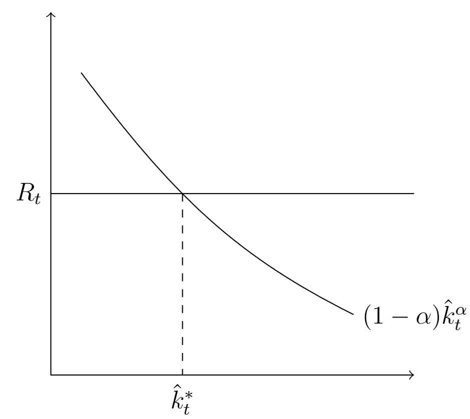
We can write these FOCs in terms of per-effective-worker variables as

$$\alpha(\hat{k}_t^d)^{\alpha-1} = R_t, \tag{46}$$

$$(1-\alpha)(\hat{k}_t^d)^{\alpha} T_t = w_t, \tag{47}$$

where $\hat{k}_t^d \equiv K_t^d/(T_t L_t^d)$ is capital per effective worker demanded by the firm.

The following plot illustrates equation (46):



The flat horizontal line represents the right-hand side of equation (46), while the downward-sloping curve represents the left-hand side (remember that $\alpha - 1 < 0$). If the level of capital per worker is below $\hat{k}_t^*$, using one additional unit of capital per worker increases production by $\alpha\hat{k}_t^{\alpha-1}$ (marginal productivity of capital), while the cost increases by R_t (marginal cost of capital). Since the first term is larger than the second, profits increase. Therefore, choosing any level below $\hat{k}_t^*$ is not optimal because profits can be made larger if more capital per worker is used.

Now think about the opposite case and assume that $\hat{k}_t$ is higher than $\hat{k}_t^*$. If capital per worker decreases a bit, production falls by the marginal productivity of capital and costs fall by the marginal cost of capital. Since the first term is lower than the second, the loss in revenues is smaller than the fall in costs, so profits increase. This means that any level of capital per worker higher than $\hat{k}_t^*$ is not optimal. Therefore, profits are maximized when $\hat{k}_t = \hat{k}_t^*$. A similar plot can be used to represent the economic intuition behind labor demand in equation (47).

Equations (46) and (47) can be explicitly written as optimal input demand functions. Rewrite $\hat{k}_t^d$ as a function of input prices to get

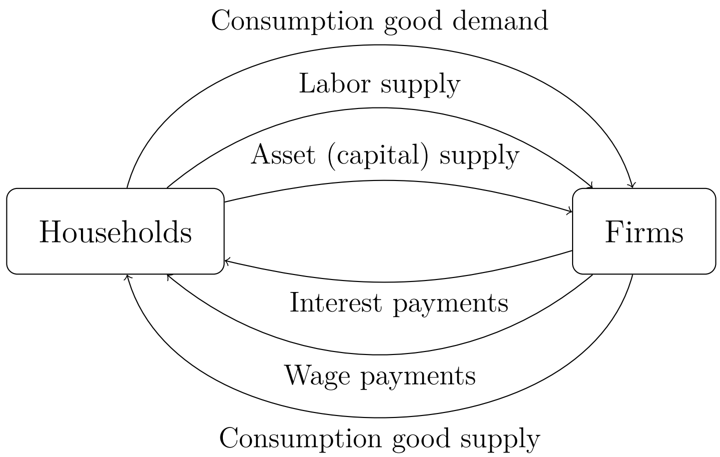
$$\hat{k}_t^d = \left(\frac{\alpha}{R_t}\right)^{\frac{1}{1-\alpha}},$$

$$\hat{k}_t^d = \left(\frac{w_t}{T_t(1-\alpha)}\right)^{\frac{1}{\alpha}}.$$

The equations above describe the optimal level of capital per worker as a function of input prices, R_t and w_t, and labor productivity, T_t. The first equation says that optimal capital per worker demand is negatively related to the price of capital, R_t. If the price of capital increases, the firm substitutes labor for capital (more labor, less capital), implying a lower ratio of capital per worker. The second equation says that the optimal capital per worker demand is positively correlated to the price of labor, w_t, and negatively correlated to labor productivity, T_t. If the wage rate increases or if labor productivity falls, the firm substitutes capital for labor (more capital, less labor), leading to a higher capital per effective worker.

2.3 General equilibrium

Market clearing In the two previous sections, we have studied the optimal behavior of households and firms. In each time period, household members supply their labor force, households choose how much to consume and how many assets to take to the next time period. Firms demand capital and labor to produce the final good in each time period. Households and firms interact in three markets: the labor market, the asset/capital market, and the consumption good's market.



This diagram shows that household members constitute the supply side in the labor and capital markets because they are the agents who supply labor and capital, while firms constitute the demand side. The opposite happens in the consumption good's market, with households functioning as the demand side and firms operating in the supply side.

The firm pays w_t for each worker it hires. Equivalently, each household member receives w_t for its labor supply. Additionally, the firm pays the capital rental rate R_t for each unit of capital, while the household receives the interest rate r_t for each unit of asset it supplies. The relationship between the interest rate and the capital rental rate is

$$R_t = r_t + \delta, \tag{48}$$

where $\delta \in (0, 1)$ is a parameter denoting the capital depreciation rate: a fraction δ of the capital stock in the economy depreciates in each period.

An intuitive way to interpret (48) is that, when the firm rents a given amount of capital from households, it promises to give back to households this same amount of capital plus the interest rate r_t per unit of capital. However, a fraction δ of the capital depreciates after its usage in production. Thus, after production takes place the firm needs to replenish the depreciated capital (so that it can return the full capital stock) and, on top of that, pay the interest rate.[17]

[17] Here's an example with numbers. Let's say that the firm rents 100 units of capital/assets, the interest rate is 10% ($r = 0.1$) and the depreciation rate is 5% ($\delta = 0.05$). The household lends 100€ and expects to receive 110€ back because the interest rate is 10%. The firm uses 100 units of capital. After production, 5 units of capital disappear because of depreciation, so the firm has 95 units of capital after the production stage. The firm needs to pay 110€ to households, so the effective cost for the firm is 5 (to replenish capital) plus 10 (interest rate), which corresponds to $100(r + \delta)$.

We saw that households and firms take the prices (w_t, r_t and R_t) as given. At the same time, prices are determined so that all markets clear – i.e, so that supply equals demand in all markets. Therefore, the general equilibrium conditions (market clearing conditions) are:

$$L_t = L_t^d \text{ for all } t = 0, 1, \dots, \tag{49}$$

$$A_t = K_t^d \text{ for all } t = 0, 1, \dots. \tag{50}$$

The first condition represents the labor market clearing, with the left-hand side being the population size (labor supply) and the right-hand side denoting labor demand from firms. The second condition represents capital market clearing. The left-hand side of (50) denotes the aggregate asset supply and the right-hand side captures the aggregate capital demand from firms.

Note that the conditions above naturally imply that market clearing holds in terms of per-capita and per-effective-labor terms, $a_t = k_t^d \equiv k_t$ and $\hat{a}_t = \hat{k}_t^d \equiv \hat{k}_t$.

The Walras' law is a common feature of macroeconomic models and says that, if an economy has $N \in \mathbb{N}$ markets and $N - 1$ markets clear, then all markets clear. In our case, the economy has three markets. Conditions (49) and (50) ensure that the labor and asset markets clear. By Walras' law, if these two conditions hold, then the consumption good market also clears. Let's see this in detail.

You are asked to show in exercise X that the firms' profits must be zero in general equilibrium. Using this fact and the market clearing conditions $L_t^d = L_t$ and $A_t = K_t^d \equiv K_t$,

$$Y_t = w_t L_t + R_t K_t.$$

Using this equation, (48) and the market clearing conditions in the household's budget constraint (5), we obtain

$$\underbrace{C_t}_{\text{Demand of consumption good}} = \underbrace{\underbrace{Y_t}_{\text{Total final goods}} - \underbrace{(K_{t+1} - (1-\delta)K_t)}_{\text{New capital}}}_{\text{Supply of consumption good}}.$$

This equation shows the aggregate demand for consumption goods in the left-hand side. The right-hand side is given the total amount of final goods, Y_t, minus the new stock of capital accumulated from t to $t + 1$, net of depreciation. The quantity of final goods produced in t that is not used for capital constitutes the aggregate supply of consumption goods. Thus, we see that, if labor and asset markets clear, then the consumption good market also clears.

System of difference equations Our next objective is to obtain general equilibrium equations only in terms of $\hat{c}$, $\hat{k}$, and exogenous variables. First, start with the Euler equation (37), and use firm's FOC for capital (46), $R_t = r_t + \delta$, and $\hat{a}_t = \hat{k}_t^d = \hat{k}_t$ to arrive at[18]

$$\hat{c}_{t+1} - \hat{c}_t = \hat{c}_t \left\{ (1+x)^{-1} [\beta(\alpha \hat{k}_{t+1}^{\alpha-1} + 1 - \delta)]^{1/\theta} - 1 \right\}. \tag{51}$$

Second, use the budget constraint (15), the firm's FOCs (46) and (47), $R_t = r_t + \delta$, $\hat{a}_t = \hat{k}_t^d = \hat{k}_t$, and $L_t^d = L_t = (1+n)^t$, to get

$$\hat{k}_{t+1} - \hat{k}_t = [(1+n)(1+x)]^{-1} \left\{ \hat{k}_t^{\alpha} - [(1+n)(1+x) - (1-\delta)]\hat{k}_t - \hat{c}_t \right\}. \tag{52}$$

The only endogenous variables in equations (51) and (52) are $\hat{c}$ and $\hat{k}$. That is, there are no prices (w, r, R) in these equations. Recall we obtained the Euler equation assuming that households take prices w and r as given. The Euler equation is related to how households optimally distribute consumption over time and, therefore, how they optimally save (supply assets). Thus, the Euler equation implicitly generates an asset supply curve as a function of prices. Similarly, the firms' FOCs were obtained under the assumption that w and R are taken as given by the firm. This generates labor and capital demand as functions of prices. Therefore, when we use the market clearing conditions $\hat{a}_t = \hat{k}_t^d$ and $L_t^d = L_t$ to obtain equations (51) and (52), we are implicitly finding the prices that clear the labor and asset/capital markets.

Equations (51) and (52) describe the changes in consumption and capital between t and $t+1$ as a function of the levels of $\hat{c}$ and $\hat{k}$. (51) and (52) are called *difference equations* with respect to $\hat{c}$ and $\hat{k}$, respectively, because these mathematical objects describe the behavior of consumption and capital through their *differences* across time; i.e, the left-hand side of these equations are the differences $\hat{c}_{t+1} - \hat{c}_t$ and $\hat{k}_{t+1} - \hat{k}_t$. The change in consumption between t and $t+1$ depends on the level of $\hat{c}$ in t and on the level of $\hat{k}$ in $t+1$, while the change in capital depends on the levels of $\hat{c}$ and $\hat{k}$ in t. Consumption and capital are interconnected, and the evolution of one these variables over time cannot be studied independently from the other variable. Because of this, we say that (51)-(52) constitute a *system of difference equations* in terms of $\hat{c}$ and $\hat{k}$.

[18] Equation (51) looks better in logs: $\ln(\hat{c}_{t+1}) - \ln(\hat{c}_t) = \frac{1}{\theta}[\ln(\alpha \hat{k}_{t+1}^{\alpha-1} + 1 - \delta) - \ln(1+\rho) - \theta \ln(1+x)]$.

Similarly to differential equations, difference equations need initial conditions to provide a full description of the unknowns. We therefore need an initial condition for capital and consumption. The initial capital in the economy, $K_0 > 0$, is a parameter of the model and it determines the initial condition for capital. Since T_0 and L_0 are also exogenous (equal to one), $\hat{k}_0 = K_0/(T_0 L_0) = K_0$ is exogenous. Let's make this clear by writing

$$\hat{k}_0 > 0 \text{ given.} \tag{53}$$

We cannot do the same for the initial consumption and say it is exogenous in the model. Remember the household chooses c_0, so $\hat{c}_0$ is an endogenous variable. As seen in the cake-eating problem, the last condition we need to impose to determine the initial consumption is the transversality condition. Let's use the TVC (40), (46), $\hat{a}_t = \hat{k}_t$, and $a_t = \hat{a}_t(1+x)^{-t}$ to obtain the transversality condition in general equilibrium:

$$\lim_{t\to\infty} \hat{k}_{t+1} \prod_{\tau=1}^{t} \left[\frac{\alpha \hat{k}_{\tau+1}^{\alpha-1} + (1-\delta)}{(1+n)(1+x)} \right]^{-1} = 0. \tag{54}$$

The transversality condition in general equilibrium works as a *terminal condition* for capital because it is related to capital in the long run ($\hat{k}_t$ when $t \to \infty$). Similarly to the mechanics of the cake-eating problem, this is a necessary condition to determine the initial consumption in general equilibrium.

To summarize, equations (51)-(54) form a system of difference equations that fully describe the behavior of consumption and capital per effective labor for each time period $t = 0, 1, \ldots$. (53) is the initial condition for capital. (54) is a terminal condition for capital, which pins down the initial condition for consumption.

The phase diagram: $\hat{k}_t = \hat{k}_{t+1}$ and $\hat{c}_t = \hat{c}_{t+1}$ locci We will study the behavior of the system of differential equations describing consumption and capital in general equilibrium. We repeat the system here for convenience:

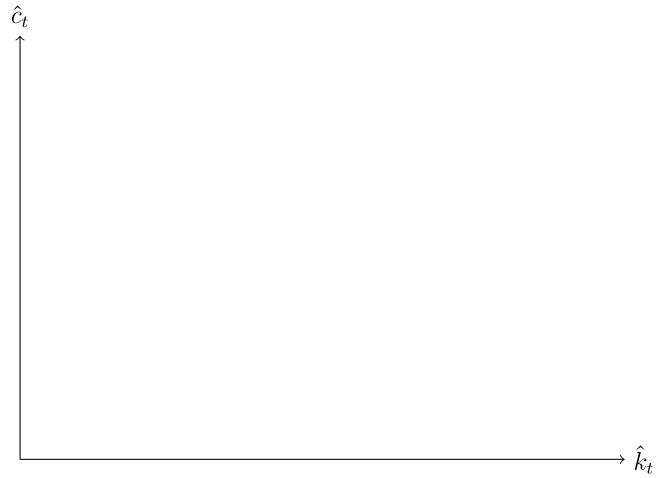
$$\hat{c}_{t+1} - \hat{c}_t = \hat{c}_t \left\{ (1+x)^{-1} [\beta(\alpha \hat{k}_{t+1}^{\alpha-1} + 1 - \delta)]^{1/\theta} - 1 \right\}, \tag{55}$$

$$\hat{k}_{t+1} - \hat{k}_t = [(1+n)(1+x)]^{-1} \left\{ \hat{k}_t^{\alpha} - [(1+n)(1+x) - (1-\delta)]\hat{k}_t - \hat{c}_t \right\}, \tag{56}$$

$$\hat{k}_0 > 0 \text{ given,} \tag{57}$$

$$\lim_{t\to\infty} \hat{k}_{t+1} \prod_{\tau=1}^{t} \left[\frac{\alpha \hat{k}_{\tau+1}^{\alpha-1} + (1-\delta)}{(1+n)(1+x)} \right]^{-1} = 0. \tag{58}$$

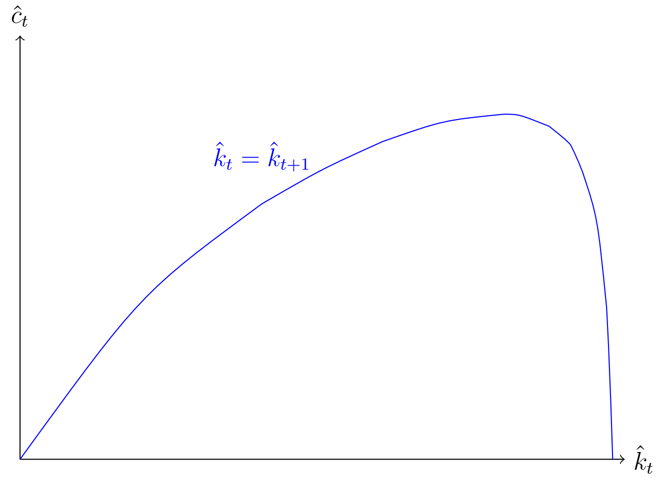
To understand the behavior of this system, we use the *phase diagram*, which is a graphical tool to analyze bivariate systems of difference equations in discrete time (or differential equations in continuous time). In our case, the phase diagram is drawn on the two-dimensional $(\hat{k}_t, \hat{c}_t)$ plane:



The first step is to study the pairs of capital and consumption implying that one of these variables (or both) don't move between two subsequent time periods. Let's start with the pairs $(\hat{k}_t, \hat{c}_t)$ implying that capital doesn't move between t and $t+1$. Making $\hat{k}_{t+1} - \hat{k}_t = 0$ in equation (56), we can write the following equation

$$\hat{c}_t = \hat{k}_t^{\alpha} - [(1+n)(1+x) - (1-\delta)]\hat{k}_t. \tag{59}$$

This equation describes a relationship between $\hat{c}$ and $\hat{k}$ ($\hat{c}$ as a function of $\hat{k}$) associated with $\hat{k}_t = \hat{k}_{t+1}$. $\hat{c}_t$ is equal to a concave and increasing function of capital, $\hat{k}_t^{\alpha}$, minus a linear function of capital, $[(1+n)(1+x)-(1-\delta)]\hat{k}_t$. Note that $(1+n)(1+x) - (1-\delta) > 0$. This implies that the pairs of consumption and capital associated with $\hat{k}_t = \hat{k}_{t+1}$ are described by an inverse U-shaped curve in the $(\hat{k}_t, \hat{c}_t)$ plane:



The blue curve above is called the "$\hat{k}_t = \hat{k}_{t+1}$ locus", and it has the following interpretation: if, at any given point in time, the economy is on the blue curve (meaning that the levels of capital and consumption satisfy equation (59)), then capital doesn't move from the current time period to the next ($\hat{k}_t = \hat{k}_{t+1}$).

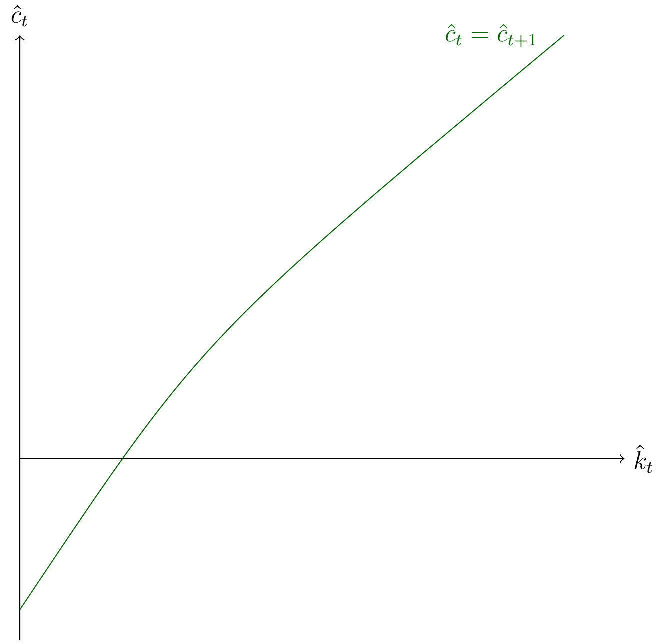
Now, let's find the pairs $(\hat{k}_t, \hat{c}_t)$ associated with $\hat{c}_t = \hat{c}_{t+1}$ using (55). But first note that (59) describes a relationship between variables in time period t. This happens because the right-hand side of (56) only contain variables in t. Note that this is different in equation (55), where the right-hand side depends on $\hat{c}_t$ and $\hat{k}_{t+1}$. Thus, isolate $\hat{k}_{t+1}$ in equation (56) and substitute in equation (55) to obtain

$$\hat{c}_{t+1} - \hat{c}_t = \hat{c}_t \left\{ (1+x)^{-1} \left[\beta\alpha \left(\phi^{-1} \left\{ \hat{k}_t^{\alpha} - [\phi - \psi]\hat{k}_t - \hat{c}_t \right\} + \hat{k}_t \right)^{\alpha-1} + \beta\psi \right]^{1/\theta} - 1 \right\}, \tag{60}$$

where $\phi \equiv (1+n)(1+x)$ and $\psi \equiv 1-\delta$. Now the right-hand side only depends on variables in t. Making $\hat{c}_{t+1} - \hat{c}_t = 0$, we can write

$$\hat{c}_t = \hat{k}_t^{\alpha} + (1-\delta)\hat{k}_t - [(1+n)(1+x)] \left(\frac{\alpha}{(1+x)^{\theta}\beta^{-1} - (1-\delta)} \right)^{\frac{1}{1-\alpha}}. \tag{61}$$

This implies that the pairs of consumption and capital associated with $\hat{c}_t = \hat{c}_{t+1}$ are described by an increasing and concave curve in the $(\hat{k}_t, \hat{c}_t)$ plane, with its intercept being negative because $(1+x)^{\theta}\beta^{-1} - (1-\delta) > 0$:



The green curve above is called the "$\hat{c}_t = \hat{c}_{t+1}$ locus", and it has the following interpretation: if, at any given point in time, the economy is on the green curve (meaning that the levels of capital and consumption satisfy equation (61)), then consumption doesn't move from the current time period to the next ($\hat{c}_t = \hat{c}_{t+1}$).

