

Part 2: Overlapping Generations Models

Chapters

| | |
|----------------------------------------------------------|-----|
| 8. Overlapping Generations Models | 115 |
| 9. An Application: Social Security and Transitions | 135 |
| 10. Unified Growth Theory | 147 |



Filipe Campante,
Federico Sturzenegger
and Andrés Velasco

Advanced Macro- economics



Suggested citation: Campante, Filipe; Sturzenegger, Federico; and Velasco, Andrés. (2021) 'Part 2: Overlapping Generations Models', in Filipe Campante, Federico Sturzenegger and Andrés Velasco, *Advanced Macro-Economics: An Easy Guide*. London: LSE Press, 2021. Chapters 8–10. <https://doi.org/10.31389/lsepress.ame.2>



Filipe Campante, Federico Sturzenegger and Andrés Velasco 2021. Published under a Creative Commons Attribution NonCommercial licence allowing users to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license does not allow for commercial use.

Overlapping Generations Models

Overlapping generations models

The neoclassical growth model (NGM), with its infinitely-lived and identical individuals, is very useful for analysing a large number of topics in macroeconomics, as we have seen, and will continue to see, for the remainder of the book. However, there are some issues that require a departure from those assumptions. An obvious example involves those issues related to the understanding of the interaction of individuals who are at different stages in their life cycles. If lives are finite and not infinite, as in the NGM, individuals are *not* the same (or at a minimum are not at the same moment in their lives). This diversity opens a whole new set of issues, such as that of optimal consumption and investment over the life cycle, and the role of bequests. It also requires a redefinition of optimality. Not only because we need to address the issue of how to evaluate welfare when agents have different utility functions, but also because we will need to check if the optimality properties of the NGM prevail. For example, if there are poor instruments to save, yet people need to save for retirement, can it be possible that people accumulate too much capital?

This richer framework will provide new perspectives for evaluating policy decisions such as pensions, taxation, and discussing the impact of demographic changes. Of course, the analysis becomes more nuanced, but the added difficulty is not an excuse for not tackling the issue, particularly because in many instances the fact that individuals are different is the key aspect that requires attention.

To study these very important issues, in the next three chapters we develop the *overlapping generations* (OLG) model, the second workhorse framework of modern macroeconomics. We will see that, when bringing in some of these nuances, the implications of the model turn out to be very different from those of the NGM. This framework will also allow us to address many of the most relevant current policy debates in macroeconomics, including low interest rates, secular stagnation, and topics in fiscal and monetary policy.

8.1 | The Samuelson-Diamond model

The Samuelson-Diamond model simplifies by assuming two generations: young and old. The young save for retirement, and this is the capital stock next period. The dynamics of capital will be summarised by a savings equation of the form $s(w, r)$. This savings equation will allow us to trace the evolution of capital over time.

How to cite this book chapter:

Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide*.

Ch. 8. 'Overlapping generations models', pp. 115–134. London: LSE Press.

DOI: <https://doi.org/10.31389/lsepress.ame.h> License: CC-BY-NC 4.0.

Here we present a discrete time model initially developed by Diamond (1965), building on earlier work by Samuelson (1958), in which individuals live for two periods (young and old). The economy lasts forever as new young people enter in every period. We first characterise the decentralised competitive equilibrium of the model. We then ask whether the market solution is the same as the allocation that would be chosen by a central planner, focusing on the significance of the golden rule, which will allow us to discuss the possibility of dynamic inefficiency (i.e. excessive capital accumulation).

8.1.1 | The decentralized equilibrium

The market economy is composed of individuals and firms. Individuals live for two periods. They work for firms, receiving a wage. They also lend their savings to firms, receiving a rental rate.

An individual born at time t consumes c_{1t} in period t and c_{2t+1} in period $t+1$, and derives utility

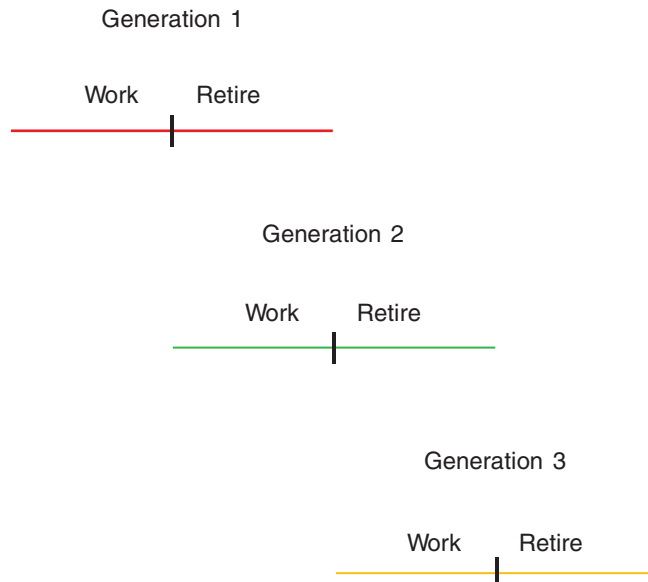
$$\left(\frac{\sigma}{\sigma-1}\right) c_{1t}^{\frac{\sigma-1}{\sigma}} + (1+\rho)^{-1} \left(\frac{\sigma}{\sigma-1}\right) c_{2t+1}^{\frac{\sigma-1}{\sigma}}, \quad \rho \geq 0, \quad \sigma \geq 0. \quad (8.1)$$

Note that the subscript “1” refers to consumption when young, and “2” labels consumption when old. Individuals work only in the first period of life, inelastically supplying one unit of labour and earning a real wage of w_t . They consume part of their first-period income and save the rest to finance their second-period retirement consumption. The saving of the young in period t generates the capital stock that is used to produce output in period $t+1$ in combination with the labour supplied by the young generation of period $t+1$.

The time structure of the model appears in Figure 8.1.

The number of individuals born at time t and working in period t is L_t . Population grows at rate n so that $L_t = L_0 (1+n)^t$.

Figure 8.1 Time structure of overlapping generations model



Firms act competitively and use the constant returns technology $Y = F(K, L)$. For simplicity, assume that capital fully depreciates after use, which is akin to assuming that $F(\cdot, \cdot)$ is a net production function, with depreciation already accounted for. As before, output per worker, Y/L , is given by the production function $y = f(k)$, where k is the capital-labour ratio. This production function is assumed to satisfy the Inada conditions. Each firm maximises profits, taking the wage rate, w_t , and the rental rate on capital, r_t , as given.

We now examine the optimisation problem of individuals and firms and derive the market equilibrium.

Individuals

Consider an individual born at time t . His maximisation problem is

$$\text{Max} \left\{ \left(\frac{\sigma}{\sigma - 1} \right) c_{1t}^{\frac{\sigma-1}{\sigma}} + (1 + \rho)^{-1} \left(\frac{\sigma}{\sigma - 1} \right) c_{2t+1}^{\frac{\sigma-1}{\sigma}} \right\} \quad (8.2)$$

subject to

$$c_{1t} + s_t = w_t, \quad (8.3)$$

$$c_{2t+1} = (1 + r_{t+1}) s_t, \quad (8.4)$$

where w_t is the wage received in period t and r_{t+1} is the interest rate paid on savings held from period t to period $t+1$. In the second period the individual consumes all his wealth, both interest and principal. (Note that this assumes that there is no altruism across generations, in that people do not care about leaving bequests to the coming generations. This is crucial.)

The first-order condition for a maximum is

$$c_{1t}^{-\frac{1}{\sigma}} - \left(\frac{1 + r_{t+1}}{1 + \rho} \right) c_{2t+1}^{-\frac{1}{\sigma}} = 0, \quad (8.5)$$

which can be rewritten as

$$\frac{c_{2t+1}}{c_{1t}} = \left(\frac{1 + r_{t+1}}{1 + \rho} \right)^{\sigma}. \quad (8.6)$$

This is the Euler equation for the generation born at time t . Note that this has the very same intuition, in discrete time, as the Euler equation (Ramsey rule) we derived in the context of the NGM.

Next, using (8.3) and (8.4) to substitute out for c_{1t} and c_{2t+1} and rearranging we get

$$s_t = \left(\frac{1}{(1 + r_{t+1})^{1-\sigma} (1 + \rho)^{\sigma} + 1} \right) w_t. \quad (8.7)$$

We can think of this as a saving function:

$$s_t = s(w_t, r_{t+1}), \quad 0 < s_w \equiv \frac{\partial s_t}{\partial w_t} < 1, \quad s_r \equiv \frac{\partial s_t}{\partial r_{t+1}} \geq 0 \text{ or } \leq 0. \quad (8.8)$$

Saving is an increasing function of wage income since the assumption of separability and concavity of the utility function ensures that both goods (i.e. consumption in both periods) are normal. The effect of an increase in the interest rate is ambiguous, however, because of the standard income and substitution

effects with which you are familiar from micro theory. An increase in the interest rate decreases the relative price of second-period consumption, leading individuals to shift consumption from the first to the second period, that is, to substitute second- for first-period consumption. But it also increases the feasible consumption set, making it possible to increase consumption in both periods; this is the income effect. The net effect of these substitution and income effects is ambiguous. If the elasticity of substitution between consumption in both periods is greater than one, then in this two-period model the substitution effect dominates and an increase in interest rates leads to an increase in saving.

Firms

Firms act competitively, renting capital to the point where the marginal product of capital is equal to its rental rate, and hiring labour to the point where the marginal product of labour is equal to the wage

$$f'(k_t) = r_t \quad (8.9)$$

$$f(k_t) - k_t f'(k_t) = w_t, \quad (8.10)$$

where k_t is the firm's capital-labour ratio. Note that $f(k_t) - k_t f'(k_t)$ is the marginal product of labour, because of constant returns to scale.

8.1.2 | Goods and factor market equilibrium

The goods market equilibrium requires that the demand for goods in each period be equal to supply, or equivalently that investment be equal to saving:

$$K_{t+1} - K_t = L_t s(w_t, r_{t+1}) - K_t. \quad (8.11)$$

The left-hand side is net investment: the change in the capital stock between t and $t+1$. The right-hand side is net saving: the first term is the saving of the young; the second is the dissaving of the old.

Eliminating K_t from both sides tells us that capital at time $t+1$ is equal to the saving of the young at time t . Dividing both sides by L_t gives us the equation of motion of capital in per capita terms:

$$(1+n)k_{t+1} = s(w_t, r_{t+1}). \quad (8.12)$$

The services of labour are supplied inelastically; the supply of services of capital in period t is determined by the savings decision of the young made in period $t-1$. Equilibrium in the factor markets obtains when the wage and the rental rate on capital are such that firms wish to use the available amounts of labour and capital services. The factor market equilibrium conditions are therefore given by equations (8.9) and (8.10).

8.1.3 | The dynamics of the capital stock

The capital accumulation equation (8.12), together with the factor market equilibrium conditions (8.9) and (8.10), implies the dynamic behaviour of the capital stock:

$$k_{t+1} = \frac{s[w(k_t), r(k_{t+1})]}{1+n}, \quad (8.13)$$

or

$$k_{t+1} = \frac{s [f(k_t) - k_t f'(k_t) + f'(k_{t+1})]}{1 + n}. \quad (8.14)$$

This last equation implies a relationship between k_{t+1} and k_t . We will describe this as the savings locus. The properties of the savings locus depend on the derivative:

$$\frac{dk_{t+1}}{dk_t} = \frac{-s_w(k_t) k_t f''(k_t)}{1 + n - s_r(k_{t+1}) f''(k_{t+1})}. \quad (8.15)$$

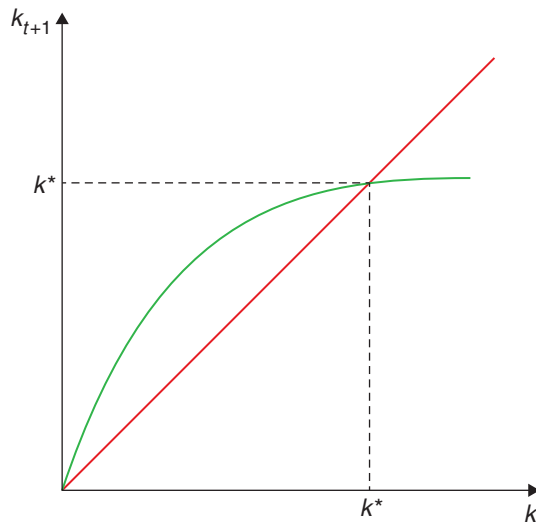
The numerator of this expression is positive, reflecting the fact that an increase in the capital stock in period t increases the wage, which increases savings. The denominator is of ambiguous sign because the effects of increases in the interest rate on savings are ambiguous. If $s_r \geq 0$, then the denominator in (8.15) is positive, and then so is dk_{t+1}/dk_t .

The savings locus in Figure 8.2 summarises both the dynamic and the steady-state behaviour of the economy. The 45-degree line in Figure 8.2 is the line along which steady states, at which $k_{t+1} = k_t$, must lie. Any point at which the savings locus s crosses that line is a steady state. We have drawn a locus that crosses the 45-degree line only once, and hence guarantees that the steady state capital stock both exists and is unique. But this is not the only possible configuration. The model does not, without further restrictions on the utility and/or production functions, guarantee either existence or uniqueness of a steady-state equilibrium with positive capital stock.

