Chapter 4

A growing economy

In the previous chapter we ignored technological progress. An incontestable fact of real life in industrialized countries is, however, the presence of a persistent rise in GDP per capita – on average between 1.5 and 2.5 percent per year since 1870 in many developed economies. In regard to UK, USA, and Japan, see Fig. 4.1; and in regard to Denmark, see Fig. 4.2. In spite of the somewhat dubious quality of the data from before the Second World War, this observation should be taken into account in a model which, like the Diamond model, aims at dealing with long-run issues. For example, in relation to the question of dynamic inefficiency, cf. Chapter 3, the cut-off value of the steady-state interest rate is the steady-state GDP growth rate of the economy and this growth rate increases one-to-one with the rate of technological progress. We shall therefore now introduce technological progress.

On the basis of a summary of "stylized facts" about growth, Section 4.1 motivates the assumption that technological progress at the aggregate level takes the Harrod-neutral form. In Section 4.2 we extend the Diamond OLG model by incorporating this form of technological progress. Section 4.3 extends the concept of the golden rule to allow for the existence of technological progress. In Section 4.4 we address what is known as the marginal productivity theory of the functional income distribution and apply an expedient analytical tool, the elasticity of factor substitution. The next section defines the concept of elasticity of factor substitution at the general level. Section 4.6 then . goes into detail with the special case of a constant elasticity of factor substitution (the CES production function). Finally, Section 4.7 concludes with some general considerations regarding the concept of economic growth.

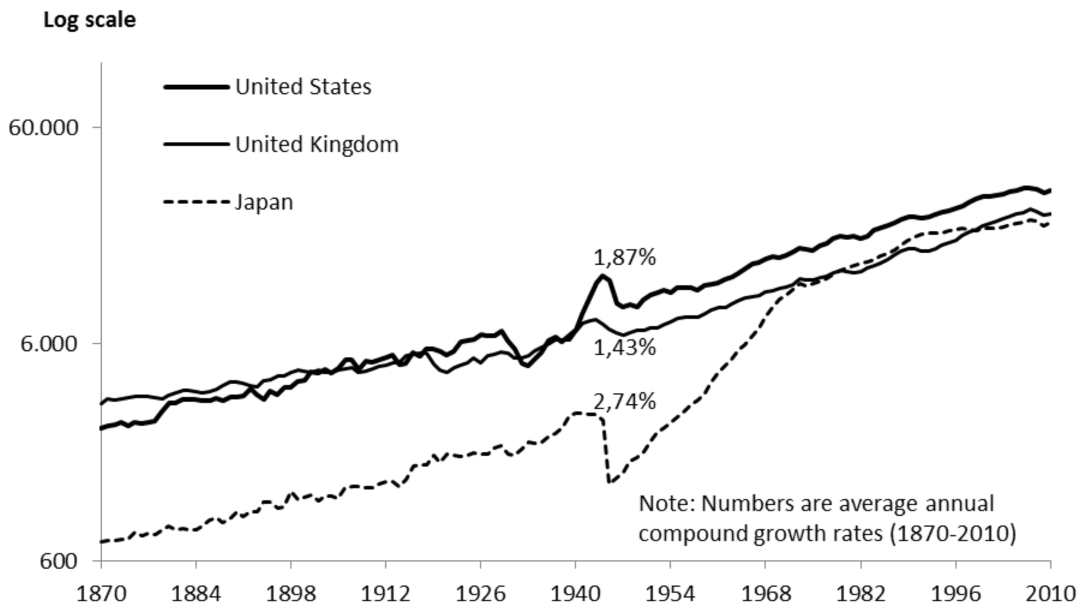


Figure 4.1: GDP per capita in USA, UK, and Japan 1870-2010. Source: Bolt and van Zanden (2013).