What happens if we make $\hat{c}_t = \hat{c}_{t+1}$ in equation (51)? Let's see:

$$0 = \hat{c}_t \left\{ (1+x)^{-1} [\beta(\alpha \hat{k}_{t+1}^{\alpha-1} + 1 - \delta)]^{1/\theta} - 1 \right\}$$
$$\Leftrightarrow \hat{k}_{t+1} = \left(\frac{\alpha}{(1+x)^{\theta}\beta^{-1} - (1-\delta)} \right)^{\frac{1}{1-\alpha}} \equiv \hat{k}_{ss}.$$

That is, if consumption doesn't change from t from $t+1$, then capital in $t+1$ must be equal to the right-hand side of first equality in the second row. Note it doesn't depend on time. This capital amount is called the *steady-state capital*, denoted by $\hat{k}_{ss}$, because we will see that this is the capital level to which the economy converges in the long-run. This long-run situation is called the steady state because the variables per effective units of labor ("hat variables", $\hat{k}$, $\hat{c}$, $\hat{y}$) are constant over time (steady).

The phase diagram: Locci intersection A natural question at this point is: where do the two locci (plural of locus) intersect with each other? Does the green curve intersect with the blue curve exactly at peak of the blue curve? To the left of the peak? To the right?

To answer this, let's first characterize the capital level associated with the peak of the blue curve. The blue curve is described by (59). Therefore, we only need to take the derivative of the right-hand side of (59) and make it equal to zero:

$$\begin{aligned} \alpha\hat{k}_t^{\alpha-1} - [(1+n)(1+x)-(1-\delta)] &= 0 \\ \Leftrightarrow \hat{k}_t = \left(\frac{\alpha}{(1+n)(1+x)-(1-\delta)}\right)^{\frac{1}{1-\alpha}} &\equiv \hat{k}_{\text{gold}}. \end{aligned}$$

The capital that maximizes consumption in the $\hat{k}_t = \hat{k}_{t+1}$ locus is called the golden rule capital and is denoted by $\hat{k}_{\text{gold}}$.[19]

To find out where the blue and green curves intersect, we need to use the transversality condition (58). Let's assume that capital converges to its steady state level (we will see later that this is indeed true). From the TVC,

$$\lim_{t\to\infty} \underbrace{\hat{k}_{t+1}}_{\text{converges to } \hat{k}_{ss}} \underbrace{\prod_{\tau=1}^{t}\left[\frac{\alpha\hat{k}_{\tau+1}^{\alpha-1}+(1-\delta)}{(1+n)(1+x)}\right]^{-1}}_{\text{needs to converge to a value smaller than one}} = 0.$$

The product term (e.g., the term "$\prod_{\tau=1}^{t}[\cdot]^{-1}$") only depends on time through capital, $\hat{k}_{\tau+1}$. To ensure the limit above is zero, the product term must converge to a value smaller than one. That's because multiplying a number infinitely many times by another number between zero and one results in a value approaching zero.

Now, note that

$$\begin{aligned} \lim_{t\to\infty}\prod_{\tau=1}^{t}\left[\frac{\alpha\hat{k}_{\tau+1}^{\alpha-1}+(1-\delta)}{(1+n)(1+x)}\right]^{-1} &< 1 \\ \Rightarrow \lim_{t\to\infty}\frac{[(1+n)(1+x)]^t}{\prod_{\tau=1}^{t}\alpha\hat{k}_{\tau+1}^{\alpha-1}+(1-\delta)} &< 1 \\ \Rightarrow \frac{(1+n)(1+x)}{\alpha\hat{k}_{ss}^{\alpha-1}+(1-\delta)} &< 1 \end{aligned}$$

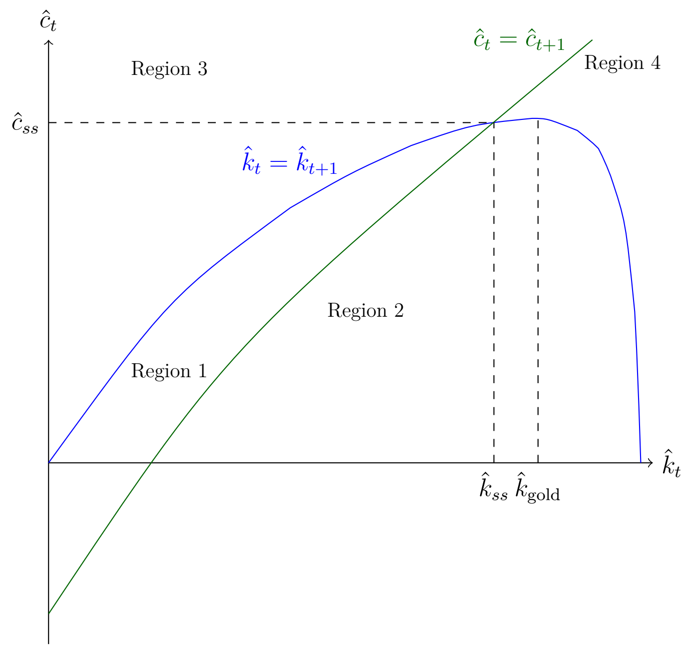
because $\hat{k}_t \to \hat{k}_{ss}$ as $t \to \infty$. The last inequality implies that

$$\hat{k}_{ss} < \left(\frac{\alpha}{(1+n)(1+x)-(1-\delta)}\right)^{\frac{1}{1-\alpha}} = \hat{k}_{\text{gold}}.$$

This derivation shows that, for the TVC to be valid, the steady state capital must

[19] Maybe something about the golden rule capital here? Mention the Solow model.

be lower than the golden rule capital. This implies that the $\hat{c}_t = \hat{c}_{t+1}$ locus intersects with the $\hat{k}_t = \hat{k}_{t+1}$ locus to the left of the peak of the $\hat{k}_t = \hat{k}_{t+1}$ locus:[20]



The intersection of the two locci is the steady state. If the economy is on this intersection, both consumption and capital per effective units of labor don't move over time. The steady state consumption is obtained by setting $\hat{c}_{t+1} - \hat{c}_t = \hat{k}_{t+1} - \hat{k}_t = 0$, or

$$\hat{c}_{ss} = \hat{k}_{ss}^{\alpha} - [(1+n)(1+x) - (1-\delta)]\hat{k}_{ss}.$$

The phase diagram: Movement direction We learned before that if the economy is on a given locus, the variable associated to this locus doesn't change in that particular time period. But what happens if the economy is not on a locus? Let's investigate.

Let's start with the $\hat{k}_t = \hat{k}_{t+1}$ locus. Take a pair $(\hat{k}'_t, \hat{c}'_t)$ on this locus. We know that

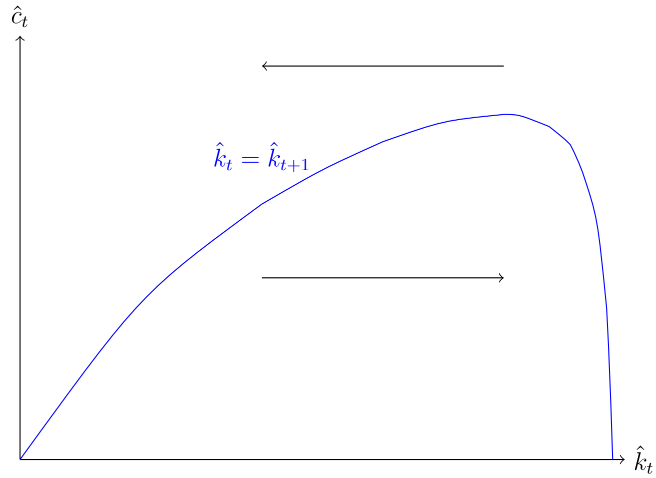
$$\hat{k}'_{t+1} - \hat{k}'_t = [(1+n)(1+x)]^{-1} \left\{ (\hat{k}'_t)^{\alpha} - [(1+n)(1+x) - (1-\delta)]\hat{k}'_t - \hat{c}'_t \right\} = 0.$$

[20]By substituting the expressions $\hat{k}_{ss}$ and $\hat{k}_{\text{gold}}$ into the inequality $\hat{k}_{ss} < \hat{k}_{\text{gold}}$, we obtain the conditions on the parameters $\beta(1+n) < (1+x)^{\theta-1}$. This represents a requirement for the model's parameters to ensure the existence of equilibrium. The economic intuition behind this condition is that...

What happens if we pick another pair with the same level of capital but with a higher consumption? That is, pick $\hat{k}''_t$ and $\hat{c}''_t$ with $\hat{k}''_t = \hat{k}'_t$ and $\hat{c}''_t > \hat{c}'_t$. Since consumption is preceded by a minus sign in the equation above, we have that

$$\hat{k}''_{t+1} - \hat{k}''_t = [(1+n)(1+x)]^{-1} \left\{ (\hat{k}''_t)^\alpha - [(1+n)(1+x) - (1-\delta)]\hat{k}''_t - \hat{c}''_t \right\} < 0.$$

That is, if at a given point in time t the economy has a pair of capital and consumption above the $\hat{k}_t = \hat{k}_{t+1}$ locus, then capital falls between t and $t+1$. This can be represented in the phase diagram through arrows:



A horizontal arrow pointing left means that capital falls (because capital is on the horizontal axis of the phase diagram), while an arrow pointing right means capital increases. Thus, the arrows in the figure above mean that that, if the economy is above the $\hat{k}_t = \hat{k}_{t+1}$ locus in t, then capital falls between t and $t+1$. Conversely, if the economy is below the blue curve in t, then capital increases between t and $t+1$.

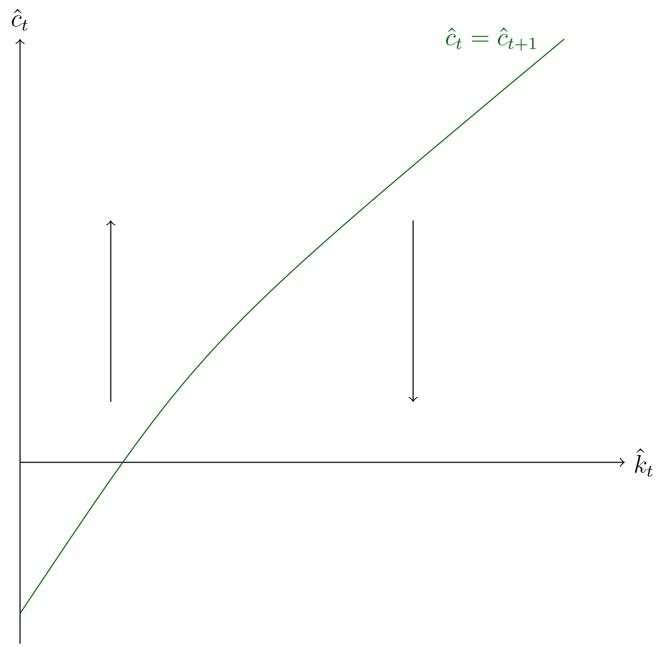
Let's follow the same logic now using the $\hat{c}_t = \hat{c}_{t+1}$ locus. Pick a pair $(\hat{k}'_t, \hat{c}'_t)$ on the $\hat{c}_t = \hat{c}_{t+1}$ locus. Using (60),

$$\hat{c}'_{t+1} - \hat{c}'_t = \hat{c}'_t \left\{ (1+x)^{-1} \left[\beta\alpha \left(\phi^{-1} \left\{ (\hat{k}'_t)^\alpha - [\phi - \psi]\hat{k}'_t - \hat{c}'_t \right\} + \hat{k}'_t \right)^{\alpha-1} + \beta\psi \right]^{1/\theta} - 1 \right\} = 0.$$

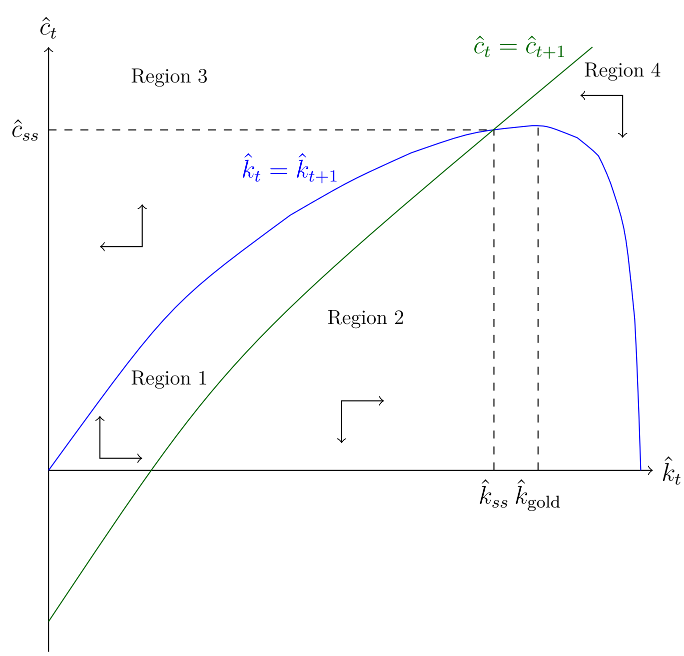
Now increase $\hat{c}$ keeping $\hat{k}$ fixed. That is, pick $\hat{k}''_t = \hat{k}'_t$ and $\hat{c}''_t > \hat{c}'_t$. Since $\alpha - 1 < 0$,

$$\hat{c}''_{t+1} - \hat{c}''_t = \hat{c}''_t \left\{ (1+x)^{-1} \left[\beta\alpha \left(\phi^{-1} \left\{ (\hat{k}''_t)^\alpha - [\phi - \psi]\hat{k}''_t - \hat{c}''_t \right\} + \hat{k}''_t \right)^{\alpha-1} + \beta\psi \right]^{1/\theta} - 1 \right\} > 0.$$

This shows that, if the economy is above the $\hat{c}_t = \hat{c}_{t+1}$ locus in t, consumption increases between t and $t+1$. Conversely, if the economy is below the green line, consumption falls from t to $t+1$. This is graphically represented by the vertical arrows in the figure below.



We can draw a phase diagram with vertical and horizontal arrows simultaneously:



It's useful to give names to some areas in the phase diagram. We will call the area below the blue curve and above the green curve region 1. Regions 2, 3, and 4 are also shown in the figure above and, similarly to region 1, are defined in terms of their positions relative to the blue and green curves. The arrows show the directions to which variables move in the four regions. For example, in region 2, capital grows (horizontal arrow pointing to the right) and consumption falls (vertical arrow pointing downward).

The phase diagram: General equilibrium What we learned before simply gives us information about what would happen if the economy happens to be at a given point in the phase diagram. A natural question is: in which points of the phase diagram can the economy actually be and still satisfy the system of differential equations (55)-(58)? That is, where in the phase diagram can the economy be *in general equilibrium*?

We first show that the economy cannot be in regions 2 or 3 and satisfy (55)-(58).

First, assume that the economy is in region 3. The arrows show that capital falls and consumption grows. Over time, the economy gets farther from the blue line, implying that capital falls by bigger magnitudes each time. Thus, two things might happen. First, equation (56) might project a fall in capital so large that it would imply a negative amount of capital in finite time period. This cannot happen because the model doesn't allow for a negative capital stock. Second, equation (56) might indicate that, in a given time period where $\hat{k}_t > 0$, capital decreases precisely to the extent that a zero capital stock is projected in the subsequent period. Considering that equation (55) involves the term $\hat{k}_{t+1}^{\alpha-1}$ and $\alpha - 1 < 0$, a capital stock nearing zero would imply infinite growth in consumption. Yet, this is unfeasible since a zero capital stock results in zero final goods, and thus zero consumption. In conclusion, the economy cannot be in region 3 while still satisfying equations (55), (56]), and the inequality $\hat{k}_t \geq 0$ for all t.

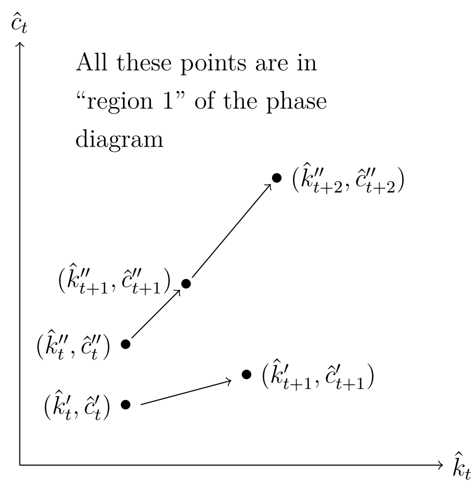
Assume now that the economy is in region 2. Both capital and consumption fall continuously. Capital will converge to the capital level associated with the point where the blue curve touches the horizontal axis to the right of the golden rule capital. Recall we showed before that the level to which capital converges should be lower than the golden rule capital for the TVC to hold. Therefore, the economy being in region 3 would violate the TVC.

We already mentioned that the economy will converge to the steady state point, where the blue and green curves intersect. Since we are interested in *economic growth*, the interesting parametric cases for us are those where the initial capital level $\hat{k}_0$ is below $\hat{k}_{ss}$. In this scenario, capital grows, implying that $\hat{y}_t$ grows and there is economic growth.

Thus, assume $\hat{k}_0 < \hat{k}_{ss}$. Since we already saw the economy cannot be in region 3, then it must start in region 1.

We will see one important property of region 1: if we compare two economies with the same parameters, starting from the same capital level, but economy one starting with a lower consumption than economy two, economy one will display a higher (lower) capital (consumption) growth than economy two for each time period when both economies are in region 1.

Let's say economy one starts with $(\hat{k}'_t, \hat{c}'_t)$ and economy two with $(\hat{k}''_t, \hat{c}''_t)$, where $\hat{k}'_t = \hat{k}''_t$ and $\hat{c}'_t < \hat{c}''_t$. Equation (56) shows that capital in the second economy will grow less than in the first because consumption has a minus sign in the right-hand side of (56). Thus, $\hat{k}'_{t+1} > \hat{k}''_{t+1}$. Equation (55) implies that consumption in the second economy grows more than in the first because the right-hand side of (55) is positively related to consumption in t and negatively related to capital in $t+1$ ($\alpha - 1 < 0$). Therefore, $\hat{c}'_{t+1} - \hat{c}'_t < \hat{c}''_{t+1} - \hat{c}''_t$. Since $\hat{c}'_t < \hat{c}''_t$, we have $\hat{c}'_{t+1} < \hat{c}''_{t+1}$. The capital-consumption pairs in time periods t and $t+1$ in the figure below illustrate this:



Let's say $(\hat{c}'_{t+1}, \hat{k}'_{t+1})$ and $(\hat{c}''_{t+1}, \hat{k}''_{t+1})$ are still in region 1. Will the same growth pattern happen between $t+1$ and $t+2$? That is, will capital (consumption) in economy one grow more (less) than in economy two? Equation (56) shows that capital growth depends on consumption and capital levels. First, a lower consumption in the first economy, compared to the second economy, contributes to a larger capital growth (because of the minus sign). Second, the capital level affects capital growth non-linearly through the expression $\hat{k}^\alpha - [(1+n)(1+x) - (1-\delta)]\hat{k}$. Note that

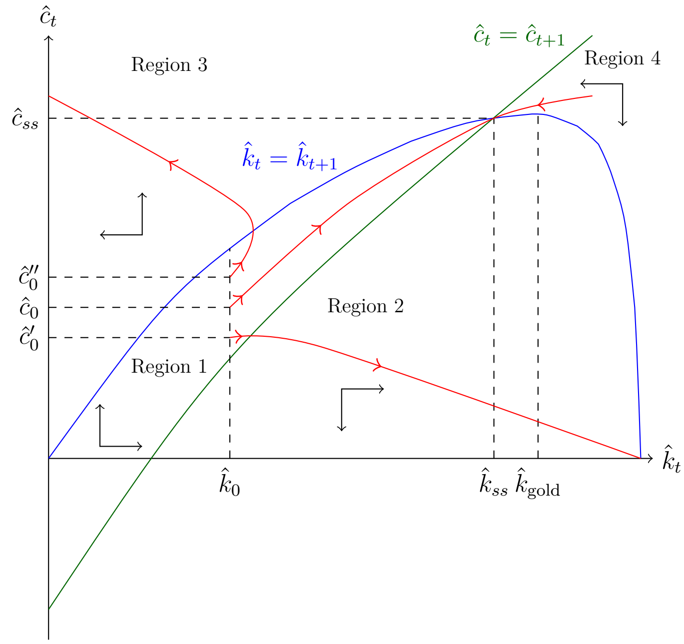
$$\frac{\partial\{\hat{k}_{t+1}^{\alpha} - [(1+n)(1+x)-(1-\delta)]\hat{k}_{t+1}\}}{\partial \hat{k}_{t+1}} = \alpha\hat{k}_{t+1}^{\alpha-1} - [(1+n)(1+x)-(1-\delta)] > 0$$
$$\Leftrightarrow \hat{k}_{t+1} < \left(\frac{\alpha}{(1+n)(1+x)-(1-\delta)}\right)^{\frac{1}{1-\alpha}} = \hat{k}_{\text{gold}}.$$

Since this last inequality is true because $(\hat{c}'_{t+1}, \hat{k}'_{t+1})$ and $(\hat{c}''_{t+1}, \hat{k}''_{t+1})$ are in region 1, a higher capital level in the first economy contributes to a larger capital growth, compared to the one in the second economy. We conclude that the lower consumption and higher capital in the first economy lead to a higher capital growth than in the second economy between $t+1$ and $t+2$.