If there exists a unique equilibrium with positive capital stock, will it be stable? To answer this, evaluate the derivative around the steady state:

$$\left. \frac{dk_{t+1}}{dk_t} \right|_{ss} = \frac{-s_w k^* f''(k^*)}{1 + n - s_r f''(k^*)}. \quad (8.16)$$

Figure 8.2 The steady-state capital stock



(Local) stability requires that $\left. \frac{dk_{t+1}}{dk_t} \right|_{SS}$ be less than one in absolute value:

$$\left| \frac{-s_w k^* f''(k^*)}{1 + n - s_r f''(k^*)} \right| < 1.$$

Again, without further restrictions on the model, the stability condition may or may not be satisfied. To obtain definite results on the comparative dynamic and steady-state properties of the model, it is necessary either to specify functional forms for the underlying utility and production functions, or to impose conditions sufficient for uniqueness of a positive steady-state capital stock.¹

8.1.4 | A workable example

In this sub-section, we analyse the properties of the OLG model under a fairly simple set of assumptions: log utility (i.e. the limit case where $\sigma = 1$) and Cobb-Douglas production. (This is sometimes referred to as the canonical OLG model.) This permits a simple characterisation of both dynamics and the steady state.

With this assumption on preferences, the saving function is

$$s_t = \left(\frac{1}{2 + \rho} \right) w_t, \quad (8.17)$$

so that savings is proportional to wage income. Notice that the interest rate cancels out in the case of log utility, but not otherwise. This is a case in which the savings rate will be constant over time (as in the Solow model), though, once again, here this is the result of an optimal choice (as in the version of the AK model that we studied in Chapter 5).

With Cobb-Douglas technology, the firm's rules for optimal behaviour (8.9) and (8.10) become

$$r_t = \alpha k_t^{\alpha-1} \quad (8.18)$$

and

$$w_t = (1 - \alpha) k_t^\alpha = (1 - \alpha) y_t. \quad (8.19)$$

Using (8.17) and (8.19) in (8.12) yields

$$k_{t+1} = \left(\frac{1 - \alpha}{2 + \rho} \right) \left(\frac{1}{1 + n} \right) k_t^\alpha, \quad (8.20)$$

which is the new law of motion for capital.

Define as usual the steady state as the situation in which $k_{t+1} = k_t = k^*$. Equation (8.20) implies that the steady state is given by

$$k^* = \left(\frac{1 - \alpha}{2 + \rho} \frac{1}{1 + n} \right)^{\frac{1}{1-\alpha}}, \quad (8.21)$$

so that we have a unique and positive steady-state per-capita capital stock. This stock is decreasing in ρ (the rate of discount) and n (the rate of population growth). Note the similarities with the NGM and the Solow model.

Similarly, we can write steady-state income per-capita as $y^* = (k^*)^\alpha$, or

$$y^* = \left(\frac{1-\alpha}{2+\rho} \frac{1}{1+n} \right)^{\frac{\alpha}{1-\alpha}}. \quad (8.22)$$

Again, this steady-state level is decreasing in ρ and n .

Will the system ever get to the steady state? Local stability requires that $\left. \frac{dk_{t+1}}{dk_t} \right|_{SS}$ be less than one in absolute value, which in this case implies

$$\alpha \left(\frac{1-\alpha}{2+\rho} \right) \left(\frac{1}{1+n} \right) (k^*)^{\alpha-1} = \alpha < 1, \quad (8.23)$$

which is always satisfied. Hence, if the initial capital stock is larger than zero it will gradually converge to k^* . Convergence is depicted in Figure 8.3. The economy starts out at k_0 and gradually moves toward the steady-state capital stock.

The effects of a shock

Suppose next that the economy is at the steady state and at some time 0 the discount rate falls from ρ to ρ' , where $\rho' < \rho$. This shock is unexpected, and will last forever.

From (8.21) we see that the new steady-state per capita capital stock will clearly rise, with $k_{new}^* > k_{old}^*$. In Figure 8.4 we show the dynamic adjustment toward the new stationary position. The economy starts out at k_{old}^* and gradually moves toward k_{new}^* . Income per capita rises in the transition and in the new steady state.

Figure 8.3 Convergence to the steady state

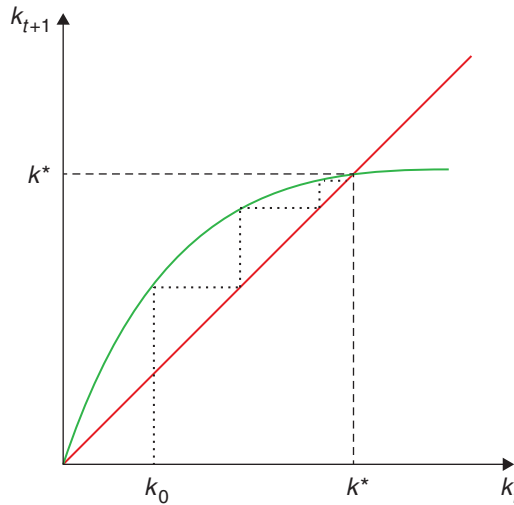
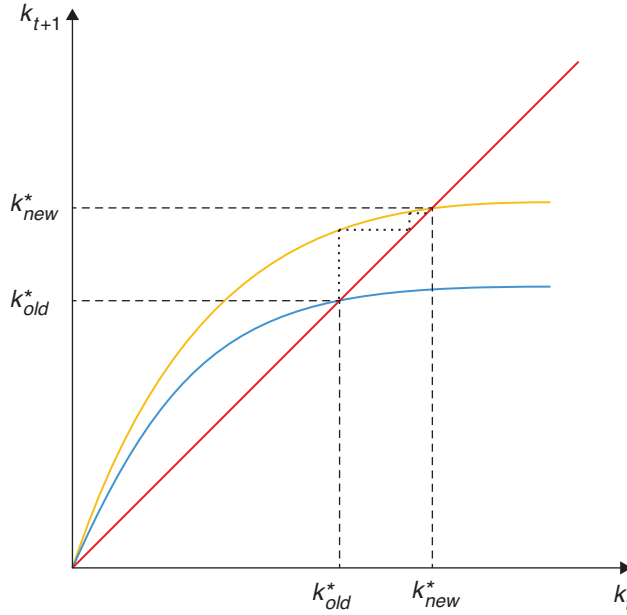


Figure 8.4 Fall in the discount rate

8.2 | Optimality

The distinctive characteristic of the OLG model is that the interest rate may be smaller than the growth rate. In this case, there is a potential gain of reducing the stock of capital. The OLG model can lead to dynamic inefficiency.

We now ask how the market allocation compares to that which would be chosen by a central planner who maximises an intertemporal social welfare function. This raises a basic question, that of the relevant social welfare function. When individuals have infinite horizons and are all alike, it is logical to take the social welfare function to be their own utility function. But here the generations that are alive change as time passes, so it is not obvious what the central planner should maximise.

8.2.1 | The steady-state marginal product of capital

In any event, as in the Solow model, there is something we can say about efficiency here. Notice that, at the steady state, the marginal product of capital is

$$f'(k^*) = \alpha (k^*)^{\alpha-1} = r^* = \left(\frac{\alpha}{1-\alpha} \right) (2 + \rho) (1 + n). \quad (8.24)$$

Notice that this interest rate depends on more parameters than in the NGM. The relationship between the discount factor and the interest rate is still there. A higher discount factor implies less savings today and a higher interest rate in equilibrium. But notice that now that the population growth affects the interest rate. Why is this the case? The intuition is simple. A higher growth rate of population

decreases the steady-state stock of capital thus increasing the marginal product of capital. How does this compare with the golden rule of $f'(k_G) = n$? From the above it is clear that $k^* > k_G$ if

$$r^* < n, \quad (8.25)$$

which in turn implies

$$\alpha < \frac{n}{n + (1 + n)(2 + \rho)}. \quad (8.26)$$

That is, if α is sufficiently low (or, alternatively, if n is sufficiently high), the steady-state capital stock in the decentralised equilibrium can exceed that of the golden rule.

Dynamic inefficiency

Suppose a benevolent planner found that the economy was at the steady state with k^* and y^* . Suppose further that $k^* > k_G$. Is there anything the planner could do to redistribute consumption across generations that would make at least one generation better off without making any generation worse off? Put differently, is this steady state Pareto efficient?

Let resources available for per-capita consumption (of the young and old), in any period t , be given by x_t . Note next that in any steady state,

$$x_{SS} = k_{SS}^\alpha - nk_{SS}. \quad (8.27)$$

Note that, by construction, k_G is the k_{SS} that maximises x_{SS} , since $\frac{\partial x_{SS}}{\partial k_{SS}} = 0$.

The initial situation is one $k_{SS} = k^*$, so that $x_{SS} = c^*$. Suppose next that, at some point $t = 0$, the planner decides to allocate more to consumption and less to savings in that period, so that next period the capital stock is $k_G < k^*$.

Then, in period 0, resources available for consumption will be

$$x_0 = (k^*)^\alpha - nk_G + (k^* - k_G). \quad (8.28)$$

In every subsequent period $t > 0$, resources available for consumption will be

$$x_t = k_G^\alpha - nk_G, \quad t > 0. \quad (8.29)$$

Clearly, in $t > 0$ available resources for consumption will be higher than in the status quo, since k_G maximises x_{SS} . Note next that $x_0 > x_t$ (this should be obvious, since at time 0 those alive can consume the difference between k^* and k_G). Therefore, in $t = 0$ resources available will also be higher than in the status quo. We conclude that the change increases available resources at all times. The planner can then split them between the two generations alive at any point in time, ensuring that everyone is at least as well off as in the original status quo, with at least one generation being better off. Put differently, the conclusion is that the decentralised solution leading to a steady state with a capital stock of k^* is not Pareto efficient. Generally, an economy with $k^* > k_G$ (alternatively, one with $r^* < n$) is known as a dynamically inefficient economy.

8.2.2 | Why is there dynamic inefficiency?

If there is perfect competition with no externalities or other market failures, why is the competitive solution inefficient? Shouldn't the First Welfare Theorem apply here as well? The reason why this isn't the case is the infinity of agents involved, while the welfare theorems assume a finite number of agents.²

An alternative way to build this intuition is that when the interest rate is below the growth rate of the economy, budget constraints are infinite and not well-defined, making our economic restrictions meaningless. This infinity gives the planner a way of redistributing income and consumption across generations that is not available to the market. In a market economy, individuals wanting to consume in old age must hold capital, even if the marginal return on capital is low. The planner, by contrast, can allocate resources between the current old and young in any manner they desire. They can take part of the fruit of the labour income of the young, for instance, and transfer it to the old without forcing them to carry so much capital. They can make sure that no generation is worse off by requiring each succeeding generation to do the same (and remember, there are infinitely many of them).³ And, if the marginal product of capital is sufficiently low (lower than n , so the capital stock is above the golden rule), this way of transferring resources between young and old is more efficient than saving, so the planner can do better than the decentralized allocation.

8.2.3 | Are actual economies dynamically inefficient?

Recall that in the decentralised equilibrium we had

$$r_{ss} = f'(k_{ss}), \quad (8.30)$$

so the rental rate is equal to the marginal product of capital. Notice also that the rate of growth of the economy is n (income per-capita is constant, and the number of people is growing at the rate n). Therefore, the condition for dynamic inefficiency is simply that r_{ss} be lower than the rate of growth of the economy, or, taking depreciation into account (which we have ignored here), that the rate of interest minus depreciation be lower than the rate of growth of the economy.

Abel et al. (1989) extend the model to a context with uncertainty (meaning that there is more than one observed interest rate, since you have to adjust for risk), and show that in this case a sufficient condition for dynamic efficiency is that net capital income exceeds investment. To understand why, notice that the condition for dynamic efficiency is that the marginal product of capital (r) exceeds the growth rate of population (n), which happens to be the growth rate of the economy g . So, rK is the total return to capital and nK is total investment, so the condition $r > g$ can be tested by comparing the return on capital vs new investment: the net flow out of firms. Their evidence from seven industrialised countries suggests that this condition seems to be comfortably satisfied in practice.

However, a more recent appraisal, by Geerolf (2013), suggests that this picture may have actually changed or never been quite as sanguine. He updates the Abel et al. data, and provides a different treatment to mixed income and land rents.⁴ With these adjustments, he finds that, in general, countries are in dynamically efficient positions, though some countries such as Japan and South Korea are definitely in a dynamically *inefficient* state! (And Australia joins the pack more recently...) In other words, it seems at the very least that we cannot so promptly dismiss dynamic inefficiency as a theoretical curiosity.

8.2.4 | Why is this important?

At this point you may be scratching your head asking why we seem to be spending so much time with the question of dynamic efficiency. The reason is that it is actually very relevant for a number of issues. For example, a dynamically inefficient economy is one in which fiscal policy has more leeway. Any debt level will eventually be wiped out by growth. Blanchard (2019) (p. 1197) takes this point seriously and argues

... the current U.S. situation, in which safe interest rates are expected to remain below growth rates for a long time, is more the historical norm than the exception. If the future is like the past, this implies that debt rollovers, that is the issuance of debt without a later increase in taxes, may well be feasible. Put bluntly, public debt may have no fiscal cost.

Not surprisingly, during 2020/2021, in response to the Covid-19 pandemic, many countries behaved as if they could tap unlimited resources through debt issuing. Dynamic inefficiency, if present and expected to remain in the future, would say that was feasible. If, on the contrary, economies are dynamically efficient, the increases in debt will require more taxes down the road.

The second issue has to do with the possibility of bubbles, that is, assets with no intrinsic value. By arbitrage, the asset price of a bubble will need to follow a typical pricing equation

$$(1 + r)P_t = P_{t+1}, \quad (8.31)$$

assuming for simplification a constant interest rate. The solution to this equation is

$$P_t = P_0(1 + r)^t, \quad (8.32)$$

(simply replace to check it is a solution). The price of the asset needs to grow at the rate of interest rate (you may hold a dividend-less asset, but you need to get your return!). In an NGM where $r > g$, this asset cannot exist, because it will eventually grow to become larger than the economy. But if $r < g$ this is not the case, and the bubble can exist. We will come back to this later. What are examples of such assets? Well, you may have heard about Bitcoins and cryptocurrency. In fact, money itself is really a bubble!

Finally, notice that the OLG model can deliver very low interest rates. So, it is an appropriate setup to explain the current world of low interest rates. We will come back to this in our chapters on fiscal and monetary policy.

Before this, however, we need to provide a continuous-time version of the OLG model, to provide continuity with the framework we have been using so far, and because it will be useful later on.

8.3 | Overlapping generations in continuous time

The OLG model can be modelled in continuous time through an ingenious mechanism: a constant probability of death and the possibility of pre-selling your assets upon death in exchange for a payment while you live. This provides cohorts and steady-state behaviour that make the model tractable. Even so, the details get a bit eerie. This section is only for the brave-hearted.

The trick to model the OLG model in a continuous-time framework is to include an age-independent probability of dying p . By the law of large numbers this will also be the death rate in the population. Assume a birth rate $n > p$. Together these two assumptions imply that population grows at the rate $n - p$.⁵ This assumption is tractable but captures the spirit of the OLG model: not everybody is the same at the same time.

As in Blanchard (1985), we assume there exist companies that allow agents to insure against the risk of death (and, therefore, of leaving behind unwanted bequests). This means that at the time of death all of an individual's assets are turned over to the insurance company, which in turn pays a return of p on savings to all agents who remain alive. If r_t is the interest rate, then from the point of view of an individual agent, the return on savings is $r_t + p$.

We will also assume logarithmic utility which will make the algebra easier. As of time t the representative agent of the generation born at time τ maximises

$$\int_t^\infty \log c_{s,\tau} e^{-(\rho+p)(s-t)} ds, \quad (8.33)$$

subject to the flow budget constraint

$$\dot{a}_{t,\tau} = (r_t + p) a_{t,\tau} + y_{t,\tau} - c_{t,\tau}, \quad (8.34)$$

where $a_{t,\tau}$ is the stock of assets held by the individual and $y_{t,\tau}$ is labour income. The other constraint is the no-Ponzi game condition requiring that if the agent is still alive at time s , then

$$\lim_{s \rightarrow \infty} a_{s,\tau} e^{-\int_t^s (r_v + p) dv} \geq 0. \quad (8.35)$$

If we integrate the first constraint forward (look at our Mathematical Appedix!) and use the second constraint, we obtain

$$\int_t^\infty c_{s,\tau} e^{-\int_t^s (r_v + p) dv} ds \leq a_{t,\tau} + h_{t,\tau}, \quad (8.36)$$

where

$$h_{t,\tau} = \int_t^\infty y_{s,\tau} e^{-\int_t^s (r_v + p) dv} ds, \quad (8.37)$$

can be thought of as human capital. So the present value of consumption cannot exceed available assets, a constraint that will always hold with equality.

With log utility the individual Euler equation is our familiar

$$\dot{c}_{s,\tau} = (r_s - \rho) c_{s,\tau}, \quad (8.38)$$

which can be integrated forward to yield

$$c_{s,\tau} = c_{t,\tau} e^{\int_t^s (r_v - \rho) dv}. \quad (8.39)$$

Using this in the present-value budget constraint gives us the individual consumption function

$$\begin{aligned} \int_t^\infty c_{t,\tau} e^{\int_t^s (r_v - \rho) dv} e^{-\int_t^s (r_v + p) dv} ds &= a_{t,\tau} + h_{t,\tau}, \\ c_{t,\tau} \int_t^\infty e^{-(\rho+p)(s-t)} ds &= a_{t,\tau} + h_{t,\tau}, \\ c_{t,\tau} &= (\rho + p)(a_{t,\tau} + h_{t,\tau}), \end{aligned} \quad (8.40)$$

so that the individual consumes a fixed share of available assets, as is standard under log utility. That completes the description of the behaviour of the representative agent in each generation.

The next task is to aggregate across generations or cohorts. Let $N_{t,\tau}$ be the size at time t of the cohort born at τ . Denoting the total size of the population alive at time τ as N_τ , we can write the initial size of the cohort born at τ (that is, the newcomers to the world at τ) as nN_τ . In addition, the probability that someone born at τ is still alive at $t \geq \tau$ is $e^{-p(t-\tau)}$. It follows that

$$N_{t,\tau} = nN_\tau e^{-p(t-\tau)}. \quad (8.41)$$

Now taking into account deaths and births, we can write the size of the total population alive at time t as a function of the size of the population that was alive at some time τ in the past: $N_t = N_\tau e^{(n-p)(t-\tau)}$. It follows that

$$\frac{N_{t,\tau}}{N_t} = ne^{-p(t-\tau)} e^{-(n-p)(t-\tau)} = ne^{-n(t-\tau)}. \quad (8.42)$$

We conclude that the relative size at time t of the cohort born at τ is simply $ne^{-n(t-\tau)}$. For any variable $x_{t,\tau}$ define the per capita (or average) x_t as

$$\begin{aligned} x_t &= \int_{-\infty}^t x_{t,\tau} \left(\frac{N_{t,\tau}}{N_t} \right) d\tau \\ x_t &= \int_{-\infty}^t x_{t,\tau} ne^{-n(t-\tau)} d\tau. \end{aligned} \quad (8.43)$$

Applying this definition to individual consumption from (8.40) we have

$$c_t = (\rho + p) (a_t + h_t), \quad (8.44)$$

so that per capita consumption is proportional to per capita assets, where

$$a_t = \int_{-\infty}^t a_{t,\tau} ne^{-n(t-\tau)} d\tau, \quad (8.45)$$

and

$$h_t = \int_{-\infty}^t h_{t,\tau} ne^{-n(t-\tau)} d\tau, \quad (8.46)$$

are non-human and human wealth, respectively. Focus on each, beginning with human wealth, which using the expression for $h_{t,\tau}$ in (8.37) can be written as

$$h_t = \int_{-\infty}^t \left\{ \int_t^\infty y_{s,\tau} e^{-\int_t^s (r_v + p) dv} ds \right\} ne^{-n(t-\tau)} d\tau. \quad (8.47)$$

Now, if labour income is the same for all agents who are alive at some time s , we have

$$h_t = \int_{-\infty}^t \left\{ \int_t^\infty y_s e^{-\int_t^s (r_v + p) dv} ds \right\} ne^{-n(t-\tau)} d\tau, \quad (8.48)$$

where the expression in curly brackets is the same for all agents. It follows that

$$h_t = \int_t^\infty y_s e^{-\int_t^s (r_v + p) dv} ds. \quad (8.49)$$

Finally, differentiating with respect to time t (with the help of Leibniz's rule) we arrive at⁶

$$\dot{h}_t = (r_t + p) h_t - y_t, \quad (8.50)$$

which is the equation of motion for human capital. It can also be written as

$$r_t + p = \frac{\dot{h}_t + y_t}{h_t}. \quad (8.51)$$

This has our familiar, intuitive asset pricing interpretation. If we think of human capital as an asset, then the RHS is the return on this asset, including the capital gain \dot{h}_t and the dividend y_t , both

expressed in proportion to the value h_t of the asset. That has to be equal to the individual discount rate $r_t + p$, which appears on the LHS.

Turn next to the evolution of non-human wealth. Differentiating a_t , from (8.45), with respect to t (again using Leibniz's rule!) we have

$$\dot{a}_t = na_t + n \int_{-\infty}^t \left\{ -a_{t,\tau} n e^{-n(t-\tau)} + e^{-n(t-\tau)} \dot{a}_{t,\tau} \right\} d\tau, \quad (8.52)$$

$$\dot{a}_t = na_{t,0} - na_t + n \int_{-\infty}^t \dot{a}_{t,\tau} e^{-n(t-\tau)} d\tau,$$

since $a_{t,0}$ is non-human wealth at birth, which is zero for all cohorts, we have

$$\begin{aligned} \dot{a}_t &= -na_t - n \int_{-\infty}^t \dot{a}_{t,\tau} e^{-n(t-\tau)} d\tau, \\ \dot{a}_t &= -na_t + \int_{-\infty}^t \left\{ (r_t + p) a_{t,\tau} + y_t - c_{t,\tau} \right\} n e^{-n(t-\tau)} d\tau, \\ \dot{a}_t &= -na_t + (r_t + p) \int_{-\infty}^t a_{t,\tau} n e^{-n(t-\tau)} d\tau + y_t \int_{-\infty}^t n e^{-n(t-\tau)} d\tau - \int_{-\infty}^t c_{t,\tau} n e^{-n(t-\tau)} d\tau, \\ \dot{a}_t &= [r_t - (n - p)] a_t + y_t - c_t. \end{aligned} \quad (8.53)$$

Notice that while the individual the rate of return is $r_t + p$, for the economy as a whole the rate of return is only r_t , since the p is a transfer from people who die to those who remain alive, and washes out once we aggregate. Recall, however, that a_t is assets per capita, so naturally $(n - p)$, the rate of growth of population, must be subtracted from r_t .

The consumption function (8.40) and the laws of motion for per capita human and non-human wealth, (8.50) and (8.53), completely characterise the dynamic evolution of this economy. It can be expressed as a two-dimensional system in the following way. Differentiate the consumption function with respect to time in order to obtain

$$\dot{c}_t = (\rho + p) (\dot{a}_t + \dot{h}_t). \quad (8.54)$$

Next use the laws of motion for both forms of wealth to obtain

$$\dot{c}_t = (\rho + p) \left[(r_t - n + p) a_t - c_t + (r_t + p) h_t \right]. \quad (8.55)$$

Write the consumption function in the following way

$$h_t = \frac{c_t}{\rho + p} - a_t, \quad (8.56)$$

and use it to substitute out h_t from the \dot{c}_t equation (8.55):

$$\begin{aligned} \dot{c}_t &= (\rho + p) \left[(r_t - n + p) a_t - c_t - (r_t + p) a_t + \frac{r_t + p}{\rho + p} c_t \right], \\ \dot{c}_t &= (\rho + p) \left[-na_t + \frac{r_t - \rho}{\rho + p} c_t \right], \\ \dot{c}_t &= (r_t - \rho) c_t - n(p + \rho) a_t. \end{aligned} \quad (8.57)$$

This is a kind of modified Euler equation. The first term is standard, of course, but the second term is not. That second term comes from the fact that, at any instant, there are n newcomers for each person alive, and they dilute assets per capita by na_t since at birth they have no assets. This slows down the average rate of consumption growth.

This modified Euler equation plus the law of motion for non-human wealth (8.53) are a two-dimensional system of differential equations in c_t and a_t . That system, plus an initial condition and a transversality condition for a_t , fully describes the behaviour of the economy.

8.3.1 | The closed economy

We have not taken a stance on what kind of asset a_t is. We now do so. In the closed economy we assume that $a_t = k_t$, and k_t is per-capita productive capital that yields output according to the function $y_t = k_t^\alpha$, where $0 < \alpha < 1$. In this context profit maximisation dictates that $r_t = \alpha k_t^{\alpha-1}$, so that our two differential equations become

$$\begin{aligned}\dot{c}_t &= (\alpha k_t^{\alpha-1} - \rho) c_t - n(p + \rho) k_t, \\ \dot{k}_t &= (1 + \alpha) k_t^\alpha - (n - p) k_t - c_t.\end{aligned}\tag{8.58}$$

In steady state we have

$$\begin{aligned}\frac{c^*}{k^*} &= \frac{n(p+\rho)}{\alpha k^{*\alpha-1}-\rho}, \\ (1 + \alpha) k^{*\alpha-1} - (n - p) &= \frac{c^*}{k^*}.\end{aligned}\tag{8.59}$$

Combining the two yields

$$(1 + \alpha) k^{*\alpha-1} = (n - p) + \frac{n(p + \rho)}{\alpha k^{*\alpha-1} - \rho},\tag{8.60}$$

which pins down the capital stock. For given k^* , the first SS equation yields consumption. Rewrite the last equation as

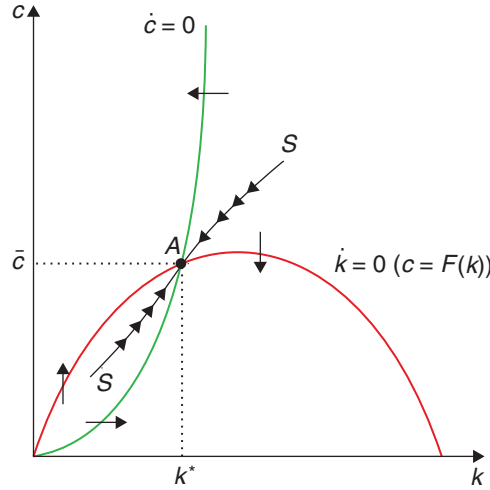
$$\alpha k^{*\alpha-1} - \rho = \frac{n(p + \rho)}{(1 + \alpha) k^{*\alpha-1} - (n - p)} > 0.\tag{8.61}$$

So the steady-state level of the (per capita) capital stock is smaller than the modified golden rule level that solves $\alpha k^{\alpha-1} = \rho$, implying under-accumulation of capital.⁷ This is in contrast to the NGM, in which the modified golden rule applies, and the discrete-time OLG model with two-period lives, in which over-accumulation may occur. Before examining that issue, consider dynamics, described in Figure 8.5.