Finally, equation (55) shows that, since $\hat{c}''_{t+1} > \hat{c}'_{t+1}$ and $\hat{k}''_{t+2} < \hat{k}'_{t+2}$, consumption in the second economy grows faster than in the first.

This argument shows that, as long as both economies are in region 1, capital growth in the first economy is higher than in the second economy in each time period, while consumption growth in the second economy is higher than in the first for each t.

We can use this property to talk about the existence and unicity of the general equilibrium path in the Ramsey economy. The figure below shows three possible levels of initial consumption, $\hat{c}'_0 < \hat{c}_0 < \hat{c}''_0$. If the initial consumption is too low, $\hat{c}'_0$, from region 1's property we just learned, consumption will increase too slowly while capital grows quickly in the first time periods. The economy will cross the green curve and transit to region 2. We saw that the system (55)-(58) cannot be satisfied in region 2. We conclude that the initial consumption $\hat{c}'_0$ cannot occur in equilibrium.



Note that, since the difference equations (55) and (56) are continuous, slightly increasing the initial consumption from $\hat{c}'_0$ to $\hat{c}'_0 + \varepsilon$, where $\varepsilon > 0$ is very small, implies that consumption will now grow a bit more and capital a bit less. Since the increase in $\hat{c}_0$ is small, the economy will again cross the green line and transit to region 2, but now the capital level in the first period when the economy hits region 2 is higher than the case with initial consumption $\hat{c}_0$.

On the other hand, if we assume that the initial consumption is too high, $\hat{c}''_0$, consumption grows quickly in the first time periods, while capital moves up slowly. The economy will cross the blue line and reach region 3. We saw this cannot happen in general equilibrium, so $\hat{c}''_0$ cannot occur in equilibrium. Slightly decreasing the initial consumption to $\hat{c}_0 - \varepsilon$ implies that consumption grows a bit less and capital a bit more. The economy will hit region 3 again, but now with a higher capital than in the scenario with initial consumption $\hat{c}''_0$.

Based on this continuity idea, we can see that there's only one possible level of initial consumption, denoted by $\hat{c}_0$ in the figure below. Starting from $(\hat{k}_0, \hat{c}_0)$ the variables will grow in each time period satisfying equations (55) and (56), while not leaving region 1. As capital and consumption get closer to the blue and green lines, their growths reduce in each time period, and the economy converges to the steady state $(\hat{k}_{ss}, \hat{c}_{ss})$.

We learned that, if the initial capital is below the steady state capital, then consumption and capital grow and converge to their steady state levels. There's only one possible equilibrium path that satisfy equations (55)-(58).

2.4 Model overview

The next page shows the fundamental equations of the neoclassical growth model. We call these equations fundamental because they are *assumptions* and should be "taken as given" by someone studying the Ramsey model. All other equations we have seen in section 2 are not "fundamental" in the sense that they are not assumptions. Instead, they follow logically from the fundamental equations; they are a *consequence*, a *result*.

For example, an assumption of the Ramsey model is that households make choices to maximize their discounted lifetime utilities. The Euler equation, then, is a result describing that, if households maximize their discounted lifetime utilities, then they want to equalize marginal utilities in different time periods, adjusting for their time discount and the interest rates. If we change the assumption on what households care about, the same Euler equation might not hold anymore.

The next page also lists all exogenous and endogenous variables in the Ramsey model.[21] Knowing the type of each variable (exogenous or endogenous) is extremely important. First, it helps understanding the "mechanics" of the Ramsey model, considering that some variables are a consequence of others. For example, if β (exogenous) changes, the whole sequences of prices in the economy, $\{w_t, r_t, R_t\}_{t=0}^{\infty}$ (endogenous), change.

Knowing the type of each variable in the Ramsey model also helps *thinking critically* about the theory embedded in the model. For example, one nice feature of the Ramsey model is its ability to describe how economic growth is affected by individuals' preferences related to intertemporal choices. This is because the model features two exogenous variables capturing households' preferences related to intertemporal choices (β and θ) and economic growth is endogenous (GDP growth). Conversely, the Ramsey model is not an appropriate tool to study how public policies affect technological innovation because the growth rate of labor productivity in the model, x, is an exogenous variable.

[21]"Exogenous variables" are also called "parameters" in economics.

The Neoclassical Growth Model in Discrete Time

- Households take as given $\{L_t, w_t, r_t\}_{t=0}^{\infty}$ and solve:

$$\max_{\{C_t, A_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t L_t u\left(\frac{C_t}{L_t}\right)$$

subject to

$$C_t + A_{t+1} = L_t w_t + (1 + r_t) A_t \text{ for all } t = 0, 1, \ldots$$

$$A_0 > 0 \text{ given}$$

$$C_t \geq 0,\ A_{t+1} \in \mathbb{R} \text{ for all } t = 0, 1, \ldots$$

$$\lim_{t \to \infty} A_{t+1} \prod_{\tau=1}^{t} (1 + r_\tau)^{-1} \geq 0$$

- Firms: in each t, firms take as given T_t, w_t and R_t, and solve:

$$\max_{K_t^d, L_t^d} (K_t^d)^\alpha (T_t L_t^d)^{1-\alpha} - R_t K_t^d - w_t L_t^d$$

- Exogenous processes (labor and technology):

$$L_{t+1} = (1 + n) L_t \text{ for all } t = 0, 1, \ldots,\ L_0 = 1$$

$$T_{t+1} = (1 + x) T_t \text{ for all } t = 0, 1, \ldots,\ T_0 = 1$$

- Relationship between R_t and r_t:

$$R_t = r_t + \delta$$

- General equilibrium conditions:

$$A_t = K_t^d \text{ for all } t = 0, 1, \ldots$$

$$L_t = L_t^d \text{ for all } t = 0, 1, \ldots$$

- Exogenous variables: $\alpha, \beta, \delta \in (0, 1)$, $n, x \geq 0$, $A_0 > 0$, utility function parameters
- Endogenous variables: $\{w_t, r_t, R_t, C_t, A_{t+1}, K_{t+1}, L_{t+1}, K_t^d, L_t^d, T_{t+1}\}_{t=0}^{\infty}$ (per-capita and per-effective-labor variables are not listed here to save space)

2.5 Mini appendix. Dynamic consistency: Does the household stick to its original plan?

We wrote in section 2.1.1 that the household solves its mathematical problem *in the first time period*, $t = 0$. That means specifically that, at $t = 0$, the household plans how much to consume and save in each time period, $t = 0, 1, \ldots$. Naturally, the household believes in $t = 0$ that, when a given future time period arrives, it will make choices as planned. But will this indeed happen? Does the household have any incentive to deviate from its initial plan?

Economic dynamic problems where the agent has incentives to deviate in the future from its own plan are called dynamically inconsistent. We will show that this is not the case in the Ramsey model. That is, households are dynamically consistent.

Let's think about the household's situation when the second time period, $t = 1$, arrives. The household starts that time period with a given level of assets a_1 chosen it $t = 0$. The problem is

$$U_1(A_1) = \max_{\{C_t, A_{t+1}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} L_t u\left(\frac{C_t}{L_t}\right) \tag{62}$$

subject to

$$C_t + A_{t+1} = L_t w_t + (1 + r_t)A_t \text{ for all } t = 1, 2, \ldots, \tag{63}$$

$$A_1 \text{ given,} \tag{64}$$

$$C_t \geq 0, \; A_{t+1} \in \mathbb{R} \text{ for all } t = 1, 2, \ldots, \tag{65}$$

$$\lim_{t\to\infty} A_{t+1} \prod_{\tau=1}^{t} (1 + r_\tau)^{-1} \geq 0. \tag{66}$$

Two things are worth noting. First, the time discount term is written as β^{t-1} because the objective function in (62) is the discounted lifetime utility measured at $t = 1$. Just like the weight given to utiliy at $t' = 0$ when the household is making choices in $t = 0$ is $\beta^{t'} = \beta^0 = 1$, the weight given to utility at $t' = 1$ when the household is making choices in $t = 1$ is $\beta^{t'-1} = \beta^0 = 1$.

Second, the no-Ponzi condition (66) is the same as before, (18). Recall we have interpreted this condition as an institutional feature where the credit market is able to keep track of all financial operations households make in order to avoid Ponzi schemes. In a given time period $t > 0$, the credit market knows past financial operations done by the household. That's why the no-Ponzi condition doesn't change: because it needs to consider all the interest rates faced by the household throughout its lifetime (past, present and future).

Equation (62) says that $U_1(A_1)$ is the maximum attainable discounted lifetime utility, measured at $t = 1$, subject to all relevant constraints the household faces. $U_1(A_1)$ is a function of A_1 because the household's optimal choices starting at $t = 1$ and going forward in time depend on the level of assets with which it starts at $t = 1$. In summary, the function $U_1(A_1)$ gives the optimal lifetime utility of the household if it chooses to start time period $t = 1$ with A_1 assets.

Now, let's go back to $t = 0$ and consider the following problem:

$$\max_{C_0, A_1} L_0 u\left(\frac{C_0}{L_0}\right) + \beta U(A_1) \tag{67}$$

subject to

$$C_0 + A_1 = L_0 w_0 + (1 + r_0) A_0, \tag{68}$$

$$A_0 \text{ given}, \; C_0 \geq 0, \; A_0 \in \mathbb{R}. \tag{69}$$

In this problem, the household is choosing what to do in $t = 0$ taking as given what it will choose to do in the future as a function of A_1. Here it is clear that the household will make the best possible choice in $t = 0$ to maximize its discounted lifetime utility measured in $t = 0$, $L_0 u(C_0/L_0) + \beta U(a_1)$, and, once $t = 1$ arrives, it will make exactly the same choices it projected to make in the future when choosing in $t = 0$. That is, problem (67)-(69) is dynamically consistent between $t = 0$ and $t = 1$.

Now, note that

$$
\begin{aligned}
& \underbrace{\max_{C_0, A_1} L_0 u(C_0/L_0) + \beta U_1(a_1)}_{\text{subject to constraints in } t=0} \\
= & \underbrace{\max_{C_0, A_1} L_0 u(C_0/L_0) + \beta \underbrace{\max_{\{C_t, A_{t+1}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} L_t u(C_t/L_t)}_{\text{subject to constraints in } t=1,2,\ldots \text{ (function of } A_1)}}_{\text{subject to constraints in } t=0} \\
= & \underbrace{\max_{C_0, A_1} L_0 u(C_0/L_0) + \underbrace{\max_{\{C_t, A_{t+1}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t} L_t u(C_t/L_t)}_{\text{subject to constraints in } t=1,2,\ldots \text{ (function of } A_1)}}_{\text{subject to constraints in } t=0} \\
= & \underbrace{\max_{\{C_t, A_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} L_t u(C_t/L_t)}_{\text{subject to constraints in } t=0,1,\ldots} .
\end{aligned}
$$

The first line of the equation above is the dynamically consistent problem, (67)-(69). The last line of the equation above is the original problem of the Ramsey model, (14)-(15). This shows that household's problem in the Ramsey model is dynamically consistent between $t = 0$ and $t = 1$. That is, the household has no incentive to deviate from its original plan when $t = 1$ is reached. This same argument can be used for showing that the household will not deviate from the plan traced in $t = 0$ in any future time period.

3 The model in continuous time

This section presents the Neoclassical Growth Model in continuous time. In many situations, continuous-time models are more tractable than their discrete-time counterparts for different reasons. For example, differential equations are easier to work with than difference equations. Additionally, in continuous time, the change of variables across time can be described through derivatives, which allows using calculus to study their behavior.

Continuous time means that time is described by a variable in the real positive line, $[0, \infty)$. In section X, we used subscripts to notate a variable's dependence on time (e.g., Y_t is GDP in time period t). In continuous time, we will use parenthesis for this (e.g., $Y(t)$ is GDP in time period t). In the discrete-time model of section X, the term $Y_{1.5}$ doesn't exist, but now our model will allow describing GDP in time period 1.5, $Y(1.5)$.

Since time is continuous, the appropriate notion to denote how a variable changes over time is through its derivative with respect to time. For example, $\partial Y(t)/\partial t$ denotes the change of GDP in time period t. In the discrete-time model of section D, the change of GDP in t is denoted as $Y(t+1) - Y(t)$ (forward change) or $Y(t) - Y(t-1)$ (backward change).[22]

The derivative of a variable with respect to time will be an extensively used mathematical object in this section. Therefore, we introduce the following notation to simplify the exposition. For any variable $X(t)$, $\dot{X}(t)$ denotes the derivative of $X(t)$ with respect to time. That is,

$$\dot{X}(t) \equiv \frac{\partial X(t)}{\partial t}.$$

The model features population and technological growth. $L(t)$ denotes the size of the population in time period t, and we assume that its growth rate is $n \geq 0$ every time period:

$$\frac{\dot{L}(t)}{L(t)} = n \text{ for all } t \geq 0. \tag{1}$$

n is an exogenous variable. Since $\dot{L}(t)$ reflects the change in population size in time period t, $\dot{L}(t)/L(t)$ denotes the growth rate of $L(t)$ or, equivalently, the proportional

[22] Appendix D shows that a continuous-time model can be though of as the limit of a discrete-time model with the duration of each time period tending to zero. Therefore, forward and backward changes in continuous time are equal if the limit giving rise to the derivative is two-sided, or $\frac{\partial Y(t)}{\partial t} = \lim_{\Delta \nearrow 0} \frac{Y(t+\Delta)-Y(t)}{\Delta} = \lim_{\Delta \searrow 0} \frac{Y(t+\Delta)-Y(t)}{\Delta} = \lim_{\Delta \nearrow 0} \frac{Y(t-\Delta)-Y(t)}{\Delta}$. All derivatives in this model are based on two-sided limits.

change in $L(t)$. The initial population size is $L(0) = 1$.

In some parts of the model, it will be useful to work with per-capita variables. Per-capita variables are denoted through lower-case variables and are obtained by dividing a given variable by the population size. For example, GDP per capita in t is given by the total GDP, $Y(t)$, divided by the population, $L(t)$:

$$y(t) \equiv \frac{Y(t)}{L(t)}.$$

Technological growth is modeled through a labor-augmenting productivity $T(t)$. The production function is

$$Y(t) = F(K(t), T(t)L(t)) = K(t)^{\alpha}[L(t)T(t)]^{1-\alpha}. \tag{2}$$

Technology grows by a fixed proportion every time period:

$$\frac{\dot{T}(t)}{T(t)} = x \text{ for all } t \geq 0,$$

where $x \geq 0$ is the exogenous technological growth parameter and the initial technology level is $T(0) = 1$.

(2) says that the total production input generated from workers is given by the number of workers, $L(t)$, times the productivity of each worker, $T(t)$. The term $L(t)T(t)$ is called "effective units of labor" or "effective labor". In some parts of the model, it will be useful to work with variables per effective units of labor. These variables are denoted by lower-case letters with a "hat". For example, GDP per effective units of labor is

$$\hat{y}(t) \equiv \frac{Y(t)}{L(t)T(t)} = \frac{y(t)}{T(t)}.$$

The other aspects of the model are similar to the model in discrete time: there are households and firms making choices to maximize equivalent objective functions and, on top of that, the general equilibrium conditions ensure market clearing in labor and asset markets.

The pace in this section will be faster than in discrete time because many of the concepts were already explained in section D. Thus, if you are starting to read these notes from here, you might want to go back to section D in case you are looking for a more comprehensive explanation of some aspect of the model.

3.1 Households

A representative households takes as given the law of motion of population (1), and the path of wages and interest rates $[w(t), r(t)]_{t=0}^{\infty}$.[23] It solves the following problem:

$$\max_{[C(t)]_{t=0}^{\infty}} \int_0^{\infty} e^{-\rho t} L(t) u\left(\frac{C(t)}{L(t)}\right) dt \tag{3}$$

subject to

$$\dot{A}(t) = w(t)L(t) + r(t)A(t) - C(t), \tag{4}$$

$$A(0) > 0 \text{ given}, \tag{5}$$

$$C(t) \geq 0 \text{ for all } t \geq 0, \tag{6}$$

$$\lim_{t\to\infty} A(t)e^{-\int_0^t r(\tau)d\tau} \geq 0. \tag{7}$$

Choice variables and the objective function Equation (3) shows that the choice variable in the household's maximization problem is consumption in each time period. The objective function (i.e., the function the household maximizes) is the discounted total utility of all household members, measured in $t = 0$. The term $e^{-\rho t}$ denotes time discounting due to impatience. $\rho > 0$ is a parameter that represents impatience. Since ρ is positive, utilities in more distant time periods (relative to $t = 0$) receive less weight in the objective function (i.e., $e^{-\rho t}$ falls as t increases). Besides, a high ρ implies that $e^{-\rho t}$ falls fast as t grows, implying that higher ρ's represent more impatient households. An important property of the time discount term $e^{-\rho t}$ is that every period it falls by the constant rate $-\rho$. That is, defining $D(t) = e^{-\rho t}$, we have that $\dot{D}(t)/D(t) = -\rho$.[24]

$L(t)$ is the number of household members in t. Since all household members are included in the household's objective function, the total utility of the household in a given time period is given by the number of members, $L(t)$, times the utility each of these members have, $u(C(t)/L(t))$. The utility function denotes the happiness that a given household member enjoys, thus it depends on per-capita consumption,

[23] The notation $[x(t)]_{t\geq 0}$ means the set of $x(t)$ for all time periods $t \in (0, \infty]$. Differently from the notation in discrete time (see footnote X), we use brackets instead of curly brackets to indicate that time is continuous.

[24] See appendices E.2 and F.1 for an explanation on why time discounting has this exponential functional form when time is continuous. Appendix F.1 explores in particular the property $\dot{D}(t)/D(t) = -\rho$.

$C(t)/L(t) = c(t)$. At each point in time, all household members consume the same amount.

The budget constraint Equation (4) is the budget constraint. It says that the change in assets in t is given by the total earnings in t (asset earnings, $r(t)A(t)$, and labor earnings, $w(t)L(t)$) minus consumption in t. The total earnings constitute the total inflow in the budget, while consumption comprises the outflow. The change in the household's asset level is given by the difference between inflows and outflows. The budget constraint shows that present consumption has future implications: if consumption is low, assets are accumulated for the future; if consumption is high, the asset level falls and there are less resources for the future.[25]

Bounds and the no-Ponzi condition (5) says that the initial level of assets is taken as given by the household (i.e., the household doesn't choose it).

(6) says that consumption cannot be negative. Similarly to the discrete-time model, a similar constraint doesn't apply to assets. That is, in principle the household can have negative assets, which would be interpreted as debt. Because of this reason, there needs to be a condition that prevents the household from increasing debt indefinitely to maintain higher and higher levels of consumption. This condition is the no-Ponzi condition (7).

As seen in section X, the no-Ponzi condition requires the present-value of assets in the infinitely distant future to be non-negative. The term $e^{-\int_0^t r(\tau)d\tau}$ is the factor to make present value calculations in continuous. Thus, $A(t)e^{-\int_0^t r(\tau)d\tau}$ denotes the present value of assets in time period t, measured in time period zero. Intuitively, a negative present value of assets as time approaches infinity would mean the household accumulates debt forever, allowing it to consume more resources than the amount of resources available for consumption. In other words, the household would be executing a Ponzi scheme. [More to be written here.][26]

[25]See Appendix E.1 for a detailed explanation on why the budget constraint in discrete time (E.1) is written as (4) in continuous time. A quick and heuristic explanation involves rewriting (E.1) as $A_{t+1} - A_t = L_t w_t + r_t A_t - C_t$ and noting that the left-hand side of this equation is the change in assets in a given time period. In continuous time, the change in assets is given by $\dot{A}(t)$, so the budget constraint becomes (4) in continuous time.

[26]See appendices E.3 and F.2 for an explanation why the present-value factor has this exponential functional form in continuous time.

Transforming the problem to per-capita terms Let's rewrite the problem (3)-(7) in terms of per-capita variables. As seen before, the population dynamics are described by

$$\frac{\dot{L}(t)}{L(t)} = n \quad \forall t \geq 0, \quad L(0) = 1.$$

The unique solution to this differential equation is[27]

$$L(t) = e^{nt} \quad \forall t \geq 0.$$

Using this, the household's objective function can be rewritten as

$$\int_0^\infty e^{-\rho t} e^{nt} u(C(t)/L(t)) dt = \int_0^\infty e^{-(\rho - n)t} u(c(t)) dt.$$

Let's now transform the budget constraint to per-capita terms. Start with the definition of assets per capita:

$$a(t) = \frac{A(t)}{L(t)}$$

Differentiating both sides with respect to time,

$$\begin{aligned} \dot{a}(t) &= \frac{\dot{A}(t)L(t) - A(t)\dot{L}(t)}{L(t)^2} \\ &= \frac{r(t)A(t) + w(t)L(t) - C(t)}{L(t)} - a(t)n \\ &= [r(t) - n]a(t) + w(t) - c(t), \end{aligned}$$

where we have used the quotient rule for derivatives in the first equality, and (4) and (1) in the second equality.

Finally, using $A(t) = a(t)L(t)$, note that

$$A(t)e^{-\int_0^t r(\tau)d\tau} = a(t)e^{nt}e^{-\int_0^t r(\tau)d\tau} = a(t)e^{\int_0^t n d\tau}e^{-\int_0^t r(\tau)d\tau} = a(t)e^{-\int_0^t [r(\tau)-n]d\tau}.$$

Therefore, the no-Ponzi condition becomes:

$$\lim_{t\to\infty} a(t)e^{-\int_0^t [r(\tau)-n]d\tau} \geq 0.$$

[27] See appendix D.

The household's problem in terms of per-capita variables is:

$$\max_{[c(t)]_{t=0}^{\infty}} \int_0^{\infty} e^{-(\rho-n)t} u(c(t)) dt \tag{8}$$

subject to

$$\dot{a}(t) = [r(t) - n]a(t) + w(t) - c(t) \tag{9}$$

$$a(0) \text{ given} \tag{10}$$

$$c(t) \geq 0 \text{ for all } t \geq 0 \tag{11}$$

$$\lim_{t\to\infty} a(t) e^{-\int_0^t [r(\tau)-n]d\tau} \geq 0 \tag{12}$$

There are two variables affecting the time discount of the household. On the one hand, the household is impatient and weights future utilities less than present utility. This mechanism is governed by the parameter ρ. On the other hand, since population grows, there are more household members in the future, so the total utility of the household tends to be higher in the future because there are more people enjoying utility from consumption. This mechanism is governed by the population growth rate, n. The net effect of these two forces on time discounting is given by $\rho - n$. We assume that $\rho > n$, so that $-(\rho - n) < 0$ and the household gives more weight to present consumption, compared to future consumption. If condition $\rho > n$ doesn't hold, the weight given to utility in t, $e^{-(\rho-n)t}$, grows with time, and the lifetime discounted utility can be infinite if the consumption path is bounded.[28]

Solving the household's problem Let's obtain the optimality conditions for the maximization problem above. Write the Hamiltonian:[29]

$$\mathcal{H}(t) = e^{-(\rho-n)t} u(c(t)) + \lambda(t)\{[r(t) - n]a(t) + w(t) - c(t)\},$$

where $\lambda(t)$ is the Hamiltonian multiplier associated to constraint (9). The optimality conditions are

$$\frac{\partial \mathcal{H}(t)}{\partial c(t)} = 0,$$

[28]Depending on the sign(s) of the values taken by a given utility function, the lifetime discounted utility can be negative or positive infinity if $\rho < n$.

[29]See Appendix C on how to use the Hamiltonian.

$$\frac{\partial \mathcal{H}(t)}{\partial a(t)} = -\dot{\lambda}(t),$$

and the transversality condition

$$\lim_{t\to\infty} \lambda(t)a(t) = 0. \tag{13}$$

The first optimality condition implies that

$$e^{-(\rho-n)t}u'(c(t)) = \lambda(t). \tag{14}$$

The second implies that

$$-\dot{\lambda}(t) = \lambda(t)[r(t) - n] \Leftrightarrow \frac{\dot{\lambda}(t)}{\lambda(t)} = -[r(t) - n]. \tag{15}$$

Take natural logs of both sides of (14)

$$\ln(\lambda(t)) = -(\rho - n)t + \ln(u'(c(t))).$$

Differentiating both sides with respect to time,

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = -(\rho - n) + \frac{u''(c(t))\dot{c}(t)}{u'(c(t))}.$$

This equation and (15) imply that[30]

$$\frac{\dot{c}(t)}{c(t)} = \left\{-\frac{u''(c(t))c(t)}{u'(c(t))}\right\}^{-1} [r(t) - \rho]. \tag{16}$$

The Euler equation The equation above is the Euler equation. It describes the growth rate of consumption resulting from the optimal choices made by the household.

(16) says that the sign of the change in consumption in a given time period is determined by $r(t)$ and ρ. The interest rate $r(t)$ is related to the benefit of saving (opportunity cost of consuming), while impatience, captured by ρ, is related to the cost of saving (or the relative benefit of present consumption). If $r(t)$ is higher than ρ, it is optimal for the household to delay consumption and save more because the benefit of saving is higher than its cost. This leads to an increase in the growth rate

[30]One can differentiate both sides of (14) without taking logs and use the resulting equation, (15) and (14) again to get to (16).

of consumption.

Why does saving imply in a rise of the consumption growth rate? Intuitively, that's because saving is delaying consumption, which means decreasing current consumption and increasing consumption in the "immediate future". Imagine the current time period is t and the immediate future time period is $t + \Delta$, where Δ is extremely small but strictly positive. The growth rate of consumption in t is $[c(t + \Delta) - c(t)]/c(t)$. Delaying consumption means increasing $c(t + \Delta)$ and decreasing $c(t)$, leading to a higher consumption growth rate in t.[31]

The first term in the right-hand side of (16) is the inverse of the elasticity of the marginal utility with respect to consumption (with a positive sign). We already saw in the discrete-time model that the concavity of the utility function affects the willingness of the household to smooth consumption over time. This is what the first term is capturing: the higher the elasticity of marginal utility with respect to c, the smaller is the change in consumption for a given difference between $r(t)$ and ρ.

The Euler equation simplifies if the utility function is CIES, $u(c) = (c^{1-\theta} - 1)/(1 - \theta)$:

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\theta}[r(t) - \rho]. \tag{17}$$

The interpretation of the equation above is the same as that of its discrete-time version (38).

One example that may help understanding equation (17) is the following. Assume that the path of the interest rate is

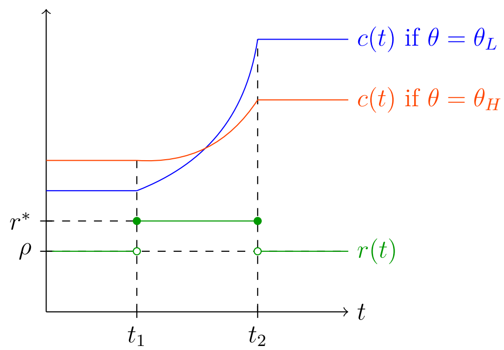
$$r(t) = \begin{cases} \rho & \text{if } t < t_1 \text{ and } t > t_2, \\ r^* & \text{if } t \in [t_1, t_2], \end{cases}$$

where $0 < t_1 < t_2$ and $r^* > \rho$. Let's also say there are two agents, one with a low theta, θ_L, and the other with a high theta, θ_H ($\theta_L < \theta_H$). Except for the difference in their thetas, these two agents are equal with respect to all other characteristics (including their initial levels of assets, wage rate paths, impatience degrees, etc.).

How do the consumption choices of these two agents differ? Since $r(t)$ is equal to ρ before t_1 and after t_2, the Euler equation says that these agents don't change their

[31] Notice that this interpretation is similar to the one in the discrete-time model (equation (35) or ()). With discrete time, the relevant interest rate for consumption choice in t is r_{t+1}. Thus, in our continuous-time interpretation, the interest rate affecting the intertemporal choice of consumption in t versus $t + \Delta$ is $r(t + \Delta)$. Since Δ is extremely small, $r(t + \Delta)$ is approximately equal to $r(t)$.

consumption over time before t_1 and after t_2. Between time periods t_1 and t_2, the interest rate is higher than the time discount parameter, meaning that these agents increase their consumptions in that time interval. (17) implies that the θ_L-agent will choose to increase consumption by a larger magnitude than that of the θ_H-agent:



There are two things worth noting in the figure above. First, in the time interval $[t_1, t_2]$, the consumption evolution of the θ_i-agent ($i \in \{L, H\}$) is described by the differential equation $\dot{c}_i(t)/c_i(t) = (1/\theta_i)(r^* - \rho)$. Since the right hand side of this equation doesn't depend on time, this differential equation says that consumption grows by a constant growth rate. We saw in appendix D that this means that consumption grows exponentially. This is why the curves describing consumption in $[t_1, t_2]$ grow in a convex fashion.

Second, why does the consumption of the θ_L-agent start at a lower level than that of the θ_H-agent? Since we have assumed that both agents are equal in all dimensions, with the exception of their θ's, both agents have the same resources to use during their lifetimes (i.e., initial asset level and wage payments). This means that, since the θ_H-agent chooses to increase consumption less in $[t_1, t_2]$, it can enjoy a higher level of consumption in the initial time periods compared to the θ_L-agent.

This example shows three things. First, in "good" time periods (when the return on assets is high), the agents exploit the gains through increasing their consumptions. Second, the blue curve is "smoother" than the red curve (fewer ups and downs), representing the fact that the θ_H-agent has a higher willingness to smooth consumption over time compared to the θ_L-agent.

Third, in this model there are two distinct aspects describing the preferences of an agent (household) toward consumption: impatience (parameter ρ) and willingness to smooth consumption (concavity of the utility function, or θ), and it might not be trivial to distinguish between these two. A layman observing the graph above could think that the red curve represents an impatient agent, since this agent chooses to consume more in initial time periods and less in distant future time periods, compared to the agent represented by the blue curve. However, since the two agents have the same time discount parameter, ρ, this interpretation is incorrect. What distinguishes the two agents are their willingnesses to smooth consumption over time.

The transversality condition and the no-Ponzi condition We haven't used the TVC (13) so far. Let's use it now. First, (15) is a differential equation whose solution is given by

$$\lambda(t) = \lambda(0)e^{-\int_0^t [r(\tau)-n]d\tau}.$$

Using this and (14) evaluated at $t = 0$ in the TVC (13),

$$\lim_{t\to\infty} a(t)e^{-\int_0^t [r(\tau)-n]d\tau} = 0. \tag{18}$$

...

Compare this equation with the no-Ponzi condition (). The left-hand side of both are the same: the present-value of assets in the long run. The difference between the two is that the no-Ponzi condition is an inequality using a "greater or equal sign", while equation (18), obtained through the TVC, is an equality.

We will see the economic intuition of (18) in the next subsections.

The household's lifetime resources [To be written]

The initial consumption [To be written]

3.2 Firms

A representative firm takes as given the current state of technology, $T(t)$, wage and interest rates $w(t)$ and $r(t)$, and chooses labor and capital to maximize profits:

$$\max_{K^d(t), L^d(t)} K^d(t)^\alpha [T(t)L^d(t)]^{1-\alpha} - w(t)L^d(t) - R(t)K^d(t),$$

where we have used the supperscript "d" to denote "demand". Since there's population growth in this model, this notation is useful to distinguish labor demand, $L^d(t)$, from labor supply (or population size), which is given by $L(t) = e^{nt}$, although both are equal in general equilibrium.

The first-order conditions to the firm's maximization problem are

$$\alpha K^d(t)^{\alpha-1}[T(t)L^d(t)]^{1-\alpha} = R(t),$$

$$(1-\alpha)K^d(t)^{\alpha}T(t)^{1-\alpha}L^d(t)^{-\alpha} = w(t).$$

Rewriting these equations in terms of per effective labor variables:

$$\alpha \hat{k}^d(t)^{\alpha-1} = R(t), \tag{19}$$

$$(1-\alpha)\hat{k}^d(t)^{\alpha}e^{xt} = w(t), \tag{20}$$

where we have used the fact that $T(t) = e^{xt}$.

3.3 General equilibrium

The general equilibrium conditions are:

$$K^d(t) = A(t),$$

$$L^d(t) = L(t).$$

The left-hand (right-hand) side of the first equation is capital demand (supply), and the left-hand (right-hand) side of the second equation is labor demand (supply). The first equation implies that $k^d(t) = a(t)$ and $\hat{k}^d(t) = \hat{a}(t)$.

As discussed in section X, the relationship between the interest rate and the capital rental rate is

$$R(t) = r(t) + \delta. \tag{21}$$

Our objective now is to write the general equilibrium equations only in terms of $\hat{c}$, $\hat{k}$, and exogenous variables.

First, using (17), (21) and (19),

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\theta}\left[\alpha\hat{k}(t)^{\alpha-1} - \delta - \rho\right]. \tag{22}$$

We need to write the left-hand side in terms consumption per effective labor, $\hat{c}$. Note that, since $\hat{c}(t) = c(t)/T(t)$,

$$\hat{c}(t) = c(t)e^{-xt}.$$

Taking logs and differentiating both sides with respect to time,

$$\frac{\dot{\hat{c}}(t)}{\hat{c}(t)} = \frac{\hat{c}(t)}{\hat{c}(t)} - x.$$

Using this in (22),

$$\frac{\dot{\hat{c}}(t)}{\hat{c}(t)} = \frac{1}{\theta}\left[\alpha\hat{k}(t)^{\alpha-1} - \delta - \rho - \theta x\right]. \tag{23}$$

Second, differentiating both sides of $\hat{a}(t) = a(t)e^{-xt}$ with respect to time, we get

$$\dot{\hat{a}}(t) = \dot{a}(t)e^{-xt} - a(t)e^{-xt}x.$$

Using (9), (19), (20), and the definition of per-effective-labor variables in the equation above,

$$\dot{\hat{k}}(t) = \hat{k}(t)^{\alpha} - \hat{c}(t) - (\delta + n + x)\hat{k}(t). \tag{24}$$

Third, use (18), (21), (19) and $a(t) = \hat{k}(t)e^{xt}$ to get to the general-equilibrium version of the transversality condition:

$$\lim_{t\to\infty} \hat{k}(t)e^{-\int_0^t [\alpha\hat{k}(\tau)^{\alpha-1} - \delta - n - x]d\tau} = 0. \tag{25}$$

System of differential equations Equations (23), (24), and (25) give constraints on the path of consumption and capital per effective units of labor in general equilibrium. We also have the fact that the initial level of aggregate capital in the economy, $K(0)$, is an exogenous variable greater than zero. Naturally, this implies that $\hat{k}(0) > 0$ is also parameter of the model because $L(0)$ and $T(0)$ are exogenous (equal to one).

Using these facts, we obtain the system of differential equations describing the general equilibrium of the Ramsey economy. Next, we repeat these four mathematical objects for convenience:

$$\dot{\hat{c}}(t) = \hat{c}(t)\frac{1}{\theta}\left[\alpha\hat{k}(t)^{\alpha-1} - \delta - \rho - \theta x\right], \tag{26}$$

$$\dot{\hat{k}}(t) = \hat{k}(t)^{\alpha} - \hat{c}(t) - (\delta + n + x)\hat{k}(t), \tag{27}$$

$$\hat{k}(0) > 0 \text{ given,} \tag{28}$$

$$\lim_{t\to\infty} \hat{k}(t)e^{-\int_0^t [\alpha\hat{k}(\tau)^{\alpha-1} - \delta - n - x]d\tau} = 0. \tag{29}$$

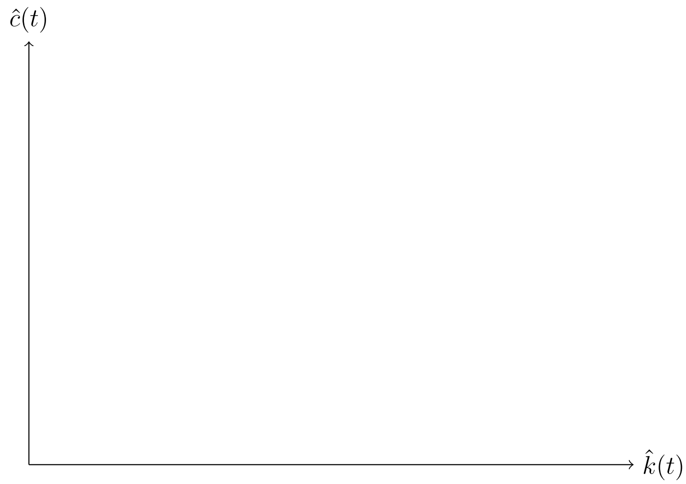
Equations (26) and (27) are differential equations with respect to, respectively, $\hat{c}$

and $\hat{k}$. The first equation describes the change in consumption in a given time period, $\dot{\hat{c}}(t)$, as a function of the levels of consumption and capital in t. The second gives the change in capital, $\dot{\hat{k}}(t)$, as a function of the levels of consumption and capital. That is, the behavior of consumption is described by a differential equation that depends on the level of capital, while capital behavior is determined by another differential equation that depends on the level of consumption. Consumption and capital are interconnected, and the differential equation of one of these variables cannot be solved independently from the other differential equation. Because of this, we say that (26)-(27) constitute a *system of differential equations* in terms of $\hat{c}$ and $\hat{k}$.

Since one independent differential equation needs an initial condition to fully describe the behavior of one variable, the bidimensional system (26)-(27) needs two initial conditions. We know the initial value of capital because that is given to us, (28). If we knew the initial level of consumption, $\hat{c}(0)$, we would have a complete system describing $[\hat{k}(t), \hat{c}(t)]_{t=0}^{\infty}$ in general equilibrium. However, remember that the full path of consumption in this model is an endogenous variable, so the initial condition for consumption is also endogenous.

As seen in X, the transversality condition (29) is a *terminal condition* for capital. As seen in Y, this condition is related to the initial condition for consumption. Therefore, the system (26)-(29) fully describes the path of $\hat{c}$ and $\hat{k}$. The only difference from a standard bidimensional system of differential equations is that, in our case, there is an intial and a terminal condition for one of the variables (capital), while the full path of the other variable (consumption) is free.

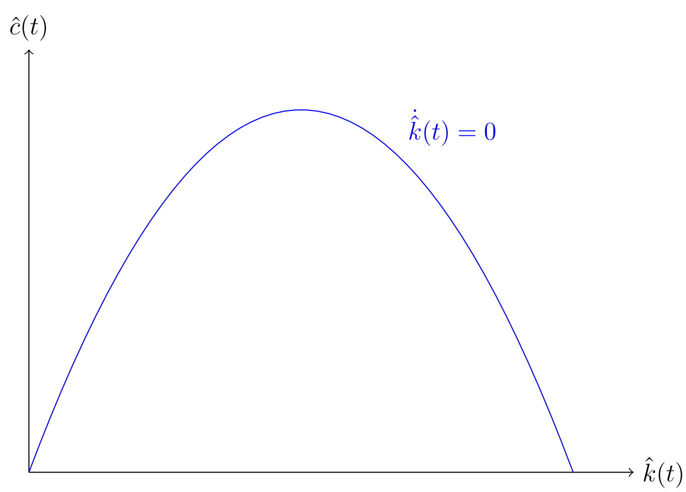
The phase diagram: $\dot{\hat{k}}(t) = 0$ and $\dot{\hat{c}}(t) = 0$ locci To better understand the behavior of the system of differential equations (26)-(29), we will use a graphical tool called *phase diagram*. In our case, the phase diagram is drawn on the $(\hat{k}(t), \hat{c}(t))$ plane:



The first step is to study the pairs of capital and consumption implying that one of these variables (or both) don't move in a given time period. For example, assuming $\dot{\hat{k}}(t) = 0$ in equation (27), we can write the following equation

$$\hat{c}(t) = \hat{k}(t)^{\alpha} - (\delta + n + x)\hat{k}(t). \tag{30}$$

This equation describes a relationship between $\hat{c}$ and $\hat{k}$ ($\hat{c}$ as a function of $\hat{k}$) associated with $\dot{\hat{k}}(t) = 0$. $\hat{c}(t)$ is equal to a concave and increasing function of capital, $\hat{k}(t)^{\alpha}$, minus a linear function of capital, $(\delta+n+x)\hat{k}(t)$. This implies that the pairs of consumption and capital associated with $\dot{\hat{k}}(t) = 0$ are described by an inverse U-shaped curve in the $(\hat{k}(t), \hat{c}(t))$ plane:



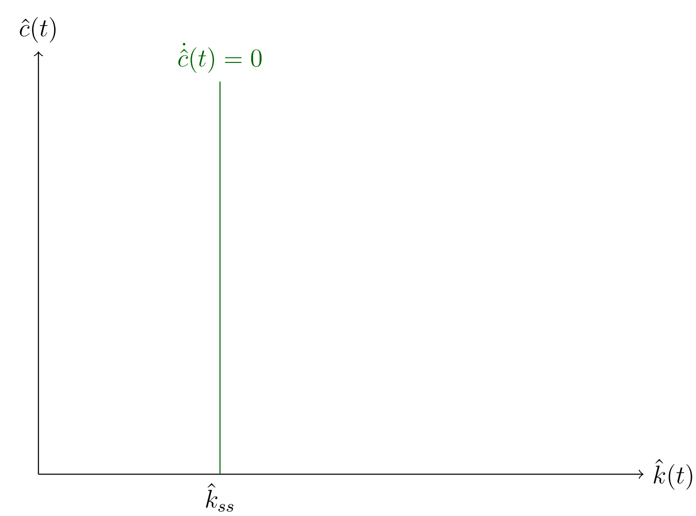
This blue curve is called the "$\dot{\hat{k}}(t) = 0$ locus", and it has the following interpretation: if, at any given point in time, the economy is on the blue curve (meaning that the levels of capital and consumption satisfy equation (30)), then capital doesn't move in that time period ($\dot{\hat{k}}(t) = 0$).