Along the saddle-path c_t and k_t move together. If the initial condition is at $k > k^*$, then consumption will start above its SS level and both c_t and k_t will gradually fall until reaching the steady-state level. If, by contrast, the initial condition is at $k < k^*$, then consumption will start below its steady-state level and both c_t and k_t will rise gradually until reaching the steady state.

8.3.2 | A simple extension

But how come we have no dynamic inefficiency in this model? Just switching to continuous time does away with this crucial result? Not really. The actual reason is that the model so far is not quite like what we had before, in another aspect: there is no retirement! In contrast to the standard OLG model, individuals have a smooth stream of labour income throughout their lives, and hence do not need to save a great deal in order to provide for consumption later in life.

Figure 8.5 Capital accumulation in the continuous time OLG model

Introducing retirement (i.e. a stretch of time with no income, late in life) is analytically cumbersome, but as Blanchard (1985) demonstrates, there is an alternative that is easily modelled, has the same flavour, and delivers the same effects: assuming labour income declines gradually as long as an individual is alive.

Let's take a look. Blanchard (1985) assumes that each individual starts out with one unit of effective labour and thereafter his available labour declines at the rate $\gamma > 0$. At time t , the labour earnings of a person in the cohort born at τ is given by $w_t e^{-\gamma(t-\tau)}$, where w_t is the market wage per unit of effective labour at time t . It follows that individual human wealth for a member of the τ generation is

$$h_{t,\tau} = \int_t^\infty w_s e^{-\gamma(s-\tau)} e^{-\int_t^s (r_v + p) dv} ds. \quad (8.62)$$

Using the same derivation as in the baseline model, we arrive at a modified Euler equation

$$\dot{c}_t = (\alpha k_t^{\alpha-1} + \gamma - \rho) c_t - (n + \gamma)(p + \rho) k_t, \quad (8.63)$$

which now includes the parameter γ .

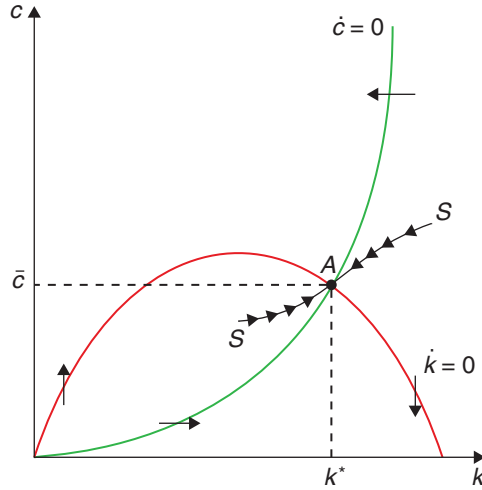
The steady state per-capita capital stock is now again pinned down by the expression

$$k^{*\alpha-1} = (n - p) + \frac{(n + \gamma)(p + \rho)}{\alpha + \gamma - \rho}, \quad (8.64)$$

which can be rewritten as

$$\alpha k^{*\alpha-1} - \rho = \frac{(n + \gamma)(p + \rho)}{k^{*\alpha-1} - (n - p)} - \gamma. \quad (8.65)$$

So if γ is sufficiently large, then the steady-state per capita capital stock can be larger than the golden rule level, which is the one that solves the equation $\alpha k^{\alpha-1} = \rho$. This would imply over-accumulation of capital. The intuition is that the declining path of labour income forces people to save more, too

Figure 8.6 Capital accumulation with retirement

much in fact. Again, intergenerational transfers would have been a more efficient way to pay for retirement, but they cannot happen in the decentralized equilibrium, in the absence of intergenerational altruism.

In this case, dynamics are given by Figure 8.6, with the steady state to the right of the modified golden-rule level of capital:

8.3.3 | Revisiting the current account in the open economy

We can also revisit the small open economy as a special case of interest. For that, let's go back to the case in which $\gamma = 0$, and consider what happens when the economy is open, and instead of being capital, the asset is a foreign bond f_t that pays the fixed world interest rate r . In turn, labour income is now, for simplicity, an exogenous endowment $y_{t,\tau} = y$ for all moments t and for all cohorts τ .

The two key differential equations now become

$$\begin{aligned}\dot{c}_t &= (r - \rho)c_t - n(p + \rho)f_t, \\ \dot{f}_t &= [r - (n - p)]f_t + y - c_t,\end{aligned}\tag{8.66}$$

with steady-state values

$$[r - (n - p)]f^* + y = c^*,\tag{8.67}$$

$$(r - \rho)c^* = n(p + \rho)f^*,\tag{8.68}$$

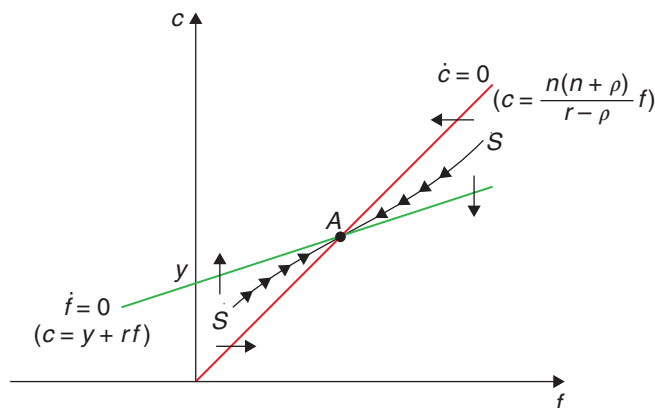
which together pin down the levels of consumption and foreign assets. The first equation reveals that in steady state the current account must be balanced, with consumption equal to endowment income plus interest earnings from foreign assets. As the second equation reveals, the steady-state stock of foreign assets can be positive or negative, depending on whether r is larger or smaller than ρ .

If $r > \rho$, individual consumption is always increasing, agents are accumulating over their lifetimes, and the steady-state level of foreign assets is positive. If $r = \rho$, individual consumption is flat and they neither save nor dissave; steady-state foreign assets are zero. Finally, if $r < \rho$, individual consumption is always falling, agents are decumulating over their lifetimes, and in the steady state the economy is a net debtor.

Equilibrium dynamics are given by Figure 8.7, drawn for the case $r > \rho$. It is easy to show that the system is saddle-path stable if $r < \rho + p$. So the diagram below corresponds to the case $\rho < r < \rho + p$. Along the saddle-path, the variables c_t and f_t move together until reaching the steady state.

In this model the economy does not jump to the steady state (as the open-economy model in Chapter 4 did). The difference is that new generations are constantly being born without any foreign assets and they need to accumulate them. The steady state is reached when the accumulation of the young offsets the decumulation of the older generation.

Figure 8.7 The current account in the continuous time OLG model



8.4 | What have we learned?

In this chapter we developed the second workhorse model of modern macroeconomics: the OLG model. This framework allows us to look at questions in which assuming a single representative agent is not a useful shortcut. We will see how this will enable us to tackle some key policy issues, starting in the next chapter.

Moreover, we have already shown how this model yields new insights about capital accumulation, relative to the NGM. For instance, the possibility of dynamic inefficiency – that is to say, of over-accumulation of capital – emerges. This is a result of the absence of intergenerational links, which entail that individuals may need to save too much, as it is the only way to meet their consumption needs as their labor income declines over their life cycle.

Notes

¹ If the production function makes the function hit the 45-degree line with a negative slope the model can give origin to cyclical behaviour around the steady-state. This cycle can be stable or unstable depending on the slope of the curve.

- ² The First Welfare Theorem can be extended to deal with an infinite number of agents, but this requires a condition that the total value of resources available to all agents taken together be finite (at equilibrium prices). This is not satisfied in the OLG economy, which lasts forever.
- ³ For those of you who are mathematically inclined, the argument is similar to Hilbert's Grand Hotel paradox. If the argument sounds counter-intuitive and esoteric, it's because it is – so much so that some people apparently think the paradox can be used to prove the existence of God! (see http://en.wikipedia.org/wiki/Hilbert's_paradox_of_the_Grand_Hotel).
- ⁴ Mixed income is that which is registered as accruing to capital, because it comes from the residual income of businesses, but that Geerolf argues should be better understood, at least partly as, returns to entrepreneurial labour. Land rents, which Abel et al. only had for the U.S., should not be understood as capital in their sense, as land cannot be accumulated.
- ⁵ Suppose, in addition, that the economy starts with a population $N_0 = 1$.
- ⁶ Leibniz's rule? Why, of course, you recall it from calculus: that's how you differentiate an integral. If you need a refresher, here it is: take a function $g(x) = \int_{a(x)}^{b(x)} f(x, s) ds$, the derivative of g with respect to x is: $\frac{dg}{dx} = f(x, b(x)) \frac{db}{dx} - f(x, a(x)) \frac{da}{dx} + \int_{a(x)}^{b(x)} \frac{df(x, s)}{dx} ds$. Intuitively, there are three components of the marginal impact of changing x on g : those of increasing the upper and lower limits of the integral (which are given by f evaluated at those limits), and that of changing the function f at every point between those limits (which is given by $\int_{a(x)}^{b(x)} \frac{df(x, s)}{dx} ds$). All the other stuff is what you get from your run-of-the-mill chain rule.
- ⁷ Because individuals discount the future ($\rho > 0$), this is not the same as the golden rule in the Solow model, which maximises consumption on the steady state. In the modified golden rule, the capital stock is smaller than that which maximises consumption, precisely because earlier consumption is preferred to later consumption.

References

- Abel, A. B., Mankiw, N. G., Summers, L. H., & Zeckhauser, R. J. (1989). Assessing dynamic efficiency: Theory and evidence. *The Review of Economic Studies*, 56(1), 1–19.
- Blanchard, O. (2019). Public debt and low interest rates. *American Economic Review*, 109(4), 1197–1229.
- Blanchard, O. J. (1985). Debt, deficits, and finite horizons. *Journal of Political Economy*, 93(2), 223–247.
- Diamond, P. A. (1965). National debt in a neoclassical growth model. *The American Economic Review*, 55(5), 1126–1150.
- Geerolf, F. (2013). Reassessing dynamic efficiency. *Manuscript, Toulouse School of Economics*.
- Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66(6), 467–482.

An application: Pension systems and transitions

Let us put our OLG framework to work in analysing the topic of pensions, a particularly suitable topic to be discussed using this framework. This is a pressing policy issue both in developed and developing countries, particularly in light of the ongoing demographic transition by which fewer working-age individuals will be around to provide for the obligations to retired individuals.

It is also a controversial policy issue because the question always looms as to whether people save enough for retirement on their own. Also, even though the models of the previous chapter suggested there may be instances in which it may be socially beneficial to implement intergenerational transfers such as pensions, this hinged on a context of dynamic inefficiency that was far from established. And then, if the economies are not dynamically inefficient, should the government interfere with the savings decisions of individuals? These are interesting but difficult policy questions. Particularly because it confronts us head-on with the difficulties of assessing welfare when there is no representative agent. Also, because, as we will see, once general equilibrium considerations are taken into account, sometimes things turn out exactly opposite to the way you may have thought they would!

So, let's tackle the basics of how pension systems affect individual savings behaviour and, eventually, capital accumulation. As in the previous chapter, the market economy is composed of individuals and firms. Individuals live for two periods (this assumption can easily be extended to allow many generations). They work for firms, receiving a wage, and also lend their savings to firms, receiving a rental rate. If there is a pension system, they make contributions and receive benefits as well.

9.1 | Fully funded and pay-as-you-go systems

There are two types of pension systems. In pay-as-you-go, the young are taxed to pay for retirement benefits. In the fully funded regimes, each generation saves for its own sake. The implications for capital accumulation are radically different.

Let d_t be the contribution of a young person at time t , and let b_t be the benefit received by an old person at time t . There are two alternative ways of organising and paying for pensions: *fully funded* and *pay-as-you-go*. We consider each in turn.

How to cite this book chapter:

Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide*.

Ch. 9. 'An application: Pension systems and transitions', pp. 135–146. London: LSE Press.

DOI: <https://doi.org/10.31389/lsepress.ame.i> License: CC-BY-NC 4.0.

Fully funded system Under a fully funded system, the contributions made when young are returned with interest when old:

$$b_{t+1} = (1 + r_{t+1})d_t. \quad (9.1)$$

This is because the contribution is invested in real assets at the ongoing interest rate.

Pay-as-you-go system Under a pay-as-you-go system, the contributions made by the current young go directly to the current old:

$$b_t = (1 + n)d_t. \quad (9.2)$$

The reason why population growth pops in is because if there is population growth there is a larger cohort contributing than receiving. Notice the subtle but critical change of subscript on the benefit on the left-hand side.