Let's now assume that $\dot{\hat{c}}(t) = 0$ in equation (26):

$$\dot{\hat{c}}(t) = 0 \Leftrightarrow \alpha \hat{k}(t)^{\alpha-1} - \delta - \rho - \theta x = 0 \Leftrightarrow \hat{k}(t) = \left(\frac{\alpha}{\delta + \rho + \theta x}\right)^{\frac{1}{1-\alpha}} \equiv \hat{k}_{ss}. \tag{31}$$

The expression above says that, if consumption doesn't move at a given point in time, then the capital level must be equal to $[\alpha/(\delta + \rho + \theta x)]^{1/(1-\alpha)}$. This capital amount is called the steady state capital, denoted by $\hat{k}_{ss}$, because we will see that this is the capital level to which the economy converges in the long-run. This long-run situation is called the steady state because the variables per effective units of labor ("hat variables", $\hat{k}$, $\hat{c}$, $\hat{y}$) are constant over time (steady).

(31) also says that, if capital is at its steady state level, then consumption doesn't move, *regardless of the level of consumption*. Because of this, the "$\dot{c}(t) = 0$ locus" (the pairs of capital and consumption associated with consumption not moving) is described by a vertical line in the $(\hat{k}(t), \hat{c}(t))$ plane:



The phase diagram: Locci intersection A natural question at this point is: where do the two locci (plural of locus) intersect with each other? Does the vertical line intersect with the blue curve exactly at peak of the blue curve? To the left of the peak? To the right?

Let's first characterize the capital level associated with the peak of the blue curve. The blue curve is described by (30). Therefore, we only need to take the derivative of the right-hand side of (30) and make it equal to zero:

$$\alpha\hat{k}(t)^{\alpha-1} - (\delta + n + x) = 0 \Leftrightarrow \hat{k}(t) = \left(\frac{\alpha}{\delta + n + x}\right)^{\frac{1}{1-\alpha}} \equiv \hat{k}_{\text{gold}}.$$

The capital that maximizes consumption in the $\dot{\hat{k}}(t) = 0$ locus is called the golden rule capital and is denoted by $\hat{k}_{\text{gold}}$.[32]

To find out where the blue curve and green line intersect, we need to use the transversality condition (29). Let's assume that capital converges to its steady state level (we will see later that this is indeed true). Note that this implies something about the TVC:

$$\lim_{t\to\infty} \underbrace{\hat{k}(t)}_{\text{converges to } \hat{k}_{ss}} \underbrace{e^{-\int_0^t [\alpha\hat{k}(\tau)^{\alpha-1} - \delta - n - x]d\tau}}_{\text{needs to converge to zero}} = 0.$$

That is, if the first factor of the multiplication goes to a finite number, then the second factor needs to converge to zero for this limit be equal to zero.

How can the exponential of a function of time go to zero? The answer is that this function of time must tend to minus infinity since that's the only way an exponential can tend to zero. That is, $\lim_{t\to\infty} - \int_0^t [\alpha\hat{k}(\tau)^{\alpha-1} - \delta - n - x]d\tau = -\infty$, or

$$\lim_{t\to\infty} \int_0^t [\alpha\hat{k}(\tau)^{\alpha-1} - \delta - n - x]d\tau = \infty.$$

Since $\hat{k}(t)$ converges to $\hat{k}_{ss}$, we need the following condition to hold for the limit above to be true: $\alpha\hat{k}_{ss}^{\alpha-1} - \delta - n - x > 0$, or[33]

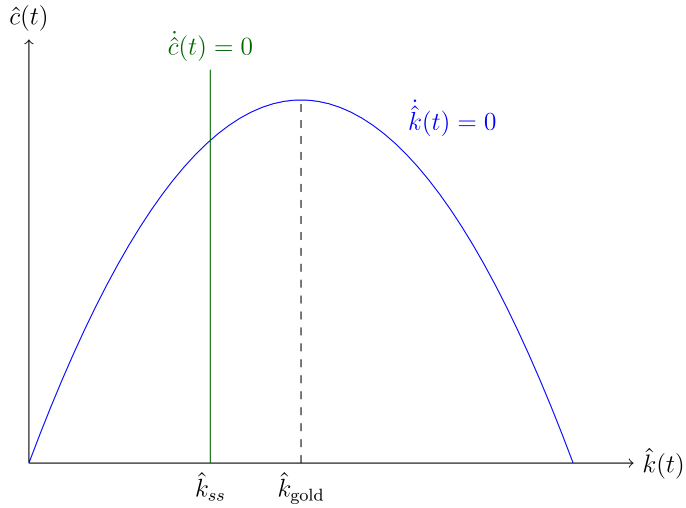
$$\hat{k}_{ss} < \left(\frac{\alpha}{\delta + n + x}\right)^{\frac{1}{1-\alpha}} = \hat{k}_{\text{gold}}.$$

This derivation shows that the steady state capital is below the golden rule capital, which implies that the $\dot{\hat{c}}(t) = 0$ locus intersects with the $\dot{k}(t) = 0$ locus to the left of the peak of the $\dot{k}(t) = 0$ locus:[34]

[32] Maybe something about the golden rule capital here? Mention the Solow model.

[33] Assume by contradiction that $\alpha\hat{k}_{ss}^{\alpha-1} - \delta - n - x < 0$. Then... Now assume that $\alpha\hat{k}_{ss}^{\alpha-1} - \delta - n - x = 0$. In the particular parametrization where $\hat{k}(0) = \hat{k}_{ss}$, the TVC cannot hold.

[34] If one substitutes the expressions for $\hat{k}_{ss}$ and $\hat{k}_{\text{gold}}$ in the inequality $\hat{k}_{ss} < \hat{k}_{\text{gold}}$, one gets the conditions on parameters X. This is a condition that the parameters of the model need to satisfy for equilibrium to exist. The intuition is that...

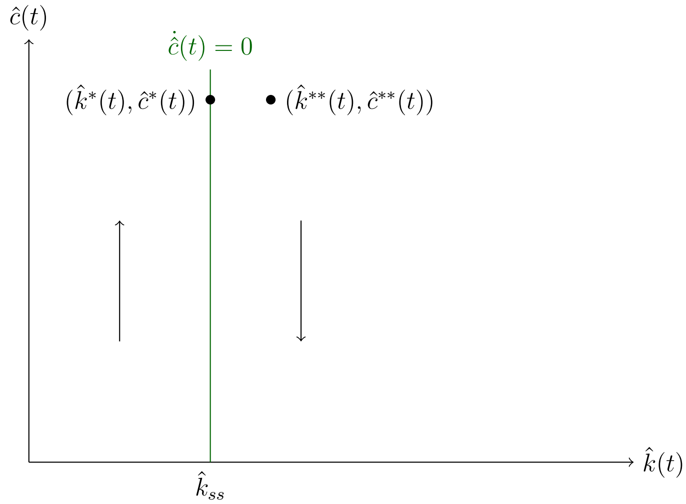


The intersection of the two locci is the steady state. If the economy is on this intersection, both consumption and capital per effective units of labor (and, therefore, output) don't move over time. The steady state consumption is obtained by setting $\dot{\hat{c}}(t) = \dot{\hat{k}}(t) = 0$, or

$$\hat{c}_{ss} = \hat{k}_{ss}^{\alpha} - (\delta + n + x)\hat{k}_{ss}.$$

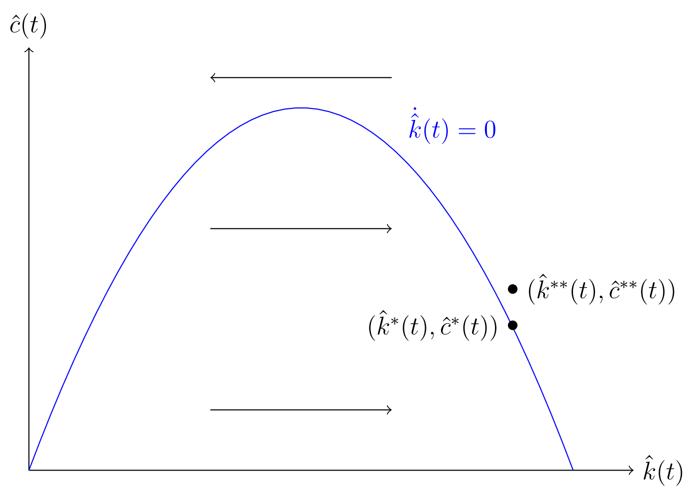
The phase diagram: Movement direction We learned before that if the economy is on a given locus, the variable associated to this locus doesn't change in that particular time period. What if the economy is not on a locus? Let's start with the $\dot{\hat{c}}(t) = 0$ locus. Take any $\hat{c}^*(t) > 0$ and a $\hat{k}^*(t)$ such that $\hat{k}^*(t) = \hat{k}_{ss}$. We know that $\dot{\hat{c}}^*(t) = \hat{c}^*(t)\frac{1}{\theta}\left[\alpha\hat{k}^*(t)^{\alpha-1} - \delta - \rho - \theta x\right] = 0$. What happens if we pick another pair with the same level of consumption but with a higher capital? That is, pick $\hat{c}^{**}(t)$ and $\hat{k}^{**}(t)$ with $\hat{c}^{**}(t) = \hat{c}^*(t)$ and $\hat{k}^{**}(t) > \hat{k}^*(t)$. Since $\alpha - 1 < 0$, we have that $\dot{\hat{c}}^{**}(t) = \hat{c}^{**}(t)\frac{1}{\theta}\left[\alpha\hat{k}^{**}(t)^{\alpha-1} - \delta - \rho - \theta x\right] < 0$. This shows that if capital is higher than the steady state capital (i.e., if the economy is to the right of the $\dot{\hat{c}}(t) = 0$ locus), consumption falls.

This can be represented in the phase diagram through arrows:

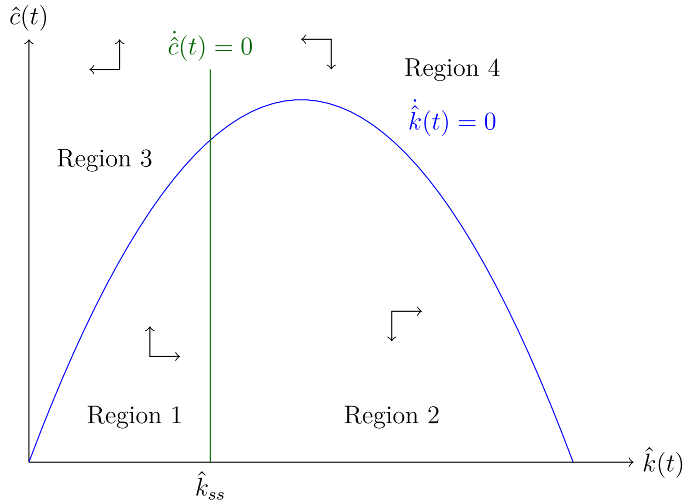


A vertical arrow pointing upward means that, if at a given point in time t the capital and consumption levels in the economy are such that $(\hat{k}(t), \hat{c}(t))$ is to the left of the vertical line, then consumption grows at t. The downward arrow has the opposite interpretation.

Let's follow the same logic now using the $\dot{\hat{k}}(t) = 0$ locus. Pick a pair $(\hat{k}^*(t), \hat{c}^*(t))$ on the $\dot{\hat{k}}(t) = 0$ locus. This means that $\dot{\hat{k}}^*(t) = \hat{k}^*(t)^\alpha - \hat{c}^*(t) - (\delta + n + x)\hat{k}^*(t) = 0$. Now increase $\hat{c}$ keeping $\hat{k}$ fixed. That is, pick $\hat{k}^{**}(t) = \hat{k}^*(t)$ and $\hat{c}^{**}(t) > \hat{c}^*(t)$. We have $\dot{\hat{k}}^{**}(t) = \hat{k}^{**}(t)^\alpha - \hat{c}^{**}(t) - (\delta + n + x)\hat{k}^{**}(t) < 0$ because there's a negative sign before $\hat{c}$ in (27). We conclude that, if the economy is above the $\dot{\hat{k}}(t) = 0$ locus, capital falls. The opposite holds for points below the $\dot{\hat{k}}(t) = 0$ locus. We can represent this graphically as:



We can draw a phase diagram with vertical and horizontal arrows simultaneously:



It's useful to give names to some areas in the phase diagram. We will call the area below the blue curve and to the left of the vertical line region 1. Regions 2, 3, and 4 are also shown in the figure above and, similarly to region 1, are defined in terms of their positions relative to the blue curve (below or above) and the green line (to the left or to the right). The arrows show the direction to which variables move in the four regions. For example, in region 2, capital grows (horizontal arrow pointing to the right) and consumption falls (vertical arrow pointing downward).

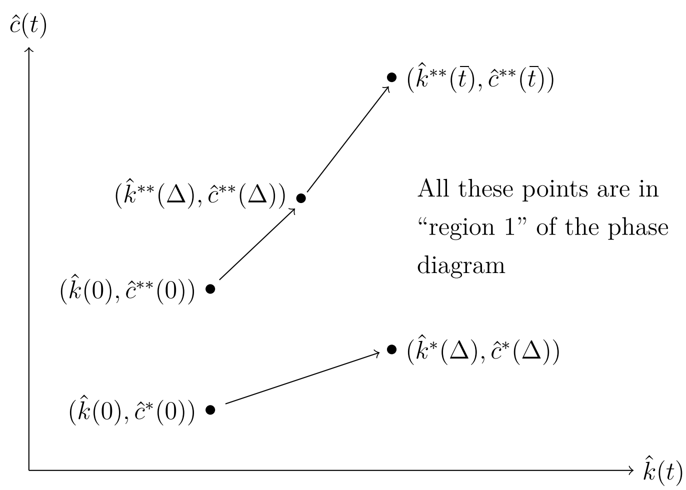
The phase diagram: Equilibrium What we learned before simply gives us information about what would happen if the economy happens to be at a given point in the phase diagram. A natural question is: in which points of the phase diagram can the economy actually be and still satisfy the system of differential equations (26)-(29)? That is, where in the phase diagram is the economy *in equilibrium*?

We can first show that the economy in equilibrium cannot be in regions 2 or 3.

First, assume that the economy is in region 2. Capital increases and consumption falls continuously. Capital will converge to the capital level associated with the point where the $\dot{\hat{k}}(t) = 0$ locus touches the horizontal axis to the right of the vertical line. This capital level is higher than the golden rule capital. Recall we showed before that the level to which capital converges should be lower than the golden rule capital for the TVC to hold. Therefore, the economy being in region 2 would violate the TVC.

Now assume that the economy is in region 3. Initially, capital falls and consumption grows continuously. Over time, the economy gets farther from the vertical line, implying that capital falls by bigger magnitudes each time. This implies that the economy gets to the situation with zero capital in finite time. If there's no capital, output is zero, implying that consumption must also be zero. Therefore, the economy discontinuously jumps from a point with positive consumption to a situation with zero consumption, and this violates the differential equation describing consumption changes (26). This shows that the economy cannot be in region 3.

We will see one more important property of this phase digram. Take two levels of initial consumption in region 1, $c^*(0)$ and $c^{**}(0)$, where $c^{**}(0) > c^*(0)$, and both consumption levels are positive but very small.[35] Let's say that Δ is a very small time length, and we want to understand where the economy will be in the phase diagram in time period Δ depending on the initial consumption. With the initial consumption $\hat{c}^*(0)$, since the economy is in region 1, both variables will increase, so let's say that at time period $t = \Delta$ the economy is at point $(\hat{k}^*(\Delta), \hat{c}^*(\Delta))$ as described in the figure below. Now let's say that the initial consumption is $\hat{c}^{**}(0)$. From equation (26) [(27)], we know that consumption [capital] grows more [less] than in the case with initial consumption $c^*(0)$ because the initial consumption is now higher and the capital level is the same. Therefore, in $t = \Delta$, the economy will be at a point $(\hat{k}^{**}(\Delta), \hat{c}^{**}(\Delta))$ where $\hat{c}^{**}(\Delta) > \hat{c}^*(\Delta)$ and $\hat{k}^{**}(\Delta) < \hat{k}^*(\Delta)$. Since the economy is still in region 1, both variables will grow, so the economy will reach the capital level $\hat{k}^*(\Delta)$ in some time period after Δ, say $\bar{t}$, and with a higher level of consumption, $(\hat{k}^{**}(\bar{t}), \hat{c}^{**}(\bar{t}))$.



The reasoning above shows that, the higher the initial consumption, the higher will be the consumption level corresponding to a capital level $\hat{k}(t) > \hat{k}(0)$ if the economy doesn't leave region 1.[36]

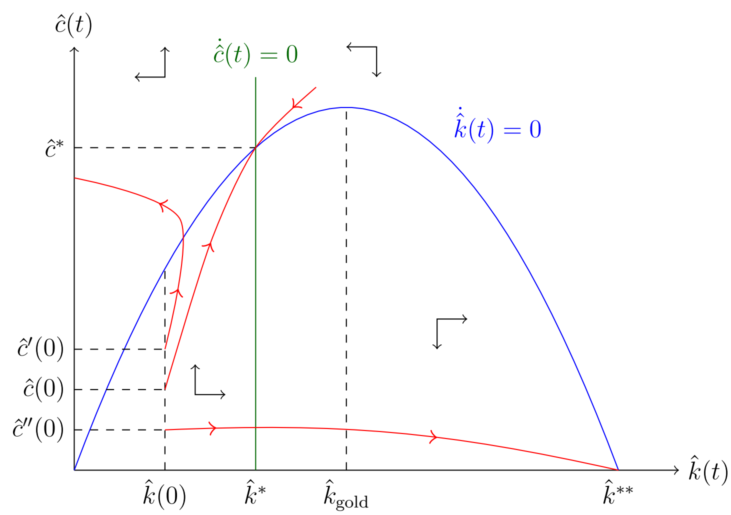
[35] See the next footnote to see why we require these consumption levels to be "small".

[36] Note that, as $\hat{c}^*(0)$ tends to zero, the economy path tends to a situation where capital grows

The reasoning above shows that, if the initial consumption is very low, the economy will follow a path where $\hat{c}$ and $\hat{k}$ grow until the vertical line is reached with a consumption level lower than the steady state consumption. Since the economy is below the $\dot{\hat{k}}(t) = 0$ locus at this point, capital grows and the economy enters region 2, which cannot happen in equilibrium.

As we increase the initial consumption continuously, the consumption level when $\hat{k}_{ss}$ is reached increases continuously because equations (26) and (27) are continuous. Thus, there's only one consumption level that makes the economy converge to the steady state $(\hat{k}_{ss}, \hat{c}_{ss})$. Let's say that $\hat{c}(0)$ is an initial consumption such that the economy converges to the steady state. If we increase $\hat{c}(0)$ infinitesimally, the economy will follow a new path where consumption (capital) will grow faster (slower), and therefore the economy will hit the $\dot{\hat{k}}(t) = 0$ locus with a consumption level lower than $\hat{c}_{ss}$. Consumption will continue growing and the economy will enter region 3.

The figure below shows the only initial consumption compatible with the equilibrium conditions, $\hat{c}(0)$, and two consumption levels that cannot happen in equilibrium, $\hat{c}'(0)$ (too high) and $\hat{c}''(0)$ (too low).



We learned that, if the initial capital is below the steady state capital, then consumption and capital grow and converge to their steady state levels. There's only one possible equilibrium path that satisfy equations (26)-(29).

and consumption stays constant at zero. In this limit case, the economy hits the vertical line with $\hat{c}(t) = 0$. Since the differential equations (26) and (27) are continuous, we have the following result: for any capital level $\hat{k}$ smaller than $\hat{k}_{ss}$, there's a (small) initial consumption such that the economy hits $\hat{k}$ in region 1. This shows that the two initial consumption levels we picked some paragraphs above, $c^*(0)$ and $c^{**}(0)$, exist.

[More to be written here]

3.4 Model overview

If you haven't done so yet, please read the beginning of Section D to understand the purpose of this "Model overview" section.

The Neoclassical Growth Model in Continuous Time

- Households take as given $[L(t), w(t), r(t)]_{t=0}^{\infty}$ and solve:

$$\max_{[C(t)]_{t=0}^{\infty}} \int_0^{\infty} e^{-\rho t} L(t) u\left(\frac{C(t)}{L(t)}\right) dt$$

subject to

$$\dot{A}(t) = w(t)L(t) + r(t)A(t) - C(t)$$

$$A(0) > 0 \text{ given}$$

$$C(t) \geq 0 \text{ for all } t \geq 0$$

$$\lim_{t\to\infty} A(t) e^{-\int_0^t r(\tau)d\tau} \geq 0$$

- Firms: in each t, firms take as given $T(t)$, $w(t)$ and $R(t)$, and solve:

$$\max_{K^d(t), L^d(t)} K^d(t)^{\alpha}[T(t)L^d(t)]^{1-\alpha} - w(t)L^d(t) - R(t)K^d(t)$$

- Exogenous processes (labor and technology):

$$\frac{\dot{L}(t)}{L(t)} = n \text{ for all } t \geq 0,\ L(0) = 1, \quad \frac{\dot{T}(t)}{T(t)} = x \text{ for all } t \geq 0,\ T(0) = 1$$

- Relationship between $R(t)$ and $r(t)$:

$$R(t) = r(t) + \delta$$

- General equilibrium conditions:

$$A(t) = K^d(t) \text{ for all } t \geq 0$$

$$L(t) = L^d(t) \text{ for all } t \geq 0$$

- Exogenous variables: $\alpha, \delta \in (0,1)$, $\rho > 0$, $n, x \geq 0$, $A(0) > 0$, utility function parameters

- Endogenous variables: $[w(t), r(t), R(t), C(t), A(t), K(t), L(t), K^d(t), L^d(t), T(t)]_{t=0}^{\infty}$ (per-capita and per-effective-labor variables are not listed here to save space)

3.5 Exercises

1. Planner's problem

Appendix A More exercises

1. The CIES utility function is

$$u(c) = \frac{c^{1-\theta} - 1}{1 - \theta},$$

where $\theta > 0$ is a parameter measuring the concavity of the utility function. Show that

$$\lim_{\theta \to 1} \frac{c^{1-\theta} - 1}{1 - \theta} = \ln(c).$$

Hint: use L'Hôpital's rule.

2. Prove the *Euler's theorem*: if $f(x, y)$ is homogeneous of degree $m \in \mathbb{N}$, or

$$f(\lambda x, \lambda y) = \lambda^m f(x, y) \quad \forall \lambda > 0,$$

then

$$mf(x, y) = f_1(x, y)x + f_2(x, y)y,$$

where $f_1(x, y) \equiv \partial f(x, y)/\partial x$ and $f_2(x, y) \equiv \partial f(x, y)/\partial y$. Furthermore, $f_1(x, y)$ and $f_2(x, y)$ are homogeneous of degree $m - 1$ in x and y.

3. Show that the Cobb-Douglas is the only constant-returns-to-scale production function with constant factor shares in a competitive setup.[37] That is, say $F(K, TL)$ is a production function with constant returns to scale, the firm takes as given w, R, and T and maximizes profits,

$$\max_{K,L} F(K, TL) - wL - RK.$$

Prove that

$$\frac{RK}{F(K, TL)} = \alpha \quad \text{and} \quad \frac{wL}{F(K, TL)} = 1 - \alpha, \tag{1}$$

where $\alpha \in (0, 1)$ doesn't depend on K, L, R, w, or T, if and only if

$$F(K, TL) = AK^\alpha (TL)^{1-\alpha}, \tag{2}$$

where A is a positive number.

[37] Based on Eric Roca's teaching notes (https://eric-roca.github.io/courses/mathematical_appendix/elasticity_of_substitution/).

Solution: proving that (1) implies (2). First, define $f(\hat{k}) \equiv F(\hat{k}, 1) = F(K/(TL), TL/(TL)) = F(K, TL)/(TL)$. From Euler's theorem, $F_1(K, TL) \equiv \partial F(K, TL)/\partial K$ is homogeneous of degree zero, so

$$F_1(K, TL) = F_1\left(\frac{K}{TL}, \frac{TL}{TL}\right) = F_1(\hat{k}, 1) = f'(\hat{k}). \tag{3}$$

From the firm's FOC,

$$F_1(K, TL) = R.$$

Multiplying both sides by $K/F(K, TL)$ and using (1),

$$\frac{F_1(K, TL)K}{F(K, TL)} = \alpha.$$

Using (3) and dividing the numerator and denominator in the left-hand side by TL,

$$\frac{f'(\hat{k})\hat{k}}{f(\hat{k})} = \alpha \Rightarrow \frac{f'(\hat{k})}{f(\hat{k})} = \frac{\alpha}{\hat{k}} \Rightarrow \frac{\partial \ln(f(\hat{k}))}{\partial \hat{k}} = \frac{\alpha}{\hat{k}}.$$

This is a differential equation, with the unknown being the function $f(\hat{k})$. Let's solve it. Integrating both sides with respect to $\hat{k}$ (indefinite integral),

$$\int \frac{\partial \ln(f(\hat{k}))}{\partial \hat{k}} d\hat{k} = \int \frac{\alpha}{\hat{k}} d\hat{k},$$

implying that

$$\ln(f(\hat{k})) = \alpha \ln(\hat{k}) + C,$$

where C is an integration constant. Isolating $f(\hat{k})$,

$$f(\hat{k}) = e^C \hat{k}^\alpha.$$

Multiplying both sides by TL,

$$F(K, TL) = e^C K^\alpha (TL)^{1-\alpha}.$$

Proving that (2) implies (1). TBW.

Appendix B Lagrangian: Applications and insights

Let's say we need to maximize a function $f(x)$, where $x = (x_1, x_2, \ldots, x_K)$ is a vector and $f(x)$ is a real number. We cannot choose any x because we face some constraints. The problem can be generically written as

$$\max_x f(x) \tag{1}$$

subject to

$$g_i(x) = 0 \quad \forall i = 1, \ldots, I \tag{2}$$

$$h_j(x) \geq 0 \quad \forall j = 1, \ldots, J. \tag{3}$$

The problem above is called a "constrained problem" because it has some constraints (the problem is "subject to" some constraints). $f(x)$ is called the "objective function" (the function we want to maximize). (2) says there are I equality constraints. (3) says there are J inequality constraints. Below you will see examples of problems that fit the description above.

To solve this problem, we write the Lagrangian as

$$\mathcal{L} = f(x) + \sum_{i=1}^{I} \lambda_i g_i(x) + \sum_{j=1}^{J} \mu_j h_j(x).$$

If x^* is the solution to the problem, it must satisfy the following optimality conditions (Kuhn-Tucker conditions):

1. First-order conditions (FOCs):

$$\frac{\partial \mathcal{L}}{\partial x_k^*} = 0 \quad \forall k = 1, \ldots, K.$$

2. Slackness conditions:

$$\mu_j h_j(x^*) = 0 \quad \forall j = 1, \ldots, J.$$

3. Sign of inequality multipliers:

$$\mu_j \geq 0 \quad \forall j = 1, \ldots, J.$$

4. Constraints:

$$g_i(x^*) = 0 \quad \forall i = 1, \dots, I$$

$$h_j(x^*) \geq 0 \quad \forall j = 1, \dots, J.$$

The slackness condition implies that, at the optimum x^*, if a given inequality constraint j doesn't hold with equality, then the multiplier associated to it must be zero, $\mu_j = 0$.

Most of the problems we study in Economics have inequality constraints (e.g. consumption cannot be negative), but most of the times we can ignore these constraints because the solutions to these problems never make the inequality constraint hold with equality. In such cases, those constraints are irrelevant. However, there are problems where the inequality constraints are extremely important and should be taken seriously. The second example below shows a problem where inequality constraints matter.

B.1 Example 1: equality constraints

Applying the Lagrangian A consumer with Cobb-Douglas utility in log chooses how much to consume out of two goods, x_1 and x_2, taking as given their prices, p_1 and p_2, and its income E. The problem is:

$$\max_{x_1, x_2 \geq 0} \ln(x_1^{\alpha} x_2^{1-\alpha}) \text{ subject to } p_1 x_1 + p_2 x_2 = E. \tag{4}$$

The consumer will never choose to consume zero out of a given good: if she does that, her utility would go to $-\infty$; which is lower than her utility in the case where she consumes a positive amount of both goods. We can then ignore the inequality constraints $x_1, x_2 \geq 0$ because none of them can hold with equality in the optimal solution.

Thus, we can think of this problem as not having inequality constraints ($J = 0$), and having one equality constraint ($I = 1$), which could be written as

$$g_1(x) = E - p_1 x_1 - p_2 x_2.$$

Also, we have $K = 2$, meaning that $x = (x_1, x_2)$ is a two-dimensional vector.

Using $\ln(x_1^\alpha x_2^{1-\alpha}) = \alpha \ln(x_1) + (1-\alpha)\ln(x_2)$, we write the Lagrangian as

$$\mathcal{L} = \alpha \ln(x_1) + (1-\alpha)\ln(x_2) + \lambda(E - p_1 x_1 - p_2 x_2). \tag{5}$$

The first-order conditions are:

$$\frac{\alpha}{x_1} = \lambda p_1, \tag{6}$$

$$\frac{1-\alpha}{x_2} = \lambda p_2. \tag{7}$$

The last optimality condition is the budget constraint

$$p_1 x_1 + p_2 x_2 = E. \tag{8}$$

Dividing (6) by (7), we get

$$\frac{\alpha/x_1}{(1-\alpha)/x_2} = \frac{p_1}{p_2} \iff \frac{\alpha}{1-\alpha}\frac{x_2}{x_1} = \frac{p_1}{p_2} \iff x_2 = x_1 \frac{p_1}{p_2}\frac{1-\alpha}{\alpha}.$$

Inserting x_2 above into the budget constraint and solving for x_1, we get

$$x_1 = \frac{\alpha E}{p_1}.$$

Using this in the budget constraint again,

$$x_2 = \frac{(1-\alpha)E}{p_2}.$$

These are the well-known demand functions of a consumer with a Cobb-Douglas utility function.

Understanding the Lagrangian Let's take a step back to appreciate what the Lagrangian is doing. We have transformed the constrained problem (4) into an unconstrained problem where we only need to maximize the Lagrangian (5). The objective function of the unconstrained problem is that of the constrained problem plus a variable denoted by λ times the budget constraint represented by the term $E - p_1 x_1 - p_2 x_2$. What is λ? What is it doing?

The unconstrained problem generates a system of three equations and three unknowns, (6), (7), and (8). We repeat this system next for convenience:

$$\begin{cases} \dfrac{\alpha}{x_1} = \lambda p_1 \\ \dfrac{1-\alpha}{x_2} = \lambda p_2 \\ p_1 x_1 + p_2 x_2 = E \end{cases}$$

The three unknowns in the system above are x_1, x_2, and λ. That is, when we write the Lagrangian, we introduce a new variable, λ. However, at first, we don't know the value of this variable. The idea is that the value of λ is the one that makes the choice variables x_1 and x_2 satisfy the third equation of the system (the equality constraint).

This works because λ functions as a weight to penalize or reward the unconstrained objective function. To see this, imagine $p_1 = 1$, $p_2 = 2$, and the income is $E = 100$ euros. If $\lambda = 100,000$, the Lagrangian is

$$\begin{aligned} \mathcal{L} &= \alpha \ln(x_1) + (1-\alpha)\ln(x_2) + (100,000)(100 - x_1 - 2x_2). \\ &= \alpha \ln(x_1) + (1-\alpha)\ln(x_2) + (100,000)(100) - (100,000)x_1 - (100,000)(2)x_2. \end{aligned}$$

Look at this expression and think about how it changes as we increase the amount of a given good, say x_1. Increasing x_1 leads to a higher $\alpha \ln(x_1)$, capturing the utility gain from increasing consumption. However, there's a big cost of increasing x_1, represented by the term $-(100,000)x_1$. That is, if λ is very high, the objective function is being penalized by a large amount for each increase in x_1.

Compare this to a Lagrangian with $\lambda = 00000.1$. Now the penalization for each increase in x_1 is much lower. If $\lambda < 0$, the variable λ is rewarding increases in x_1.

What is the "correct" value of λ? It is the one that makes the budget constraint hold with equality! That is, we need to penalize/reward the unconstrained objective function exactly in a magnitude that makes x_1 and x_2 satisfy the budget constraint. This is what the Lagrangian does.

The FOCs can be rewritten as

$$\underbrace{\frac{\alpha}{x_1}}_{\text{Mg. benefit of increasing } x_1} - \underbrace{\lambda p_1}_{\text{Mg. cost of increasing } x_1} = 0$$

$$\underbrace{\frac{1-\alpha}{x_2}}_{\text{Mg. benefit of increasing } x_2} - \underbrace{\lambda p_2}_{\text{Mg. cost of increasing } p_2} = 0.$$

Increasing x_1 and x_2 has marginal benefits and costs. The marginal benefit is related to utility gains, while the marginal costs are related to how increasing the quantity of a given good affects the budget constraint. The marginal cost has two dimensions. First, it is related to the sensitivity of the budget to a given good (i.e., this good's price). The second part of the marginal cost is the variable λ, which *scales the marginal cost* by the proportion necessary to make the budget constraint hold with equality.

We saw before that the solution to the problem is

$$x_1 = \frac{\alpha E}{p_1} \text{ and } x_2 = \frac{(1-\alpha)E}{p_2}.$$

Plugging the solution into the FOC (6) or (7), we find that

$$\lambda = \frac{1}{E}.$$

That is, the Lagrange multiplier λ is the inverse of the consumer's income. This makes sense: if the income is very high, we don't need to penalize the unconstrained objective function much, and vice-versa.

Understanding the Lagrangian (2): going further Let's try something unusual now. What happens if we assume that λ has a different value, say $\tilde{\lambda} = 2\lambda = 2/E$? If we use this new λ at the FOCs (6) and (7), we get

$$x_1 = \frac{1}{2}\frac{\alpha E}{p_1}, \quad x_2 = \frac{1}{2}\frac{(1-\alpha)E}{p_2}.$$

The consumer is choosing half of what it chooses if $\lambda = 1/E$. How much do these choices of x_1 and x_2 imply in terms of expenditures?

$$p_1 x_1 + p_2 x_2 = p_1 \frac{1}{2}\frac{\alpha E}{p_1} + p_2 \frac{1}{2}\frac{(1-\alpha)E}{p_2} = \frac{E}{2}.$$

These choices imply that the household is only spending half of its income. The idea is that the new value of the Lagrange multiplier is penalizing the unconstrained objective function too much (two times more), so the consumer is not choosing to use all of her income.

A last approach to visualize the workings of the Lagrangian is to maximize the unconstrained objective function assuming the "correct" value for λ ($\lambda = 1/E$):

$$\mathcal{L} = \alpha \ln(x_1) + (1-\alpha)\ln(x_2) + \frac{1}{E}(E - p_1 x_1 - p_2 x_2).$$

Now, the FOCs are sufficient to obtain the optimal choices:

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0 \Rightarrow \frac{\alpha}{x_1} - \frac{p_1}{E} = 0 \Rightarrow x_1 = \frac{\alpha E}{p_1},$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 0 \Rightarrow \frac{1-\alpha}{x_2} - \frac{p_2}{E} = 0 \Rightarrow x_2 = \frac{(1-\alpha)E}{p_2}.$$

B.2 Example 2: inequality constraints

Applying the Lagrangian Let's slightly modify the economic problem of the previous section. A consumer has (a variant of) Stone-Geary preferences for two goods x_1 and x_2, given by

$$u(x_1, x_2) = \ln[x_1^\alpha (c + x_2)^{1-\alpha}],$$

where $c \geq 0$ is a parameter. The consumer takes as given the prices of the two goods, p_1 and p_2, and its income E, and chooses x_1 and x_2 to maximize its utility. One important constraint is that x_1 and x_2 cannot be negative. The problem can be written as:

$$\max_{x_1, x_2} \ln[x_1^\alpha (c + x_2)^{1-\alpha}] \text{ subject to } p_1 x_1 + p_2 x_2 = E,\ x_1, x_2 \geq 0. \tag{9}$$

The new aspect compared to the previous section is the parameter $c \geq 0$. One way to interpret this parameter is that c is some quantity of the second good that the consumer gets for free. An important point of this section is that the consumer might optimally choose to make $x_2 = 0$. Think about the extreme case where E is extremely low (the consumer is very poor) and c is extremely high (the consumer gets a lot of the second good for free): the consumer will choose to spend all its income on x_1. That is, in this problem, depending on the parameters (α, c, p_1, p_2, E), there can be a corner solution where $x_2 = 0$.

Since this problem might feature a corner solution for x_2, the Lagrangian needs to take into account the constraint $x_2 \geq 0$. There's no need to worry about the constraint

$x_1 \geq 0$ because the consumer will never choose to make $x_1 = 0$.

Writing this problem using the format of the problem described in (1), we have $K = 2$ (x is a bidimensional vector composed of x_1 and x_2), $I = 1$ (one equality constraint – the budget constraint), $J = 1$ (one inequality constraint, which is $x_2 \geq 0$),

$$g_1(x) = E - p_1 x_1 - p_2 x_2$$

and

$$h_1(x) = x_2.$$

The Lagrangian can be written as

$$\mathcal{L} = \alpha \ln(x_1) + (1-\alpha)\ln(c + x_2) + \lambda(E - p_1 x_1 - p_2 x_2) + \mu x_2. \tag{10}$$

The optimality conditions are two first-order conditions (FOCs),

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0 \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial x_2} = 0, \tag{11}$$

the slackness condition,

$$\mu x_2 = 0, \tag{12}$$

and the constraint that the inequality multiplier cannot be negative,

$$\mu \geq 0. \tag{13}$$

The FOCs are given by:

$$\frac{\alpha}{x_1} - \lambda p_1 = 0,$$

$$\frac{1-\alpha}{c + x_2} - \lambda p_2 + \mu = 0.$$

Let's eliminate λ from the two equations above. Isolate λ and use the two resulting equations to get

$$\frac{\alpha}{x_1}\frac{p_2}{p_1} = \frac{1-\alpha}{c + x_2} + \mu. \tag{14}$$

Now, to solve this problem, we need to split the solution in two cases.

Case 1. First, let's assume that $x_2 > 0$. From the slackness condition, $\mu = 0$. Using this in (14),

$$x_1 = \frac{\alpha}{1-\alpha}\frac{p_2}{p_1}(c + x_2), \tag{15}$$

or

$$x_2 = \frac{1-\alpha}{\alpha}\frac{p_1}{p_2}x_1 - c. \tag{16}$$

Substituting (16) in the budget constraint,

$$x_1 = \frac{E\alpha}{p_1} + \alpha\frac{p_2}{p_1}c. \tag{17}$$

Substituting (15) in the budget constraint,

$$x_2 = \frac{E(1-\alpha)}{p_2} - c\alpha. \tag{18}$$

These are the solutions for x_1 and x_2 if x_2 is strictly positive. The negative sign in the solution for c_2 above should raise your eyebrows. If c, α, p_2 is high and/or E is low, the equation above might imply a negative c_2. The interpretation for a negative c_2 is that the consumer would want to sell the second good. The economic intuition is clear: if the consumer gets a lot of good two for free (high c), or if the consumer likes good one much more than good two (high α), or if the second good is too expensive (high p_2), or if the consumer is very poor (low E), the consumer would like to sell the second good to buy the first good.

Remember the constraint $x_2 \geq 0$. The consumer is not allowed to sell the second good. Let's check the condition for x_2 being strictly positive (recall we assume $x_2 > 0$ in case 1). Using (18),

$$x_2 > 0 \iff \frac{E(1-\alpha)}{p_2} > c\alpha \iff E(1-\alpha) > c\alpha p_2. \tag{19}$$

Ok. Let's keep all that in mind and go to case 2.