There are many questions that can be asked about the effects of such pension programs on the economy. Here we focus on only one: Do they affect savings, capital accumulation, and growth?¹

With pensions, the problem of an individual born at time t becomes

$$\max \log(c_{1t}) + (1 + \rho)^{-1} \log(c_{2t+1}), \quad (9.3)$$

subject to

$$c_{1t} + s_t + d_t = w_t, \quad (9.4)$$

$$c_{2t+1} = (1 + r_{t+1})s_t + b_{t+1}. \quad (9.5)$$

The first-order condition for a maximum is still the Euler equation

$$c_{2t+1} = \left(\frac{1 + r_{t+1}}{1 + \rho} \right) c_{1t}. \quad (9.6)$$

Substituting for c_{1t} and c_{2t+1} in terms of s , w , and r implies a saving function

$$s_t = \left(\frac{1}{2 + \rho} \right) w_t - \frac{(1 + r_{t+1})d_t + (1 + \rho)b_{t+1}}{(2 + \rho)(1 + r_{t+1})}. \quad (9.7)$$

Again, savings is an increasing function of wage income, and is a decreasing function of contributions and benefits – leaving aside the link between those, and the general equilibrium effects through factor prices. These will mean, however, that savings will be affected by the pension variables in a complicated way.

With Cobb-Douglas technology, the firm's rules for optimal behaviour are

$$r_t = \alpha k_t^{\alpha-1}, \quad (9.8)$$

and

$$w_t = (1 - \alpha) k_t^\alpha = (1 - \alpha) y_t. \quad (9.9)$$

9.1.1 | Fully funded pension system

Fully funded systems do not affect capital accumulation. What people save through the pension system they dissave in their private savings choice.

Let us start by looking at the effect of this kind of program on individual savings. (The distinction between individual and aggregate savings will become critical later on.) We can simply insert (9.1) into (9.7) to get

$$s_t = \left(\frac{1}{2 + \rho} \right) w_t - d_t. \quad (9.10)$$

Therefore,

$$\frac{\partial s_t}{\partial d_t} = -1. \quad (9.11)$$

In words, holding the wage constant, pension contributions decrease private savings exactly one for one. The intuition is that the pension system provides a rate of return equal to that of private savings, so it is as if the system were taking part of that individual's income and investing that amount itself. The individual is indifferent about who does the saving, caring only about the rate of return.

Hence, including the pension savings in total savings, a change in contributions d leaves overall, or aggregate savings (and, therefore, capital accumulation and growth) unchanged. To make this clear, let's define aggregate savings as the saving that is done privately plus through the pension system. In a fully funded system the aggregate savings equals

$$s_t^{agg} = s_t + d_t = \left(\frac{1}{2 + \rho} \right) w_t. \quad (9.12)$$

This is exactly the same as in Chapter 7, without pensions.

9.1.2 | Pay-as-you-go pension system

Pay-as-you-go pension schemes reduce the capital stock of the economy.

To see the effect of this program on savings, insert (9.2) into (9.7) (paying attention to the appropriate time subscripts) to get

$$s_t = \left(\frac{1}{2 + \rho} \right) w_t - \frac{(1 + r_{t+1}) d_t + (1 + \rho)(1 + n) d_{t+1}}{(2 + \rho)(1 + r_{t+1})}. \quad (9.13)$$

This is a rather complicated expression that depends on d_t and d_{t+1} – that is, on the size of the contributions made by each generation. But there is one case that lends itself to a simple interpretation. Assume $d_t = d_{t+1} = d$, so that contributions are the same per generation. Then equation (9.13) becomes

$$s_t = \left(\frac{1}{2 + \rho} \right) w_t - d \left[\frac{(1 + r_{t+1}) + (1 + \rho)(1 + n)}{(2 + \rho)(1 + r_{t+1})} \right]. \quad (9.14)$$

Note that, from an individual's perspective, the return on her contributions is given by n , and not r . This return depends on there being more individuals to make contributions to the pension system in each period – you can see how demographic dynamics play a crucial role here!

From (9.14) we have

$$\frac{\partial s_t}{\partial d_t} = -\frac{(1 + r_{t+1}) + (1 + \rho)(1 + n)}{(2 + \rho)(1 + r_{t+1})} < 0. \quad (9.15)$$

We can see contributions decrease individual savings – and, in principle, aggregate savings, as here they coincide (see the caveat below). Why do private and aggregate savings coincide? Because the pension system here is a transfer scheme from young to old, and not an alternative savings scheme. The only source of capital is private savings s_t .

9.1.3 | How do pensions affect the capital stock?

So far we have asked what happens to savings holding interest and wages constant – that is to say, the partial equilibrium effect of pensions. In the case of a fully funded system, that is of no consequence, since changes in contributions leave savings – and hence, capital accumulation, wages, and interest rates – unchanged. But it matters in the case of a pay-as-you-go system.

To examine the general equilibrium effects of changes in contributions within the latter system, recall that capital accumulation is given by

$$k_{t+1} = \frac{s_t}{1 + n}. \quad (9.16)$$

Substituting (9.14) into this equation we have

$$k_{t+1} = \left(\frac{1}{2 + \rho} \right) \frac{w_t}{1 + n} - h(k_{t+1})d, \quad (9.17)$$

where

$$h(k_{t+1}) = \frac{1 + (1 + \rho)(1 + n)(1 + r_{t+1})^{-1}}{(1 + n)(2 + \rho)}, \quad (9.18)$$

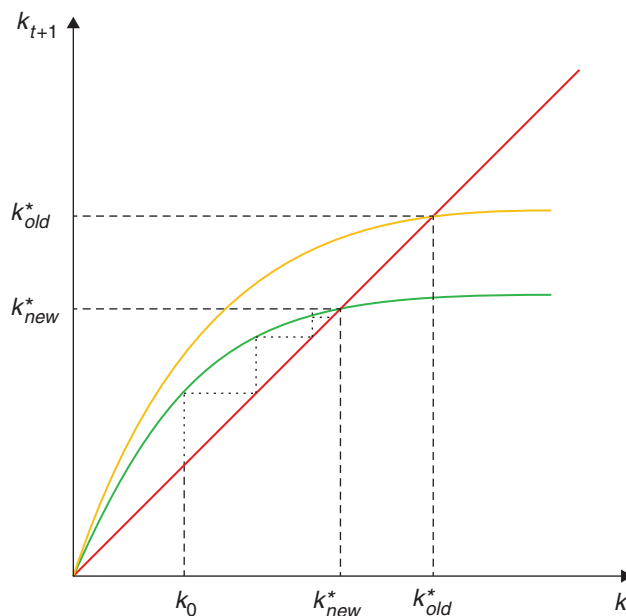
$$= \frac{1 + (1 + \rho)(1 + n)(1 + \alpha k_{t+1}^{\alpha-1})^{-1}}{(1 + n)(2 + \rho)}, \quad (9.19)$$

and where $h'(k_{t+1}) > 0$. (Note the use of (9.8) above.)

Next, totally differentiating (9.17), holding k_t constant, and rearranging, we have

$$\frac{dk_{t+1}}{dd_t} = -\frac{h(k_{t+1})}{1 + h'(k_{t+1})d} < 0. \quad (9.20)$$

Therefore, the effect of an increase in contributions in a pay-as-you-go system is to shift down the savings locus. The consequences appear in Figure 9.1. The new steady-state capital stock is lower. If the capital stock at the time of the policy shock is to the left of the new steady state, the economy continues to accumulate capital, but at a rate slower than before the change.

Figure 9.1 Introduction of pay-as-you-go social security

9.1.4 | Pensions and welfare

Is this a desirable outcome? Does it raise or lower welfare? Suppose before the change $d_t = 0$, so the change amounts to introducing pensions in a pay-as-you-go manner. Who is better off as a result?

The old at time t , who now receive total benefits equal to $(1+n)d_t$ and contribute nothing, are clearly better off. What about other generations? If r was less than n before the introduction of pensions, then the policy change reduces (perhaps totally eliminates) dynamic inefficiency, and all other generations benefit as well. In that case, introducing pensions is Pareto improving. The recent work that we saw in the last chapter suggests that this possibility is not as remote as one may have previously thought. In fact, this idea has coloured some recent policy thinking about reform in places like China.²

But if r is equal to or larger than n before the introduction of the pension system, then the policy change creates a conflict. The old at time t still benefit, but other generations are worse off. In this case, introducing pensions is not Pareto improving. Even if that is the case, this by no means implies that it is always a bad idea politically, or even that it is always socially undesirable. The point is that there will be winners and losers, and the relative gains and losses will have to be weighed against one another somehow.

9.2 | Moving out of a pay-as-you-go system

The effects on the capital stock from transitioning from a pay-as-you-go system to a fully funded system depend on how the transition is financed. If it is financed with taxes on the young, the capital stock increases. If it is funded by issuing debt, the capital stock may decrease.

There are several transitions associated with the introduction or revamping of pensions systems, and that we may want to analyze. For example, you could move from no pension system and implement a full capitalisation system. As aggregate saving behaviour does not change, we do not expect anything really meaningful to happen from such change in terms of capital accumulation and growth. (That is, of course, to the extent that rational behaviour is a good enough assumption when it comes to individual savings behaviour. We will get back to this pretty soon when we talk about consumption.) Alternatively, as discussed above, if we implement a pay-as-you-go system, the initial old are happy, while the effect for future generations remains indeterminate and depends on the dynamic efficiency of the economy.

However, in recent years it has become fashionable to move away from pay-as-you-go systems to fully funded ones. The reasons for such change is different in each country, but usually can be traced back to deficit and sometimes insolvent systems (sometimes corruption-ridden) that need to be revamped.³ But one of the main reasons was to undo the capital depletion associated with pay-as-you-go systems. Thus, these countries hoped that going for a capitalisation system would increase the capital stock and income over time.

In what remains of this chapter we will show that what happens in such transitions from pay-as-you-go to fully funded systems depends very much on how the transition is financed. There are two options: either the transition is financed by taxing the current young, or it is financed by issuing debt. Both have quite different implications.

To make the analysis simple, in what follows we will keep $n = 0$. (Note that this puts us in the region where $r > n$, i.e. that of dynamic efficiency.) Aggregate savings without pensions or with a fully funded system are

$$s_t^{agg} = \left(\frac{1}{2 + \rho} \right) w_t. \quad (9.21)$$

With a pay-as-you-go system, they are

$$s_t^{agg} = s_t = \left(\frac{1}{2 + \rho} \right) w_t - \frac{(1 + r_{t+1})d + (1 + \rho)d}{(2 + \rho)(1 + r_{t+1})}, \quad (9.22)$$

which is trivially lower (we knew this already). So now the question is how savings move when going from a pay-as-you-go to a fully funded system. You may think they have to go up, but we need to be careful: we need to take care of the old, who naturally will not be part of the new system, and their retirement income has to be financed. This, in turn, may have effects of its own on capital accumulation.

9.2.1 | Financing the transition with taxes on the young

If the transition is financed out of taxes, the young have to use their wages for consumption (c_{1t}), private savings (s_t), to pay for their contributions (d and also for taxes τ_t):

$$c_{1t} + s_t + d + \tau_t = w_t. \quad (9.23)$$

Future consumption is in turn given by

$$c_{2t+1} = (1 + r_{t+1})s_t + (1 + r_{t+1})d, \quad (9.24)$$

as we are in a fully funded system. Because taxes here are charged to finance the old, we have $\tau_t = d$ (remember we have assumed population growth to be equal to zero). If you follow the logic above, it can be shown that in this case we have

$$s_t^{agg} = \frac{(w_t - \tau_t)}{(2 + \rho)}. \quad (9.25)$$

You may notice that this is lower than the steady-state savings rate (next period, i.e. in 30 years, there are no more taxes), but you can also show that it is higher than in the pay-as-you-go system. To do so, replace τ_t with d in (9.25) and then compare the resulting expression with that of (9.22).

So savings goes up slowly, approaching its steady-state value. These dynamics are what supports World Bank recommendations that countries should move from pay-as-you-go to fully capitalised systems. Notice however that the reform hurts the current young that have to save for their own and for the current old generation. Then remember that one period here is actually one generation, so it's something like 30 years. What do you think would be the political incentives, as far as reforming the system, along those lines?

9.2.2 | Financing the transition by issuing debt

Now let's think about how things would change if the transition is financed by issuing debt. (Maybe that is a politically more palatable option!) In this case, for the current young there are no taxes, and debt is another asset that they can purchase:

$$c_{1t} + s_t + d + g_{debt} = w_t, \quad (9.26)$$

so consumption in old age can be

$$c_{2t+1} = (1 + r_{t+1}) s_t + (1 + r_{t+1}) d + (1 + r_{t+1}) g_{debt}. \quad (9.27)$$

Following the same logic as before, private savings are

$$s_t = \frac{w_t}{(2 + \rho)} - d - g_{debt}. \quad (9.28)$$

How about aggregate savings? Note that contributions to the fully funded system d , work as savings from an aggregate perspective: they are available to finance the accumulation of capital. However, the amount of debt issued by the government is in fact not used for capital accumulation, but rather for consumption, because it is a transfer to the old. As such, aggregate savings are given by

$$s_t^{agg} = s_t + d = \frac{w_t}{(2 + \rho)} - g_{debt} = \frac{w_t}{(2 + \rho)} - d, \quad (9.29)$$

where in the last step we use the fact that (under no population growth) the government issues $g_{debt} = d$ of debt to pay benefits to the current old.