Case 2. Let's now assume that $x_2 = 0$. In this case, the consumer uses all its income to buy the first good, so $x_1 = E/p_1$. Use (14) with $x_1 = E/p_1$ and $x_2 = 0$ to solve for μ:

$$\mu = \frac{\alpha p_2}{E} - \frac{1-\alpha}{c}. \tag{20}$$

Remember that one optimality condition is that μ cannot be negative. Let's check the condition for this to be true:

$$\mu \geq 0 \iff \frac{\alpha p_2}{E} \geq \frac{1-\alpha}{c} \iff c\alpha p_2 \geq E(1-\alpha).$$

This inequality is the exact opposite of inequality (19).

We conclude that the format of the solution to this problem depends on a condition related to the parameters of the model. If $E(1-\alpha)/p_2 > c\alpha$, the solution is

$$x_1 = \frac{E\alpha}{p_1} + \alpha\frac{p_2}{p_1}c \quad \text{and} \quad x_2 = \frac{E(1-\alpha)}{p_2} - c\alpha.$$

Otherwise, if $E(1-\alpha)/p_2 \leq c\alpha$, then

$$x_1 = \frac{E}{p_1} \quad \text{and} \quad x_2 = 0.$$

Understanding the Lagrangian As in the previous section, let's us now try to understand what the Lagrangian is doing to solve this maximization problem with non-trivial equality and inequality constraints. First, the constrained problem (9) is converted to the unconstrained problem of maximizing the Lagrangian (10), with the associated optimality conditions being (11)-(13).

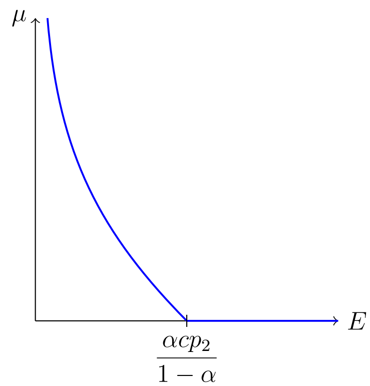
What is the inequality multiplier μ doing in (10)? We saw in the solution process that, if $E(1-\alpha) \leq c\alpha p_2$ (the condition for $x_2 = 0$), μ is given by (20). That is, μ is higher if α is high, or if p_2 is high, or if E is low, or if c is high. These are the same conditions for c_2 being smaller/more negative, as we've seen in (18). That is, if the parameters are such that the consumer would like to make x_2 very negative, μ assumes a large positive value, so that the term "μx_2" in the Lagrangian (10) is very negative (because μ is positive and large and x_2 is very negative). That is, μ penalizes the objective function, so that choosing a negative value for x_2 actually leads to a "bad" (low or very negative) value for the objective function.

In the opposite case where $E(1-\alpha) > c\alpha p_2$, the unconstrained objective function (10) doesn't need to be penalized by μ for negative values of x_2, so μ equals zero.

Another way to see this is to analyze how μ changes explicitly as a function of parameters α, p_2, E, and c. Let's take E, for example. Remember the intuition that, if the consumer is very poor (low E), the consumer would want to sell a lot of x_2 to buy x_1. Therefore, the optimal μ in (20) penalizes more the Lagrangian if E is higher. For values of E larger than a threshold value, μ doesn't need to penalize the Lagrangian because the consumer is rich enough to decide to buy x_2 (instead of sell it):

$$\mu \geq 0 \iff E \geq \frac{\alpha c p_2}{1-\alpha}.$$

Therefore, we can plot μ as a function of E as:



Note that, the poorer the consumer is, the higher μ is to penalize the Lagrangian for the consumer choosing to sell a large quantity of x_2. If E is greater than a threshold level, μ becomes zero.

Appendix C Using the Hamiltonian

- This section based on the mathematical appendix of "Economic Growth" by Barro and Sala-i-Martin.

The dynamic problems in continuous time we see in the course can be written in general as:

$$\max_{[c(t)]_{t\geq 0}} \int_0^\infty v(k(t), c(t), t) dt \tag{1}$$

subject to

$$\dot{k}(t) = g(k(t), c(t), t) \tag{2}$$

$$k(0) \text{ given} \tag{3}$$

$$\lim_{t\to\infty} k(t)e^{-\bar{r}(t)t} \geq 0. \tag{4}$$

$v(\cdot)$ is called the felicity function (usually $v(k(t), c(t), t) = e^{-(\rho-n)t}u(c(t))$, where $u(\cdot)$ is an instantaneous utility function, and ρ and n are parameters), $c(t)$ is the choice variable (or control variable), and $k(t)$ is called the state variable. Equation (2) is called the transition equation, and (3) says that the initial condition of the state variable is exogenous. The last equation (4) is generally associated to a no-Ponzi constraint. In some problems, the constraint (4) is not needed (such as in the social planner's problem in the neoclassical growth model).

Here is a cookbook procedure to find the optimality conditions associated to this problem.

1. Construct the Hamiltonian function by adding to the felicity function, $v(\cdot)$, a Lagrange multiplier times the right-hand side of the transition equation:

$$\mathcal{H}(t) = v(k(t), c(t), t) + \lambda(t)g(k(t), c(t), t). \tag{5}$$

2. Take the derivative of the Hamiltonian with respect to the control variable and set it to 0:

$$\frac{\partial \mathcal{H}(t)}{\partial c(t)} = \frac{\partial v(k(t), c(t), t)}{\partial c(t)} + \lambda(t)\frac{\partial g(k(t), c(t), t)}{\partial c(t)} = 0. \tag{6}$$

3. Take the derivative of the Hamiltonian with respect to the state variable (the variable that appears with a dot above it in the transition equation) and set it

to equal the negative of the derivative of the multiplier with respect to time:

$$\frac{\partial \mathcal{H}(t)}{\partial k(t)} = \frac{\partial v(k(t), c(t), t)}{\partial k(t)} + \lambda(t) \frac{\partial g(k(t), c(t), t)}{\partial k(t)} = -\dot{\lambda}(t). \tag{7}$$

4. Use the transversality condition:

$$\lim_{t \to \infty} \lambda(t) k(t) = 0. \tag{8}$$

If we combine equations (6) and (7) with the transition equation (2), we can form a system of two differential equations in the variables λ and k. Alternatively, we can use equation (5) to transform the differential equation for $\dot{\lambda}$ into a differential equation for $\dot{c}$. For the system to be determinate, we need two boundary conditions. One initial condition is given by the starting value of the state variable, $k(0)$, and one terminal condition is given by the transversality condition (8).

Appendix D Differential equations

Introduction A differential equation is an equation involving a function and its derivative(s). For example,

$$f(x) - \frac{\partial f(x)}{\partial(x)} = 0$$

is a differential equation with respect to $f(x)$ because it contains $f(x)$ and its derivative, $\partial f(x)/\partial x = f'(x)$. We say that $f(x)$ is the *unknown* of this differential equation. A solution to a differential equation is a function $f(x)$ that satisfies the differential equation. Since the derivative of $f(x) = e^x$ is $f'(x) = e^x$, the function $f(x) = e^x$ is a solution to the differential equation above. The solution to a differential equation cannot be written in terms of its derivatives.

We will learn how to solve a differential equation of the following type

$$\dot{y}(t) + b(t)y(t) + x(t) = 0 \quad \forall t \geq 0, \tag{1}$$

where $y(t)$ is the function that we want to solve for, and $a(t)$, $b(t)$, and $x(t)$ are known functions. t denotes time and, since time starts at zero and runs forever, the differential equation holds for all $t \geq 0$.

A solution to the differential equation (1) is a function $y(t)$ that tells us the value of y at each point in time t. The solution $y(t)$ can depend on an *initial condition*. For example, the solution $y(t)$ can depend on $y(0)$. The solution $y(t)$ can also depend on $b(t)$, $x(t)$, t, or other parameters of the model. The solution $y(t)$ *cannot* depend on $\dot{y}(t)$.

Here are three differential equations that appear in the neoclassical growth model and that can be written as (1), in order of simplicity (from the simplest to the less simple):

1. The equation that describes how population size evolves across time:

$$\frac{\dot{L}(t)}{L(t)} = n \quad \forall t \geq 0. \tag{2}$$

   $L(t)$ is the size of population at time t, and $n \geq 0$ is population growth. The equation above says that population grows by a *rate* n in each period. In other words, the *relative growth* of $L(t)$ (relative to its current value) is n. For example, if $n = 0.1 = 10\%$, it means that population grows by 10% in each period.

We will see that the solution to this equation is:

$$L(t) = L(0)e^{nt}.$$

Using $y(t) = L(t)$, $b(t) = -n$, and $x(t) = 0$ in equation (1), we get equation (2).

2. The Euler equation for the CIES utility function:

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\theta}[r(t) - \rho] \quad \forall t \geq 0. \tag{3}$$

The interpretation of this equation is similar to that in item 1: consumption changes at rate $(1/\theta)[r(t) - \rho]$ at time t. However, note now that the relative growth of consumption may change over time because $r(t)$ may vary over time. If, for example, $\theta = 3$, $\rho = 0.02$, and $r(t) = 0.04$ for a given t, then consumption grows by $3^{-1} \times (0.04 - 0.02) = 0.006 = 0.6\%$ in period t.
We will see that the solution to this equation is:

$$c(t) = c(0)e^{\int_0^t (1/\theta)[r(\tau)-\rho]d\tau}.$$

Using $y(t) = c(t)$, $b(t) = -(1/\theta)[r(t) - \rho]$, and $x(t) = 0$ in equation (1), we get equation (3).

3. The household's budget constraint:

$$\dot{a}(t) = w(t) + [r(t) - n]a(t) - c(t) \quad \forall t \geq 0. \tag{4}$$

We will see that the solution $a(t)$ to this equation satisfies

$$e^{-\int_0^t [r(\tau)-n]d\tau} a(t) - a(0) = \int_0^t e^{-\int_0^\tau [r(v)-n]dv} [w(\tau) - c(\tau)]d\tau.$$

Note that we could isolate $a(t)$ in the equation above, but we don't do it to have a simpler expression. This equation has an interesting economic intuition that we see in section X. Using $y(t) = a(t)$, $b(t) = -[r(t) - n]$, and $x(t) = c(t) - w(t)$ in equation (1), we get equation (4).

A simpler solution for a simpler equation Let's start by solving a simpler equation. Assume that $x(t) = 0$. The differential equation (1) becomes

$$\dot{y}(t) + b(t)y(t) = 0. \tag{5}$$

The two first examples previously discussed fall into this simpler case.

Rewrite the equation above as

$$\frac{\dot{y}(t)}{y(t)} = -b(t).$$

Assume that this equation holds for all $t \geq 0$. We can write the following:

$$\frac{\dot{y}(\tau)}{y(\tau)} = -b \qquad \forall \tau \geq 0. \tag{6}$$

That is, we only switched t for τ, and we wrote "for all $\tau \geq 0$" to make this fact explicit.

Since this holds for all $\tau \geq 0$, we can integrate both sides of the equation from time periods between 0 and t, where t is *any given time period* greater or equal to zero:

$$\int_0^t \frac{\dot{y}(\tau)}{y(\tau)} d\tau = -\int_0^t b(\tau)\ d\tau. \tag{7}$$

This equation holds for any time period $t \geq 0$.

Note the following fact:

$$\frac{\partial\{\ln(y(t)\}}{\partial t} = \frac{\dot{y}(t)}{y(t)}.$$

That is, the function $\ln(y(t))$ is the *primitive function*, or the *antiderivative*, of function $\dot{y}(t)/y(t)$. Thus, (7) implies that

$$\ln(y(t)) - \ln(y(0)) = -\int_0^t b(\tau)\ d\tau.$$

Isolating $y(t)$,

$$y(t) = y(0)e^{-\int_0^t b(\tau)d\tau}. \tag{8}$$

This is the solution to differential equation (5). If we know the value of $y(0)$, we can know the value of y at any point in time t by evaluating function (8) at t.

In the simpler case where $b(t) = b$ doesn't depend on time, (8) simplifies to

$$y(t) = y(0)e^{-bt}.$$

We have started assuming that (5) holds and, using logics and algebra, we arrived at (8). This can be summarized in mathematical notation as

$$\dot{y}(t) + b(t)y(t) = 0 \quad \forall t \geq 0 \Rightarrow y(t) = y(0)e^{-\int_0^t b(\tau)d\tau} \quad \forall t \geq 0. \tag{9}$$

This mathematical statement says that (8) is the *unique* solution to the differential equation (5): if $y(t)$ is described by (8), then $y(t)$ *must be written as* $y(t) = y(0)e^{-\int_0^t b(\tau)d\tau}$.

It's easy to prove that, if we start assuming (8), we arrive at (5). This is written as

$$y(t) = y(0)e^{-\int_0^t b(\tau)d\tau} \quad \forall t \geq 0 \Rightarrow \dot{y}(t) + b(t)y(t) = 0 \quad \forall t \geq 0. \tag{10}$$

We can write (10) and (9) together as

$$\dot{y}(t) + b(t)y(t) = 0 \quad \forall t \geq 0 \Leftrightarrow y(t) = y(0)e^{-\int_0^t b(\tau)d\tau} \quad \forall t \geq 0. \tag{11}$$

This logical expression implies that defining $y(t)$ through (8) is equivalent through defining $y(t)$ implicitly through the differential equation (5).

The solution we found for the differential equation has an intuitive interpretation. Let's start with the simpler case and say that $y(t) = L(t)$ and $b(t) = n$. We have the differential equation describing population growth in the Ramsey model:

$$\frac{\dot{L}(t)}{L(t)} = n \quad \forall t \geq 0.$$

This equation says that population grows by a fixed proportion every time period. The conclusion we get is that $L(t)$ starts at $L(0)$ and, from there, *grows exponentially* with time:

$$L(t) = L(0)e^{nt}.$$

This is intuitive: if population grows by, say, 10% every time period, it accumulates larger increments in each subsequent period. Consequently, its growth exhibits an exponential pattern. A numerical example and a discrete-time approximation might make this clearer. Let's say that population starts at 100 increases by 10%. In the "next time period", population is 110. Then it grows again by 10%, or 11 units. In the "next time period", population size is 121. Now it grows by 12.1... and so on. This progression demonstrates that in each time period population grows by more and more. This is why its growth is exponential. This is the essence of exponential growth and the reason why the Euler constant, e, appears in the solution to the differential equation.

In a more general case where the growth rate changes with time, $n(t)$, the solution to the differential equation is

$$L(t) = L(0)e^{\int_0^t n(\tau)d\tau}.$$

The expression $e^{\int_0^t n(\tau)d\tau}$ symbolizes the compounding of growth rates over the continuous time span. The term "compounding" conveys the notion that the growth rate for each time period is aggregated, or *summed up*. Since time is continuous, this *summation* is represented by the integral.[38]

The Leibniz rule The solution method for the more general differential equation makes use of the Leibniz rule. This rule is related to the derivative of an integral. The Leibniz rule says that, under continuity conditions,

$$\frac{\partial \left\{\int_{a(x)}^{b(x)} f(x,y)dy\right\}}{\partial x} = f(x,b(x))\frac{\partial b(x)}{\partial x} - f(x,a(x))\frac{\partial a(x)}{\partial x} + \int_{a(x)}^{b(x)} \frac{\partial f(x,y)}{\partial x}dy.$$

We will make use of a simpler version of this rule, where function $f(x,y)$ doesn't depend on y, $a(x)$ doesn't depend on x, and $b(x) = x$. Relabeling some variables so that they fit our context, this simpler Leibniz rule is

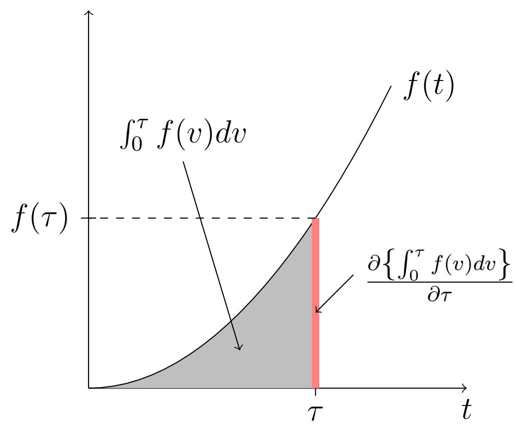
$$\frac{\partial \left\{\int_a^\tau f(v)dv\right\}}{\partial \tau} = f(\tau). \tag{12}$$

This rule has an intuitive geometric interpretation. Let's say that $f(t)$ is a function of time, and we are integrating this function between $t = 0$ and $t = \tau$, for some $\tau \geq 0$. Recall the integral of a function is the area below the integrand function over the interval determined by the integration bounds.

The first thing is to note that the integral $\int_a^\tau f(v)dv$ is a function of the upper integration bound, τ: if τ changes, the area below $f(\cdot)$ changes, and so the integral

[38]See appendices D and F.2 for more on compounding time-varying rates in continuous time.

changes. The grey area in the figure below denotes $\int_a^\tau f(v)dv$. The figure makes clear that, as τ is moved to the right, the gray area increases. The Leibniz rule (12) answers the following question: by how much does the area below $f(\cdot)$ change if we increase the upper integration bound? The left-hand side of (12) represents this question. Stated differently, by how much does the grey area in the figure below change if the upper bound of the integral τ increases? The answer is that this area must change exactly by the heigh of the red bar in the figure. In other words, the derivative of $\int_a^\tau f(v)dv$ with respect to τ must be equal to the integrand $f(\cdot)$ evaluated at the upper integration bound, $f(\tau)$.



Solving the more general differential equation Now the equation we want to solve is:

$$\dot{y}(t) + b(t)y(t) + x(t) = 0 \quad \forall t \geq 0. \tag{13}$$

Put all terms that have to do with $y(t)$ in the left-hand side, and all others in the right-hand side:

$$\dot{y}(t) + b(t)y(t) = -x(t). \tag{14}$$

Multiply both sides by the *integrating factor*, $e^{\int_0^t b(v)dv}$ (why? Just wait a few lines):

$$e^{\int_0^t b(v)dv}\left[\dot{y}(t) + b(t)y(t)\right] = -e^{\int_0^t b(v)dv}x(t).$$

Since this equation holds for all $t \geq 0$, we can write the following:

$$e^{\int_0^\tau b(v)dv}\left[\dot{y}(\tau) + b(\tau)y(\tau)\right] = -e^{\int_0^\tau b(v)dv}x(\tau) \quad \text{for all } \tau \geq 0.$$

That is, we only switched t for τ, and we wrote "for all $\tau \geq 0$" to make this fact explicit.

Since this holds for all $\tau \geq 0$, we can integrate both sides of the equation from time periods between 0 and t, where t is *any given time period*:

$$\int_0^t e^{\int_0^\tau b(v)dv} \left[\dot{y}(\tau) + b(\tau)y(\tau)\right] d\tau = -\int_0^t e^{\int_0^\tau b(v)dv} x(\tau)d\tau. \tag{15}$$

This equation holds for any time period $t \geq 0$.

Now note that

$$\begin{aligned} \frac{\partial \left\{ e^{\int_0^t b(v)dv} y(t) \right\}}{\partial t} &= e^{\int_0^t b(v)dv} \dot{y}(t) + y(t) e^{\int_0^t b(v)dv} \frac{\partial \left\{ \int_0^t b(\tau)d\tau \right\}}{\partial t} \\ &= e^{\int_0^t b(v)dv} \left[y(t) + b(t)y(t) \right]. \end{aligned} \tag{16}$$

In the second equality above, we made use of a simpler version of the Leibniz rule, discussed previously:

$$\frac{\partial \left\{ \int_0^t b(\tau)d\tau \right\}}{\partial t} = b(t).$$

Calculation (16) shows that $e^{\int_0^t b(v)dv} y(t)$ is the *primitive function*, or the *antiderivative*, of function $e^{\int_0^t b(v)dv} \left[\dot{y}(t) + b(t)y(t)\right]$. *This is why we multiplied by both sides of (14) by the integrating factor*: because this way we can solve the integral in the left-hand side of (15) and get rid of the term $\dot{y}(t)$.

Using this, we can rewrite (15) as

$$e^{\int_0^t b(v)dv} y(t) - y(0) = -\int_0^t e^{\int_0^\tau b(v)dv} x(\tau)d\tau. \tag{17}$$

The function $y(t)$ is the solution to the differential equation (13).