Let's see how this compares to the pay-as-you-go savings. Rewriting equation (9.22) which shows the savings rate in a pay-as-you-go system

$$s_t^{agg} = s_t = \left(\frac{1}{2 + \rho} \right) w_t - d \frac{(1 + r_{t+1}) + (1 + \rho)}{(2 + \rho)(1 + r_{t+1})}. \quad (9.30)$$

Notice that if

$$\frac{(1 + r_{t+1}) + (1 + \rho)}{(2 + \rho)(1 + r_{t+1})} < 1, \quad (9.31)$$

then in this case savings is even lower than in the pay-as-you-go system, which happens because the government now pays r on its debt, which in this case is higher than n .

Another way to see this is if the government imposed a fully funded system but then makes the pension firms purchase government debt that is used for the current old (i.e. for consumption). There is no way this type of reform can increase the capital stock.

9.2.3 | Discussion

The above discussion embodies the dimensions of intergenerational equity, the potential efficiency effects, and also the importance of how policies are implemented. Moving from a pay-as-you-go system to a fully funded one is not immune to the way the transition is financed. This should capture your attention: you need to work out fully the effects of policies!

Pension reform has been an important debate in developed and developing countries alike. In the 1990s there was an emerging consensus that moving to a fully funded system would be instrumental in the development of local capital markets. This view triggered reforms in many countries. Here we want to discuss two cases that turned out very different: those of Argentina and Chile.⁴

Chile, for many years, was considered the poster-child for this reform. It implemented a change to a fully funded system in 1980. Furthermore, this was done at a time of fiscal consolidation. In the framework of the previous section, this is akin to the current working-age generation saving for their own retirement, as well as to pay for their contemporaneous old. As the theory suggested, the resources were deployed into investment, the savings rate, and Chile had a successful growth spurt, which many observers associated with the reform.

Argentina, on the other hand, also migrated to a fully funded system, but rather than streamlining the budget, the deficit increased. In fact, it increased by an amount not very different from the loss in pension funds that were now going to private accounts. In the framework of the previous section, this is akin to financing the transition with debt.

As we saw, in this case the reform reduces savings and, in fact, there was no discernible development of Argentine capital markets. The inflow of contributions into the pension funds went directly to buy government debt. But it was even worse: the bloating of the deficit increased the government's explicit debt. Of course, the counterpart was a reduction in future pension liabilities. But the market was not persuaded, and in 2001 Argentina plunged into a debt crisis. Many observers associated this macroeconomic crisis to the pension reform. A few years later, Argentina renationalized the pension system, moving away from a fully funded approach. The temptation to do so was big. The current generation basically ate up the accumulated, albeit little, capital stock, again, as predicted in our simple OLG specification.

While the contrast with Chile could not be starker, the system there eventually also came under attack. The math is simple. If the return to investments is 5%, an individual that contributes 10% of her wage to a pension system for say, 42 years, and has an expected pension life of 25 years, can actually obtain a replacement ratio of close to 100% (the exact number is 96%). But reality turned out to be far from that. When looking back at the evidence, the average retirement age in Chile has been around 62 years, and the pension life 25 years. However, people reach retirement with, on average, 20 years of contributions, not 42. This allows for a replacement ratio of only 24%. It is this low replacement ratio that has been the focus of complaints. Thus, most of these attempts eventually introduced some sort of low income protection for the elderly.

9.2.4 | Do people save enough?

The above setup assumes that agents optimise their savings to maximise intertemporal utility, with or without pensions. However, there are several reasons why this may not be the case. We will get into more detail in Chapters 11 and 12, but it is worth going over some of the possibilities here, in the context of pensions.

First and foremost, people may believe that if they don't save, the government will step in and bail them out. If you believe that, why would you save? This time inconsistency feature (the government may tell you to save but, if you don't, will feel tempted to help you out) may lead to suboptimal savings.

Even in the case of the U.S., where these considerations may not be so relevant, there has been ample discussion about the intensity with which people save for retirement. On the one hand, Scholz et al. (2006) shows that the accumulation of assets of people roughly matches the life cycle hypothesis (which we will see in detail in Chapter 11); on the other hand, there is evidence that suggests that consumption levels drop upon retirement (Bernheim et al. 2001), which is inconsistent with optimal savings. One possible reconciliation of these two facts is given by including the dimension of what type of assets people use for savings. Beshears et al. (2018) show that people save sizable amounts, but they tend to save in *illiquid* assets. Illiquid assets may provide unusually high returns (for example, owning your house provides steady rental income). Kaplan et al. (2014) estimate that housing services provides an after-tax risk adjusted rate of return of close to 8%. In such a world agents hold a large share of illiquid assets but consumption tracks income while they use some potentially expensive mechanisms to partially smooth consumption.

Present bias has also been mentioned as a reason why people tend to save less than the optimal level. In this case, imposing a pension system that forces people to save may be a *ex ante* optimal commitment device. We will discuss present bias in detail in Chapter 12.

Finally, recent research on savings for retirement has delivered some interesting new ideas and policy suggestions. One typical way of saving in the U.S. is the 401K programs, where you save with the benefit of a tax deferral: your income is taxed when withdrawing the funds. These programs are typically arranged with your employer, which matches the contributions with a vesting period to entice labour stability. Yet it has been found that matching is a fairly inefficient way to stimulate savings. Madrian (2013) finds that a matching contribution of 25% increases savings by 5%. In contrast default setting seems, to have a much stronger effect. Madrian and Shea (2001) show that when a company shifted from a default where, unless the worker would opt out, it would start contributing 3% of its salary to a 401K program, they found that fifteen months after the change, 85% of the workers participated, and 65% contributed 3% of their wages. This compared with only 49% participation for those workers hired previously in which only 4% contributed 3%. In short, default standards may be powerful (and cheap) tools for incentivizing savings.

9.3 | What have we learned?

In this chapter we applied the standard OLG model to study the issue of social security and pensions. We saw that the implications for capital accumulation can vary dramatically depending on the nature of the system. While fully-funded systems simply offer an alternative mechanisms to private savings, pay-as-you-go systems are essentially intergenerational transfers. These reduce the incentive for private savings, and reduce capital accumulation.

That said, we have also seen that the effects of policy reforms hinge very dramatically on implementation details. For instance, transitioning from pay-as-you-go to a fully funded system can even reduce capital accumulation, if the transition is financed with debt.

Very importantly, the welfare effects of policy interventions are hard to pin down, because there will be winners and losers. Things get even harder if we depart from fully rational behaviour. This also means that the political incentives can be tricky, and they must be taken into consideration.

9.4 | What next?

As we anticipated, the OLG framework has become increasingly used in macroeconomics. Many years ago, Auerbach and Kotlikoff (1987) provided an extension of this model to, realistically, allow for 50 generations of working age population. That model became the starting point of a more policy-oriented simulations, which were mostly applied to discussing taxations issues. Azariadis (1993) summarised our knowledge of these models, and is a good starting point for those interested in reviewing standard applications in macroeconomic theory, and understanding the potential of the OLG model to discuss business fluctuations and monetary policy. Ljungqvist and Sargent (2018) provide a more recent update.

But the interesting action has to do with the applications of the OLG model as a workhorse from modern macroeconomics in the age of low interest rates. As we will look into this in later chapters, we defer to the bibliography on this until then.

Notes

- ¹ See Feldstein and Bacchetta (1991) for a good non-technical introduction to some of the other issues, including distribution, risk, and labour market implications.
- ² As an example, check out this headline: ‘China hopes social safety net will push its citizens to consume more, save less’ (*Washington Post*, July 14, 2010).
- ³ Chile is perhaps the best-known example, with its pioneering move in the early 1980s. (See also Feldstein’s discussion.) For a discussion of the real-world pitfalls, Google this NYT article from January 2006: “Chile’s Candidates Agree to Agree on Pension Woes”.
- ⁴ Maybe because two of us are from Argentina and Chile?

References

- Auerbach, A. J. & Kotlikoff, L. J. (1987). Evaluating fiscal policy with a dynamic simulation model. *The American Economic Review*, 77(2), 49–55.
- Azariadis, C. (1993). Intertemporal macroeconomics.
- Bernheim, B. D., Skinner, J., & Weinberg, S. (2001). What accounts for the variation in retirement wealth among U.S. households? *American Economic Review*, 91(4), 832–857.
- Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2018). Behavioral household finance. *Handbook of behavioral economics: Applications and foundations 1* (pp. 177–276). Elsevier.
- Feldstein, M. & Bacchetta, P. (1991). National saving and international investment. *National saving and economic performance* (pp. 201–226). University of Chicago Press.
- Kaplan, G., Violante, G. L., & Weidner, J. (2014). *The wealthy hand-to-mouth*. (tech. rep.). National Bureau of Economic Research.

- Ljungqvist, L. & Sargent, T. J. (2018). *Recursive macroeconomic theory*. MIT Press.
- Madrian, B. C. & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics*, 116(4), 1149–1187.
- Scholz, J. K., Seshadri, A., & Khitatrakun, S. (2006). Are Americans saving “optimally” for retirement? *Journal of Political Economy*, 114(4), 607–643.

Unified growth theory

You will recall that among the key stylised facts we set out to explain in our study of economic growth was the very existence of growth in living standards: output per worker increases over time. This, however, has only been true for a very short span in human history, starting with the Industrial Revolution and the birth of modern economic growth.

For most of history, the prevailing situation was one that we may call Malthusian stagnation. In other words, there obviously were massive increases in productivity – the wheel, agriculture, domesticated animals, ships, double-entry bookkeeping – but these did not really translate into increased living standards, or into sustained productivity growth. Instead, as per the Malthusian assumption (recall our discussion of the Kremer (1993) paper in (Chapter 6), those increases in productivity mostly translated into increases in population.

Then at some point, around the 18th century, the Great Divergence happened (see Figure 10.1): a few Western European countries, and then the Western offshoots in the New World, took off and never looked back. They inaugurated the age of sustained economic growth, which eventually spread to most other countries around the world, and which has been the object of our study in this first part of the book.

But this begs the question: how did that transition happen? Is there any way in which we can understand within a single framework the growth process in the Malthusian and modern eras *combined*? Can we understand why the latter emerged in the first place, how, from stability, suddenly growth popped up?¹ This is the object of unified growth theory, a somewhat grandiosely named attempt at understanding growth from the perspective of millennia. This chapter constitutes a brief introduction to these ideas, following the presentation in Galor (2005). The fact that it uses some features of the OLG model explains why we have this discussion now.

10.1 | From Malthus to growth

Growth seems to have experienced a kink around the end of the 18th century, when it accelerated dramatically.

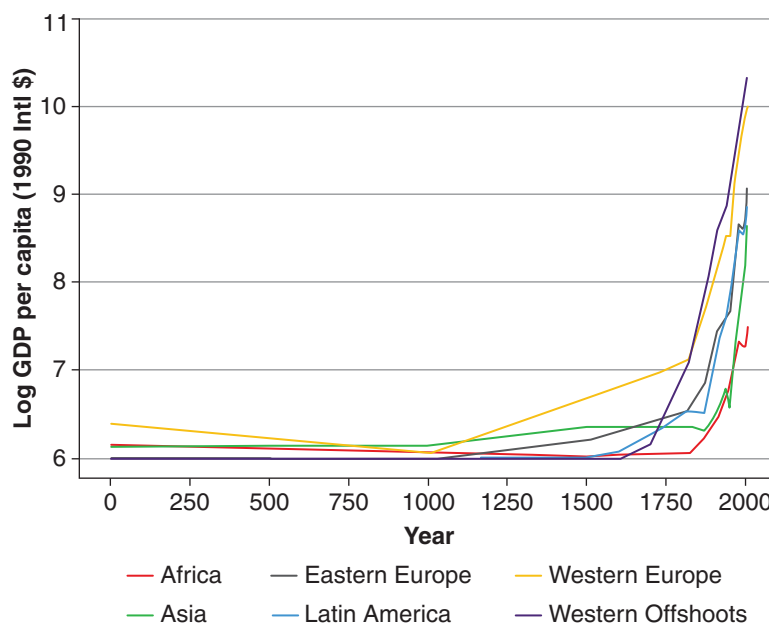
While we have argued above that technological progress and increases in productivity have been features of human history for millennia, it is pretty much undeniable that the pace at which such progress took place until the last couple of centuries was much, much slower than what we have come to expect

How to cite this book chapter:

Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide*.

Ch. 10. 'Unified growth theory', pp. 147–158. London: LSE Press.

DOI: <https://doi.org/10.31389/lsepress.ame.j> License: CC-BY-NC 4.0.

Figure 10.1 The evolution of regional income per capita over the years 1–2008

in recent times. The slow pace of productivity increases, combined with the Malthusian effect of those increases on population growth, meant that output per worker (i.e. living standards) grew very slowly over this period. Put simply, whenever things got a little better, people would starve less, live a little longer, have more kids, more of whom would reach adult age, etc. We would then simply end up sharing the increased output among a larger number of people, with essentially the same amount going for each of these people as before things had gotten better. In other words, increases in income per capita were essentially temporary, and eventually dissipated. It may sound odd to us today, but these were times when parents and sons lived in the same world, and so did their children and grandchildren. Progress was just too slow to be noticeable.

Mind you, stagnation doesn't mean that these were boring times when nothing particularly interesting happened. Much to the contrary, hovering around subsistence level meant that negative productivity shocks (say, droughts) were devastating. The flipside of the adjustment described above would clearly hold: bad shocks would mean people dying early, fewer kids reaching adulthood, etc., bringing population down.

By the same token, there were substantial differences across countries and regions. Naturally, some places were more productive (fertile land, natural resources, location) and technology varied substantially. However, more productive countries or regions were not richer on a per capita basis – they just had more people, as in China or the Indian subcontinent. In sum, this scenario had a natural implication for cross-country comparisons: population (or more precisely, population density) is the right measure of development in this era.

As such, how do we know that, from a global perspective, technological progress and economic performance were rather unimpressive? Because global population growth was very slow over the period, according to the best historical and archaeological evidence.

10.1.1 | The post-Malthusian regime

Then industrialisation happened.² Gradually, first picking up steam (figuratively, but also literally!) in England in the mid- to late-18th century, and spreading over Western Europe and the U.S. through the 19th century, the growth rate of output per capita started a sustained increase by an order of magnitude. Eventually, this reached most other places in the world somewhere along the 20th century. Still, this initial takeoff was marked by the Great Divergence – a testament to the power of growth rates in changing living standards over time. The first countries to industrialise started to grow richer and richer, leaving behind the laggards (Figure 10.2). It also brought along a marked increase in urbanisation, with people flocking from rural areas to the more dynamic urban centers, which could now sustain substantially larger (and increasing) populations.

Still, this is not what we would call the full-on, modern, sustained growth regime. Why? Because a remnant of the Malthusian past remained. There remained a positive link between increased productivity and income per capita, on one side, and increased population on the other, as can be seen readily, for different parts of the world, in Figure 10.3. As such, part of the sustained increase in productivity growth was still being dissipated over more capitas.

However, a very important transition was taking place at the same time: the rise of human capital. The acceleration in productivity growth and income per capita was accompanied by rising literacy rates, schooling achievement, and improving health. This increase in human capital seems to be driven

Figure 10.2 The differential timing of the take-off across regions

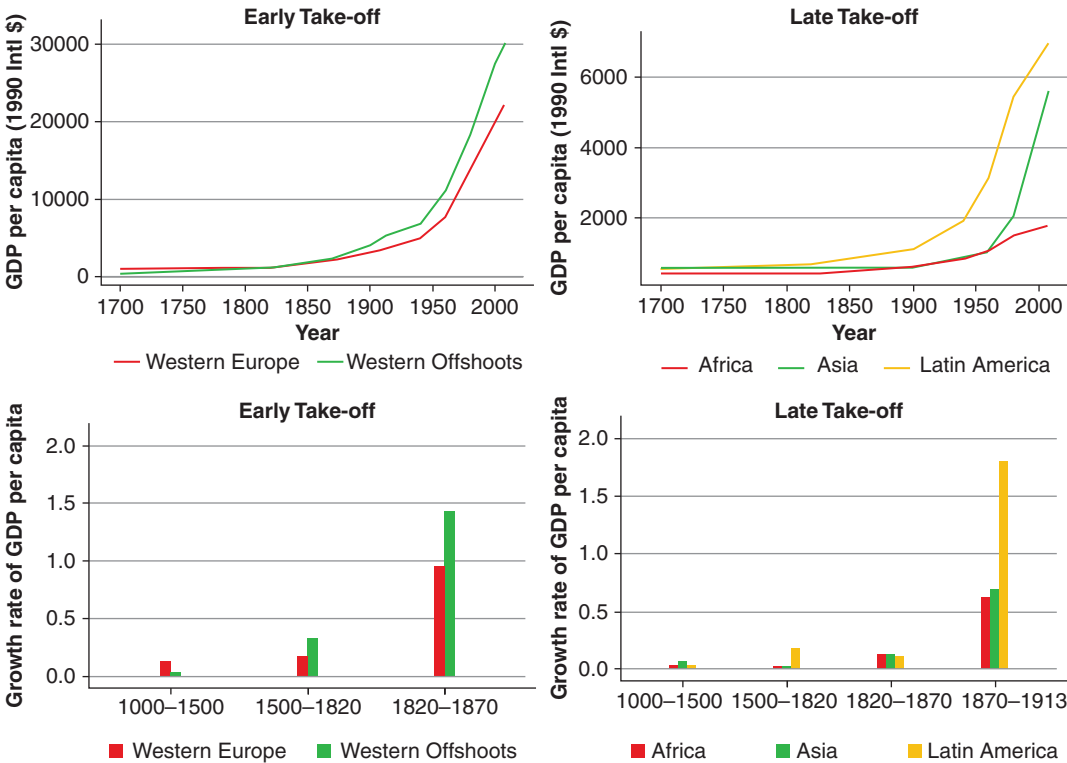
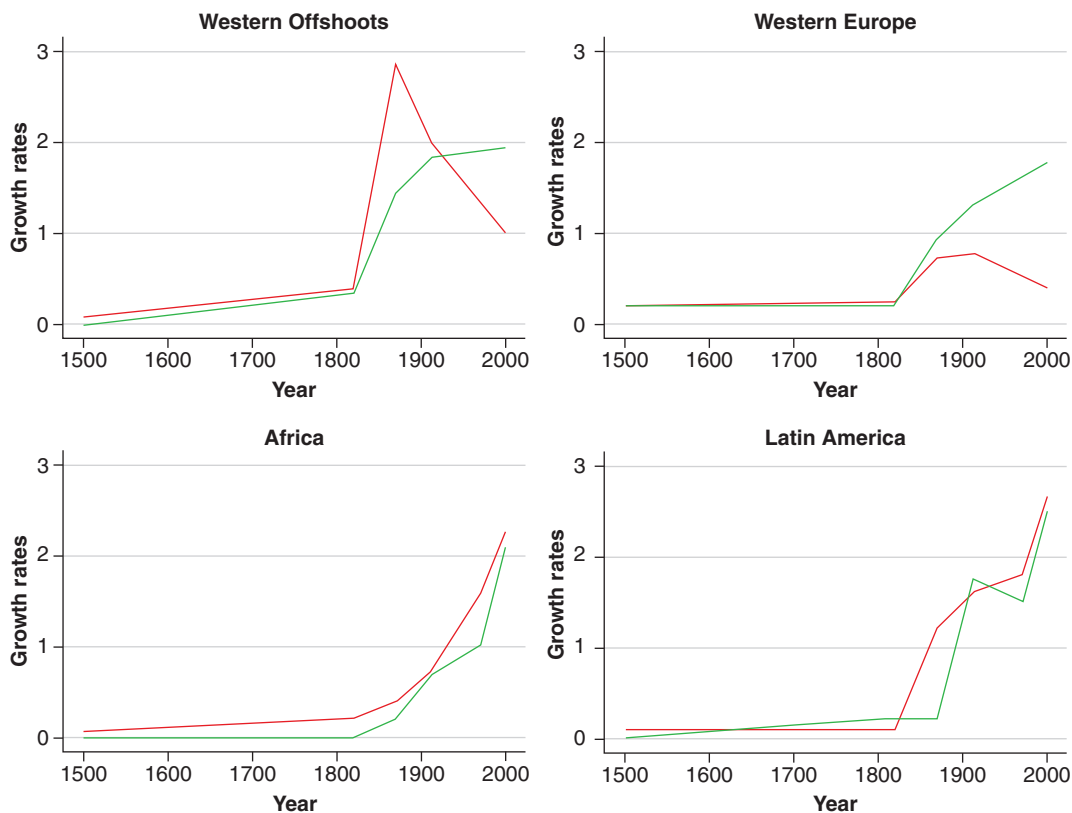


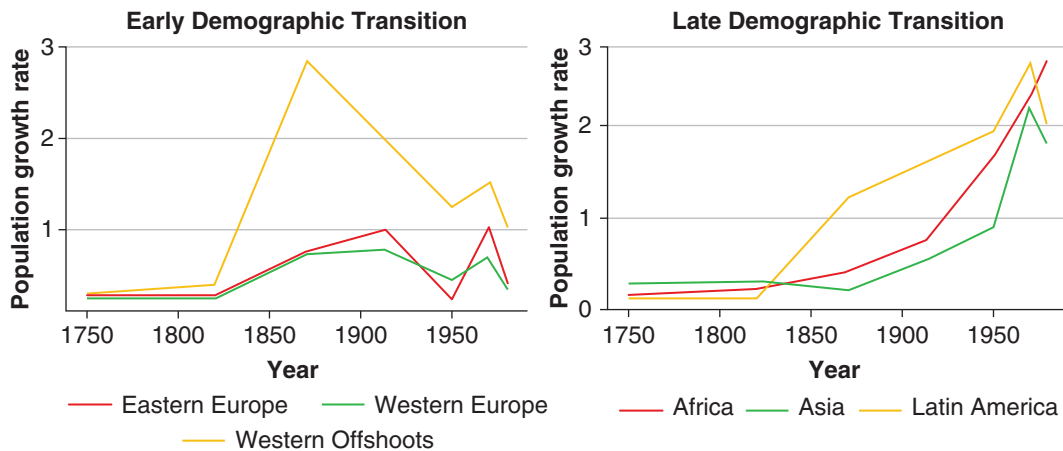
Figure 10.3 Regional growth of GDP per capita (green line) and population (red line) 1500–2000

by the fact that the industrialisation process, particularly in its later stages, requires more and more “brain” relative to “brawn” – an educated, decently healthy workforce becomes a *sine qua non*.

10.1.2 | Sustained economic growth

Then the demographic transition happened. At some point, human societies escaped definitively from the Malthusian shackles. Population growth ceased to be positively related to income per capita, and the relationship was actually reversed, with sharp declines in fertility rates. Even from a basic arithmetic perspective, this opens up the way for a historically astonishing rate of increase in living standards. In fact, the regions that first went into this transition reach a sustained speed of about 2% annual growth in income per capita over the last century or so – a rate at which living standards double in the space of one generation (35 years). Simultaneously, the relative importance of human capital increased even further, which was met by the first efforts in mass public education.

The demographic transition, i.e. the decline in fertility rates (accompanied by lower mortality and higher life expectancy), happened first in the leading industrialised nations, but then eventually reached the latecomers, as illustrated in Figure 10.4. (The exception for the moment, as we can see, is still Africa.) It actually had a three-fold impact on the growth process. First, and most obviously,

Figure 10.4 The differential timing of the demographic transition across regions

it ceased the dilution of capital (and land). Second, it actually enabled the increased investment in human capital. Put simply, the tradeoff between quantity of children and the quality of the investment made in each one of them turned decisively towards the latter. Third, and more temporarily, it yielded a demographic dividend: a relatively large labour force, with relatively few young dependents.

10.2 | A “unified” theory

Unified growth theory attempts to provide a single framework that can explain why growth may accelerate. One such mechanism relates to family decisions, i.e. whether to raise more kids or to raise fewer but more educated kids. Growth will force parents to educate their children at the expense of quantity of children, setting off an accelerating pattern.

The challenge is thus to come up with a framework that encompasses all the stylised facts spelled out in the previous section. In other words, we need a theory that has room for the Malthusian era, but then explains (i.e. endogenously generates) the transition to a post-Malthusian equilibrium with sustained productivity growth brought about by industrialisation and the attending increased importance of human capital. It must also account for the demographic transition and the possibility of sustained increases in living standards.

One could have told a story in which we combine the neoclassical growth model (which, after all, has zero growth in living standards in the long run) with the eventual takeoff of productivity being accounted for by the world of the endogenous growth models. However, how do we explain when and why we would move from one world to the other? And how do we incorporate the demographic aspect that seems to be such a central part of the story, and which lies thoroughly outside of the theories that we have seen so far?

We will start by sketching a model of parental investment that explains the links between demographic trends and productivity, while keeping the evolution of productivity exogenous. We will then endogenise productivity growth in order to complete our full very-long-run picture.

10.2.1 | A simple model of the demographic transition

Consider a discrete-time framework in which individuals live for two periods, as in our basic OLG model, but now each individual, instead of coming into the economy out of nowhere, has a single parent. In the first period of life (childhood), individuals consume a fraction of their parent's endowment of time (normalised to 1). This fraction will have to be greater to increase the child's quality. In the second period of life (parenthood), individuals are endowed with one unit of time, which they allocate between child-rearing and labour force participation. At this stage, the individual's problem is to choose the optimal mixture of quantity and quality of (surviving) children and supply their remaining time in the labour market, consuming their wages.

Let us consider the building blocks of this model in order.

Production

We assume that output is produced using land (exogenous and constant over time) and (efficiency units of) labour. We capture this using a constant returns to scale (CRS) production function:

$$Y_t = H_t^\alpha (A_t X)^{1-\alpha}, \quad (10.1)$$

where Y_t is total output, H_t is total efficiency units of labour, A_t is (endogenous) productivity, and X is land employed in production. We can write this in per worker terms as

$$y_t = h_t^\alpha x_t^{1-\alpha}, \quad (10.2)$$

where $y_t \equiv \frac{Y_t}{L_t}$, $h_t \equiv \frac{H_t}{L_t}$, and $x_t \equiv \frac{A_t X}{L_t}$ (we can think of x_t as effective (land) resources per worker).