If needed, one can isolate $y(t)$ and write it as

$$y(t) = e^{-\int_0^t b(v)dv} y(0) - e^{-\int_0^t b(v)dv} \int_0^t e^{\int_0^\tau b(v)dv} x(\tau)d\tau. \tag{18}$$

Appendix E From discrete to continuous time: approach one (Deltas and limits)

Section 2 presents the neoclassical growth model in discrete time and section 3 develops the model in continuous time. Some equations are very similar across the two models, such as

$$R_t = r_t + \delta$$

in the discrete-time model and

$$R(t) = r(t) + \delta$$

in the continuous-time model. However, other expressions are significantly different, such as the discounted utility of households, which reads as

$$\sum_{t=0}^{\infty} \beta^t u(c_t) \tag{1}$$

in discrete time and

$$\int_0^{\infty} e^{-\rho t} u(c(t)) dt \tag{2}$$

in continuous time.

The objective of this section is to understand why some mathematical objects are written in significantly different ways depending on whether the model is in discrete or continuous time.

The most usual approach to understanding how some equations in continuous time differ from those in discrete time starts with the observation that *a model in continuous time is the limit of a discrete-time model with the time period length tending to zero*. We call this method the "Delta-limit approach", and it can be split into two steps, the first being the "Delta step" and the second the "limit step".

The time periods in the model of section 2 are $t = 0, 1, \dots$. That is, the length between any two consecutive time periods is $(t + 1) - (t) = 1$. Let's say we want to understand why lifetime discounted utility looks like (2) in continuous time. In the first step (Delta step), we start by writing the discrete-time expression for lifetime utility (1) assuming that the time period length is not necessarily one, but Δ, where Δ can be anything strictly positive, $\Delta > 0$. Second, we take the limit of this expression with Δ tending to zero (this is the second step, or the limit step). The resulting limit is the expression in continuous time.

To execute the "Delta step", we need to understand the different types of variables used in the Ramsey model:

1. *Stock variables* are those that exist for more than one period of time (for example, assets and capital).

2. *Flow variables* are those that exist for only one period of time (for example, consumption and output).

3. *Rate variables* are... rates. For example, if we assume that population grows by a fixed rate over time, equal to an exogenous variable $n \geq 0$,

$$\frac{L_{t+1} - L_t}{L_t} = n \tag{3}$$

then n is a rate variable.

The "Delta step" consists in writing a given mathematical expression in discrete time assuming that the time period length is a generic $\Delta > 0$. To do this, we use the following rule, which we will refer to as the "flow-rate-stock rule":[39][40]

1. Flow and rate variables should be multiplied by Δ.

2. Stock variables should not be multiplied by Δ.

As an example, let's apply this method to understand how the equation describing population growth (3) should look like in continuous time. Equation (3) involves two variables: population size L_t, which is a stock variable because it denotes the *stock* of agents alive in t, and the population growth rate n, which is a rate variable. Using the flow-rate-stock rule, if the length of a time period is $\Delta > 0$, equation (3) reads

$$\frac{L_{t+\Delta} - L_t}{L_t} = \Delta n. \tag{4}$$

We have replaced the subscript "$t+1$" for "$t+\Delta$" because the time length changed from one to Δ, and we have multiplied the rate variable n by Δ, following the rule described above. The intuition of (4) is that, if population grows by, say, $n = 10\%$ every year

[39] This rule works for the Ramsey model, but it is not a general rule that works for converting any model from discrete to continuous time. See footnote 41 for more details.

[40] Thank Pontus Rendahl for this intuition. Slides: "Advanced Tools in Macroeconomics: Continuous time models (and methods)", August 21, 2017.

(assuming a time length of one unit corresponds to one year), then population grows by (approximately) $0.5 \times n = 5\%$ every semester (assuming we want to write the model with a time period length of a half of a year, $\Delta = 0.5$).[41]

Rewrite (4) as

$$\frac{1}{L_t}\frac{L_{t+\Delta} - L_t}{\Delta} = n.$$

The final "limit step" involves taking the limit of both sides of the equation above with $\Delta \to 0$:

$$\lim_{\Delta\to 0}\frac{1}{L(t)}\frac{L(t+\Delta) - L(t)}{\Delta} = \lim_{\Delta\to 0} n.$$

We have switched the notation from L_t to $L(t)$ to make even more explicit that L is a function of time. The right-hand side doesn't depend on Δ, so we can eliminate the limit operator. For the same reason, we can take the first factor of the multiplication in the left-hand side outside of the limit operator:

$$\frac{1}{L(t)}\lim_{\Delta\to 0}\frac{L(t+\Delta) - L(t)}{\Delta} = n.$$

Remember the formal definition of a derivative: $\partial f(x)/\partial x = \lim_{\Delta\to 0}[f(x+\Delta) - f(x)]/\Delta$. Thus,

$$\frac{\dot{L}(t)}{L(t)} = n, \tag{5}$$

where $\dot{L}(t) \equiv \partial L(t)/\partial t$.

To summarize, we have used the "Delta-limit approach" to understand why the equation describing a constant population growth (3) in discrete time is written as (5) in continuous time.

E.1 The budget constraint

The budget constraint in discrete time is

$$c_t + a_{t+1} = w_t + (1 + r_t)a_t.$$

[41] The flow-rate-stock rule works for our purposes, but it is not a general rule to convert any discrete-time model to continuous time. For example, in the Solow model, the exogenous savings rate s should not be multiplied by Δ when converting the model to time length Δ. The idea is that if agents save, say, 10% of the GDP every year, it's reasonable to assume that they also save 10% of GDP every semester. The semestral GDP, however, should be converted and, thus, would be half of the yearly GDP.

This equation assumes that the time period length is one: assets in the current time period are a_t, and assets in the *next* time period are a_{t+1}. Using the flow-rate-stock rule to convert the time length to $\Delta > 0$, we obtain

$$\Delta c_t + a_{t+\Delta} = \Delta w_t + (1 + \Delta r_t)a_t.$$

Note we have multiplied the variables c_t (flow), w_t and r_t (rates) by Δ, but not the asset variables (stocks). Note also that "$t+1$" becomes "$t+\Delta$".

To understand the equation above, assume that the frequency of the unitary time length model is annual. Assume also that $\Delta = 0.5$, so that the model with time length Δ is half of a year, or a semester. The idea behind the equation above is that, if c_t is the consumption flow during one year, then the consumption flow during a semester is $\Delta c_t = 0.5c_t$, or half the consumption made during one year. A similar logic applies to w_t and r_t. The stock variable a_t is not multiplied by Δ. The idea is that if, for instance, the unitary time length model indicates that the assets at the beginning of the year 2020 are €$1,000$, then the semestral model ($\Delta = 0.5$) should also indicate that the assets at the beginning of 2020 are €$1,000$. That is, the frequency of the model should not alter its predictions for assets in a given date.

Rewrite the equation above as

$$\frac{a_{t+\Delta} - a_t}{\Delta} = w_t + r_t a_t - c_t.$$

Since this equation holds for any $\Delta > 0$, it should hold at the limit with Δ tending to zero. Therefore,

$$\dot{a}(t) = \lim_{\Delta \to 0} \frac{a(t+\Delta) - a(t)}{\Delta} = w(t) + r(t)a(t) - c(t).$$

The first equality makes use of the definition of a derivative.

E.2 Time discounting

Let's understand why the household's discounted utility in discrete time, $\sum_{t=0}^{\infty} \beta^t u(c_t)$, is written as $\int_0^\infty e^{-\rho t} u(c(t))dt$ in continuous time.

Let's start by writing the discounted lifetime utility of a household in time zero:

$$\underbrace{u(c_0)}_{\text{utility in } t=0} + \beta \underbrace{\{u(c_1) + \beta u(c_2) + \dots\}}_{\text{discounted lifetime utility in } t=1}.$$

Similarly, when time period $t = 1$ arrives, the discounted lifetime utility can be written as

$$\underbrace{u(c_1)}_{\text{utility in } t=1} + \beta \underbrace{\{u(c_2) + \beta u(c_3) + \dots\}}_{\text{discounted lifetime utility in } t=2},$$

and so on. This shows that, in every time period, the household discounts future utilities by the constant factor $\beta \in (0, 1)$:

$$\begin{gathered}\text{Current discounted lifetime utility} = \\ (\text{Current utility}) + \beta(\text{Discounted lifetime utility in the next period}).\end{gathered} \tag{6}$$

Now let's think about discrete and continuous time. In discrete time, the discounted lifetime utility is obtained as a sum, while in continuous time it is written as an integral. Sums and integrals are very closely related mathematical operations. Both are used to calculate the total of a set of values. The sum is used when the set of values is "discrete", while the integral is used when the set is "continuous". This intuition is enough for understanding why the "$\sum$" symbol becomes "$\int$". Therefore, now we only need to understand why the term β^t in discrete time is written as $e^{-\rho t}$ in continuous time.

Define the exogenous variable ρ through $\beta = (1 + \rho)^{-1}$, or $\rho \equiv \beta^{-1} - 1$. The household's lifetime utility can be written as

$$\sum_{t=0}^{\infty}(1+\rho)^{-t}u(c_t) = u(c_0) + (1+\rho)^{-1}u(c_1) + (1+\rho)^{-2}u(c_2) + \dots. \tag{7}$$

Now the discount factor is $(1 + \rho)^{-t}$ instead of β^t. Why have we rewritten the discount factor this way? In section 2.1.1 we saw that the factor to make present-value calculations with monetary units is $\prod_{\tau=1}^{t}(1 + r_\tau)^{-1}$. Note that if the interest rate in all time periods is ρ ($r_\tau = \rho$ for all τ), the present-value factor becomes

$$\prod_{\tau=1}^{t}(1+r_\tau)^{-1} = \prod_{\tau=1}^{t}(1+\rho)^{-1} = (1+\rho)^{-t},$$

which is exactly equal to the discount factor in (7). That is, ρ is the implicit interest rate in the present-value factor used for utility calculations, derived from the time discount parameter β. We show in section X that there's a close connection between ρ and the market interest rate in the Ramsey model.

Let's use the "Delta-limit approach" to convert the discrete-time model to continuous time. The first step is to use the flow-rate-stock rule. We saw that ρ an implied interest rate, implying it is a rate variable. Therefore, the term "ρ" will show up as "$\Delta\rho$" in the expressions for the model with time length Δ. Second, applying the idea in equation (6) to the model with time length Δ,

$$\text{Current discounted lifetime utility} =$$
$$(\text{Current utility}) + (1+\Delta\rho)^{-1}(\text{Discounted lifetime utility in the next period}).$$

Therefore, the discounted lifetime utility of the household in the model with a time period length of Δ is

$$u(c_0) + (1+\Delta\rho)^{-1}u(c_\Delta) + (1+\Delta\rho)^{-2}u(c_{2\Delta}) + \dots. \tag{8}$$

Note that the present-value discount factor in a given time period t, where $t \in \{0, \Delta, 2\Delta, \dots\}$, is

$$(1+\Delta\rho)^{-t/\Delta}. \tag{9}$$

For example, if $t = 2\Delta$, the discount factor is $(1+\Delta\rho)^{-2\Delta/\Delta} = (1+\Delta\rho)^{-2}$, as seen in equation (8).

To get the continuous-time expression, take the limit of the expression in (9) with Δ tending to zero:[42]

$$\lim_{\Delta\to 0}(1+\Delta\rho)^{-t/\Delta} = e^{-\rho t}.$$

This shows that the term β^t in discrete time becomes $e^{-\rho t}$ in continuous time.

E.3 The present-value factor (no-Ponzi condition)

The no-Ponzi condition in discrete time is

$$\lim_{t\to\infty} a_{t+1}\prod_{\tau=1}^{t}(1+r_\tau)^{-1} \geq 0,$$

while in continuous time it is

$$\lim_{t\to\infty} a(t)e^{-\int r(\tau)d\tau} \geq 0. \tag{10}$$

[42]We use the famous result $\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^{nb} = e^{ab}$ making the substitution of variable $\Delta = 1/n$.

Thus, to understand why (10) is written like this in continuous time, we need to understand why the present-value factor for monetary calculations in continuous time is $e^{-\int r(\tau)d\tau}$ using the fact that in discrete time it is given by $\prod_{\tau=1}^{t}(1+r_\tau)^{-1}$.

Let's say that each time period in the discrete-time model with unitary time period length corresponds to a year. Let's assume that $t=0$ corresponds to the year 2020. The present-value factor to convert monetary units from 2021 to 2020 is $(1+r_1)^{-1}$ because saving $(1+r_1)^{-1}$ euros in 2020 leads to one euro in 2021.[43]

How should the present-value factor look like if we assume that the time period length is a half instead of one? That is, $\Delta = 0.5$ and each period in the model corresponds to one semester. Using the flow-rate-stock rule, the interest rate is Δr_t per semester. If we save one euro in the first semester of 2020, we get $(1+\Delta r_1)$ in the second semester of 2020. Then, if we save this amount in the second semester of 2020, we get $(1+\Delta r_1)(1+\Delta r_1) = (1+\Delta r_1)^2$ in the first semester of 2021. Therefore, the present-value of one euro in the beginning of 2021 is the inverse of that amount, $(1+\Delta r_1)^{-2}$. The square in the exponent comes from the fact that we have split the time period into two (going from an yearly model to a semester frequency), so the interest rate is compounded twice if we want to save money in the beginning of 2020 to receive it back in the beginning of 2021. If we had split the time period by 3, this would correspond to $\Delta = 1/3$ and the formula would be $(1+\Delta r_1)^{-3}$. Therefore, the generic formula for the present-value of one euro in the beginning of 2021 for the model with a time period length of Δ is $(1+\Delta r_1)^{-1/\Delta}$.

Using this insight, the present-value factor for time period t in the model with a time period length of Δ, denoted by D_t, is

$$D_t = \prod_{\tau=1}^{t}(1+\Delta r_\tau)^{-1/\Delta}.$$

Apply natural logarithms to both sides of the equation,

$$\ln(D_t) = -\sum_{\tau=1}^{t}\frac{\ln(1+\Delta r_\tau)}{\Delta}. \tag{11}$$

Note that[44]

$$\lim_{\Delta\to 0}\frac{\ln(1+\Delta r_\tau)}{\Delta} = r_\tau.$$

[43] See section 2.1.1 for more details on present-value calculations.

[44] First, $\ln(1+\Delta r_\tau)/\Delta = \ln[(1+\Delta r_\tau)^{1/\Delta}]$. Second, $\lim_{\Delta\to 0}\ln[(1+\Delta r_\tau)^{1/\Delta}] = \ln[\lim_{\Delta\to 0}(1+\Delta r_\tau)^{1/\Delta}] = \ln(e^{r_\tau}) = r_\tau$, where, similarly to footnote 42, we have used $\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^{nb} = e^{ab}$.

Taking the limit with $\Delta \to 0$ in equation (11) and using the limit result above, we get the present-value factor in continuous time (substitute the sum for the integral because now time is continuous)

$$\ln(D(t)) = -\int_0^t r(\tau)d\tau,$$

or

$$D(t) = e^{-\int_0^t r(\tau)d\tau}.$$

Appendix F From discrete to continuous time: approach two (differential equations)

Section E shows the usual approach to understanding the conversion of a discrete-time model into continuous time. This section presents an alternative approach that some readers might find easier. Let's say we want to understand why a given mathematical object (e.g., time discount factor, no-Ponzi condition, ...) is written in a specific way in continuous time. The approach in this section relies in noting some property that this mathematical object has in discrete time, writing this property in continuous time, and using this property in continuous time to obtain the expression for the mathematical object in continuous time. We call this method the "differential equation approach" because generally it involves differential equations and its solutions.

F.1 Time discounting

The utility discount factor for time period t in discrete time is β^t. By how much does this discount factor change proportionally between two consecutive time periods?

$$\frac{\beta^{t+1} - \beta^t}{\beta^t} = -(1 - \beta) < 0.$$

Since the right-hand side doesn't depend on time, this shows that the utility discount factor falls by a constant proportion of $-(1 - \beta)$ between any two consecutive time periods. For example, if $\beta = 0.96$, a commonly used value for beta in yearly models, the utility discount factor changes by $-(1 - 0.96) = -4\%$ in each period.

This means that in continuous time we need to model the utility discount factor as a variable that falls by a constant proportion every time period. Denoting by $D(t)$ this discount factor and ρ the negative of the growth rate of $D(t)$, we can write

$$\frac{\dot{D}(t)}{D(t)} = -\rho. \tag{1}$$

$\dot{D}(t) = \partial D(t)/\partial t$ is the change of $D(t)$ in time period t. Since we divide $\dot{D}(t)$ by $D(t)$ in the left-hand side, the equation above says that $\dot{D}(t)/D(t)$, the relative change of $D(t)$ in t, needs to be constant over time and equal to $-\rho$.

Now, the question is: is there a function of time, $D(t)$, that satisfies the (differential) equation (1)? In Appendix D, we show that there's only one solution to this differential equation, which is given by

$$D(t) = D(0)e^{-\rho t}. \tag{2}$$

If you want to check that the equation above solves (1), compute $\dot{D}(t)/D(t)$ using the function $D(t)$ above and you should get $-\rho$.

Finally, we only need to determine $D(0)$. Simply note that the discount factor for $t = 0$ in discrete time is $\beta^0 = 1$. Therefore, we must have $D(0) = 1$ so that the right-hand side of (2) equals one when evaluated at $t = 0$. We conclude that the discount factor for time period t in continuous time should be written as

$$D(t) = e^{-\rho t}$$

if we want it to fall by a constant proportion in each time period, which is a property of the discount factor β^t in discrete time.

Let's go a bit further now. What is the connection between β and ρ? Appendix E.2 shows that defining ρ through $\beta = (1 + \rho)^{-1}$ implies that ρ can be interpreted as the implicit interest rate of the present-value discount factor for utility calculations.

$(1 + \rho)^{-t}$ is the discount factor for time t because it is equal to β^t. Note that the relative change of this discount factor between a final and an initial time period (these two time periods being consecutive), *relative to the final time period*, is

$$\frac{(1 + \rho)^{-(t+1)} - (1 + \rho)^{-t}}{(1 + \rho)^{-(t+1)}} = -\rho. \tag{3}$$

This shows that defining ρ through $\beta = (1 + \rho)^{-1}$ implies that the relative change of the discount factor between two time periods is constant and equal to $-\rho$, which is exactly the interpretation for ρ that we see in (1).

You might be thinking that we normally write the relative change of a variable in terms of its initial value. So why does (3) use the final time period in the denominator? The pragmatic answer to this is: because that's the only way to make the calculation work for our purposes (showing that the relative change of the discount factor $(1+\rho)^{-t}$ equals $-\rho$). However, remember that our objective here is to understand why some mathematical expressions look like the way they do in continuous time. In continuous time, there's no notion of "next" or "previous" time period, as compared to discrete time, where $t-1$ is the "previous" time period and $t+1$ is the "next" time period. This means that if we are thinking about discrete time to understand something related to continuous time, we can safely interchange between $t-1$, t, and $t+1$ without much concern. After all, our objective here is not to provide a formal mathematical proof of something, but simply an intuitive heuristic argumentation.

F.2 The present-value factor (no-Ponzi condition)

Section 2 shows that the no-Ponzi condition in discrete time is

$$\lim_{t\to\infty} a_{t+1} \prod_{\tau=1}^{t} (1+r_\tau)^{-1} \geq 0.$$

The equivalent condition in continuous time is

$$\lim_{t\to\infty} a(t) e^{-\int_0^t r(\tau)d\tau} \geq 0.$$

One first difference between the two expressions is that the term a_{t+1} shows up in the first expression, while the term $a(t)$ shows up in the second. Remember from the previous section F.1 that, in continuous time, there's no notion of "next" $(t+1)$ or "previous" $(t-1)$ time period and, because of that, we can interchange between $t-1$, t, and $t+1$ without much concern when trying to understand a continuous-time expression based on a discrete-time expression. For our purposes, this is enough for understanding why "a_{t+1}" becomes "$a(t)$".

The interesting part is understanding why the term $\prod_{\tau=1}^{t}(1+r_\tau)^{-1}$ (the factor to make present-value calculations with monetary units) in discrete time becomes $e^{-\int_0^t r(\tau)d\tau}$ in continuous time. Let's proceed similarly to the last section and ask the question "how does the present-value factor changes proportionally between two consecutive time periods in discrete time?". Denoting $D_t = \prod_{\tau=1}^{t}(1+r_\tau)^{-1}$, we have that

$$\frac{D_t - D_{t-1}}{D_t} = -r_t.$$

Denoting by $D(t)$ the present-value factor in continuous time, we thus want $D(t)$ to have the following property:

$$\frac{\dot{D}(t)}{D(t)} = -r(t).$$

In Appendix E.2, we show that the only solution to the differential equation above is

$$D(t) = D(0)e^{-\int_0^t r(\tau)d\tau}.$$

The last step is to determine $D(0)$. Simply note that the present-value factor to convert a variable in time period 0 to time period 0 is one. Therefore, $D(0) = 1$.

We learn that the factor to make present-value calculations with monetary units in the continuous-time model is $e^{-\int_0^t r(\tau)d\tau}$