Preferences

An individual who is a parent in time t cares about consumption c_t , and also about the number of children n_t , and their quality h_{t+1} . We summarise this in the following utility function:

$$u^t = c_t^{1-\gamma} (n_t h_{t+1})^\gamma. \quad (10.3)$$

We assume that individuals need to consume at least a subsistence level \bar{c} , which will be very important for the results. Children are passive recipients of their parents' decisions. This means that this is not a full-fledged OLG model, such as the ones we have seen before, in which young and old made meaningful economic decisions, which in turn entailed intergenerational conflicts of interest. The OLG structure here is more meant to highlight the demographic structure of the problem.

Budget constraint

Let τ be the amount of time needed to raise a child, regardless of quality. Any additional quality that an individual parent in time t bestows upon each child (to be reflected in the quality h_{t+1}) requires some additional effort, e_{t+1} (education). Whatever is left of the unit time endowment will be supplied in the labour market, in exchange for a wage of w_t (per efficiency unit of labour). We assume that there are no property rights over land, so that the individual rate of return on that is zero.

For an individual who was given quality h_t by her parent, we thus have

$$c_t \leq w_t h_t (1 - n_t \tau - n_t e_{t+1}). \quad (10.4)$$

This simply states that the individual can consume up to the level of income that she obtains from paid work.

Note that we can rearrange this budget constraint as follows:

$$c_t + w_t h_t n_t (\tau + e_{t+1}) \leq w_t h_t. \quad (10.5)$$

The RHS of this inequality, which we may call $z_t \equiv w_t h_t$, corresponds to total potential income, i.e. the income that the individual would obtain if all of her time were devoted to paid work. We can then reinterpret the budget constraint as stating that the individual can choose to devote her potential income to consumption or to raising her children, and the cost of raising children is the foregone income.

10.2.2 | Investing in human capital

We take human capital to be determined by a combination of individual quality and the technological environment. Specifically, we posit a function

$$h_{t+1} = h(e_{t+1}, g_{t+1}), \quad (10.6)$$

where $g_{t+1} \equiv \frac{A_{t+1} - A_t}{A_t}$ is the rate of technological progress. The idea is that $\frac{\partial h}{\partial e} > 0$ (more education leads to more human capital), and $\frac{\partial h}{\partial g} < 0$ (faster technological progress erodes previously acquired human capital by making it obsolete). We also assume that more education increases adaptability to technological progress, so that $\frac{\partial^2 h}{\partial g \partial e} > 0$. In the absence of investment in quality, each individual has a basic-level human capital that is normalised to 1 in a stationary technological environment: $h(0, 0) = 1$.

Solution

We can substitute (10.4) and (10.6) into (10.3) to obtain:

$$u^t = (w_t [1 - n_t (\tau + e_{t+1})] h_t)^{1-\gamma} (n_t h(e_{t+1}, g_{t+1}))^\gamma. \quad (10.7)$$

Parents will choose n_t and e_{t+1} (how many children to have, and how much to invest in the education of each one of them) in order to maximise this utility. The FOC with respect to n_t will yield:

$$(1 - \gamma) c_t^{-\gamma} (n_t h(e_{t+1}, g_{t+1}))^\gamma w_t h_t (\tau + e_{t+1}) = \gamma c_t^{1-\gamma} n_t^{\gamma-1} h(e_{t+1}, g_{t+1})^\gamma \Rightarrow \quad (10.8)$$

$$\frac{c_t}{w_t h_t} = \frac{1 - \gamma}{\gamma} n_t (\tau + e_{t+1}). \quad (10.9)$$

Note, from (10.4), that the LHS of this equation is the fraction of time devoted to labour market activities, $1 - n_t (\tau + e_{t+1})$. It follows immediately that the individual will choose to devote a fraction γ of her time to child-rearing, and a fraction $1 - \gamma$ to labour.

The FOC characterises an interior solution, however, and we must take into account the subsistence consumption constraint. If $(1 - \gamma) w_t h_t < \tilde{c}$, it follows that the interior solution would not be enough to sustain the individual. In that case, the subsistence constraint is binding, and the optimal thing to do is to work as much as needed to reach \tilde{c} , and devote whatever is left to child-rearing. In other words, any individual whose potential income is below $\tilde{z} \equiv \frac{\tilde{c}}{1-\gamma}$ will be constrained to subsistence consumption, and to limited investment in their kids.

This means that, below subsistence level, increases in potential income generate no increase in consumption, and an increase in the time spent raising kids (number and quality). After the subsistence threshold, it is consumption that increases, while the time spent in child-rearing stays flat.

How about the choice of education, e_{t+1} ? With some algebra, you can show that the FOC with respect to e_{t+1} can be written as

$$\frac{\gamma}{1-\gamma} \frac{c_t}{n_t w_t h_t} = \frac{h(e_{t+1}, g_{t+1})}{\frac{\partial h}{\partial e}}. \quad (10.10)$$

Substituting the LHS using (10.9), we get the following:

$$\tau + e_{t+1} = \frac{h(e_{t+1}, g_{t+1})}{\frac{\partial h}{\partial e}}, \quad (10.11)$$

which implicitly defines e_{t+1} as a function of the productivity growth rate g_{t+1} . Using the implicit function theorem, we can see that, as long as h is decreasing in g , and concave in e , as we have assumed, e_{t+1} will be increasing in g_{t+1} . In other words, parents will choose to invest in more education when productivity grows faster, because education increases adaptability and compensates for the erosion of human capital imposed by technological change.

But how about the number of children? It is easy to see that there is no link between wages and the choice between quantity (n_t) and quality (e_{t+1}) in child-rearing: potential income doesn't show up in (10.11), and (10.9) only speaks to the overall amount of time (i.e. number times quality). However, (10.11) shows that productivity growth does affect that tradeoff: n_t doesn't show up in that equation, but e_{t+1} does. This means that an increase in g will increase the quality of kids, and reduce their quantity to keep consistency with (10.9).

In sum

This simple model therefore predicts a number of key stylised facts regarding the demographic transition:

1. An increase in (potential) income raises the number of children, but has no effect on their quality, as long as the subsistence constraint is binding.
2. An increase in (potential) income does not affect the number of children and their quality, as long as the subsistence constraint ceases to be binding.
3. An increase in the rate of technological progress reduces the number of children and increases their quality.

The first two results are driven by the subsistence requirement. As long as it is binding, any increase in potential income will imply that fewer hours of work are needed to obtain subsistence and more time can be devoted to having children. However, if the rate of technological change is constant, this will be translated into a greater number of kids, and not in higher quality: the point about investing in quality is to counteract the erosion of human capital imposed by faster technological change. When the subsistence constraint is no longer binding, then increased potential income will be reverted into increasing consumption.

10.2.3 | The dynamics of technology, education and population

In order to close the picture, we need to have a model of how technological progress takes place. Inspired by the models of endogenous growth that we have seen, we think of productivity growth being driven by the accumulation of knowledge, which is enhanced by human capital and by scale effects. Quite simply, we posit

$$g_{t+1} \equiv \frac{A_{t+1} - A_t}{A_t} = g(e_t, L_t), \quad (10.12)$$

where $g(\cdot, \cdot)$ is increasing and concave in both arguments: more and more educated people increase the rate of growth, at decreasing rates. (Remember that our discussion of scale effects argued that there is good evidence for their presence in this very long-run context, as per the models of innovation studied in Chapter 6 and synthesised in Kremer (1993).)

Obviously, the evolution of the size of the adult population is described by $L_{t+1} = n_t L_t$, since n_t is the number of children that each individual adult alive at time t chooses to have. The model of demographic decisions shows that n_t is a function of the rate of technological progress and, when the subsistence constraint is binding, also of potential income. Potential income, in turn, is a function of the existing technology, education levels, and the amount of effective resources per worker.

The one thing that is missing is the evolution of potential resources per worker, $x_t \equiv \frac{A_t X}{L_t}$, but that is easy to state: $x_{t+1} = \frac{1+g_{t+1}}{n_t} x_t$, that is, effective resources per worker grow at the rate of productivity, adjusted by the growth of population.

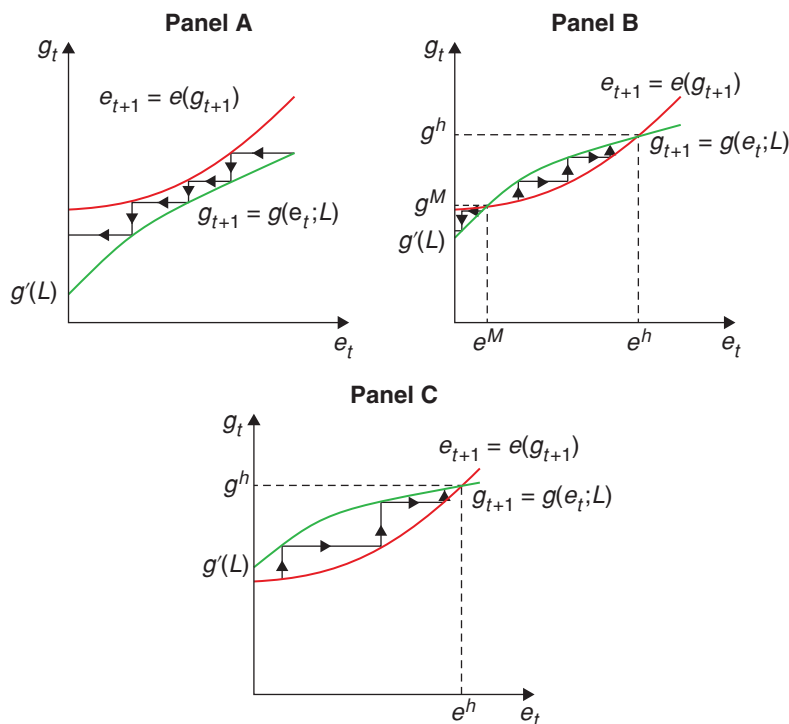
We now have a dynamic system in four variables: g_t , x_t , e_t , and L_t . These capture all of our variables of interest, namely productivity growth, productivity levels, human capital investment, and population.

Closing the model

The dynamics of the model can be analysed graphically, and we start by looking at the joint evolution of education and technology. Since the dynamics of education and technological progress are not affected by whether the subsistence constraint is binding, we can analyse them independently, as depicted in Figure 10.5. If population is small (Panel A), the rate of technological progress is slow because of scale effects. Then the only steady state is with zero levels of education: it is not worth investing in the quality of children, since the erosion caused by technological progress is really small. However, as population size grows, which the g curve shifts up, until we end up in the world of Panel B. Here there are two additional positive-education steady states (e_M and e_h , which is stable). As population grows even larger, we end up in Panel C, with a unique, stable steady state with high levels of educational investment.

10.3 | The full picture

Figure 10.5 encompasses all of our story, from the Malthusian regime to sustained growth. Consider an economy in early stages of development. Population size is relatively small and the implied slow rate of technological progress does not provide an incentive to invest in the education of children: this is the world of Figure 10.5A – the Malthusian Regime. Over time, the slow growth in population that takes place in the Malthusian Regime raises the rate of technological progress and shifts the g in

Figure 10.5 Dynamics of the model

Panel A enough to generate a qualitative change to Panel B. This is characterised by multiple, history-dependent, stable steady-state equilibria: some countries may take off and start investing in human capital, while others lag behind. However, since the economy started in the Malthusian steady state, it initially sticks in the vicinity of that steady state, which is still stable in the absence of major shocks.

Eventually, the continued increase in population leads us to Panel C: the Malthusian steady state disappears, and the economy starts converging to the steady state with high education. The resulting increase in the pace of technological progress, due to the increased levels of education, has two opposing effects on the evolution of population. On the one hand, it allows for the allocation of more resources for raising children. On the other hand, it induces a reallocation of these additional resources toward child quality. Initially, there is low demand for human capital, and the first effect dominates: this is the Post-Malthusian Regime.

The interaction between investment in human capital and technological progress generates a virtuous circle: human capital formation pushes faster technological progress, which further raises the demand for human capital, which fosters more investment in child quality. Eventually, the subsistence constraint seizes to bind, triggering a demographic transition and sustained economic growth.

10.4 | What have we learned?

In sum, this theory explains growth over the very long run by marrying the insights of endogenous growth theory (the role of scale effects and human capital in increasing the growth rate of productivity) to a theory of endogenous population growth and human capital investment. In the presence of a subsistence level of consumption, this marriage produces an initial period of stagnation, and the eventual transition to sustained economic growth with high levels of human capital investment and the demographic transition.

10.5 | What next?

For those interested in this transition Galor (2005) is a great starting point.

Notes

¹ Notice that the Kremer model does not solve this riddle. Even though it starts from a Malthusian framework, by assuming a specific process of economic growth, it only explains the divergent part of the story.

² While industrialisation is considered the turning point for economic growth, arguably their impact was possible (or at least complemented) by two other world-changing events occurring at the same time. First, the French Revolution that dismantled the segmented labour market of the middle ages, opened up labour so that anybody could find his or her best use of their own abilities, thus increasing productivity dramatically. Second, the U.S. Constitution; a key attempt to have a government controlled by its citizens and not the other way around. Reigning in the authoritarian tendencies or the possibility of capricious use of resources also provided significant productivity gains through better infrastructure and better public goods provision.

References

- Galor, O. (2005). From stagnation to growth: Unified growth theory. *Handbook of Economic Growth*, 1, 171–293.
- Kremer, M. (1993). Population growth and technological change: One million B.C. to 1990. *The Quarterly Journal of Economics*, 108(3), 681–716.