4.1 Harrod-neutrality and Kaldor's stylized facts

To allow for technological change, we may write aggregate production this way

$$Y_t = \tilde{F}(K_t, L_t, t), \tag{4.1}$$

where Y_t, K_t, and L_t stand for output, capital input, and labor input, respectively. Changes in technology are here represented by the dependency of the production function $\tilde{F}$ on time, t. For fixed t, the production function may still be for instance neoclassical with respect to the role of the factor inputs, the first two arguments. Often we assume that $\tilde{F}$ depends in a smooth way on time such that the partial derivative, $\partial \tilde{F}_t/\partial t$, exists and is a continuous function of (K_t, L_t, t). When $\partial \tilde{F}_t/\partial t > 0$, technological change amounts to technological *progress:* for K_t and L_t held constant, output increases with t.

A particular form of the time-dependency of the production function has attracted the attention of macroeconomists. This is known as *Harrod-neutral technological progress* and is present when we can rewrite $\tilde{F}$ such that

$$Y_t = F(K_t, T_t L_t), \tag{4.2}$$

where the "level of technology" is represented by a coefficient, T_t, on the labor input, and this coefficient is rising over time. An alternative name for this is *labor-augmenting* technological progress. The name "labor-augmenting" may sound as

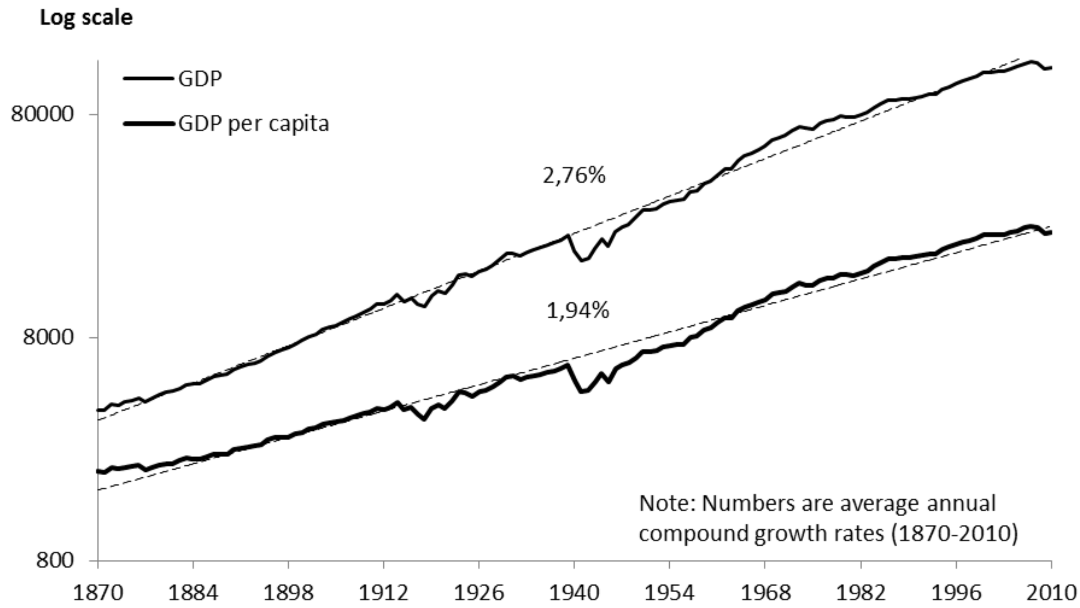


Figure 4.2: GDP and GDP per capita. Denmark 1870-2006. Sources: Bolt and van Zanden (2013); Maddison (2010); The Conference Board Total Economy Database (2013).

if *more* labor is required to reach a given output level for given capital. In fact, the opposite is the case, namely that T_t has risen so that *less* labor input is required. The idea is that the technological change − a certain percentage increase in T − affects the output level *as if* the labor input had been increased exactly by this percentage, and nothing else had happened.

The interpretation of Harrod neutrality is not that something miraculous happens to the labor input. The content of (4.2) is just that technological innovations are assumed to predominantly be such that not only do labor and capital *in combination* become more productive, but this happens to *manifest itself* in the form (4.2), that is, *as if* an improvement in the *quality* of the labor input had occurred.[1]

Kaldor's stylized facts

The reason that macroeconomists often assume that technological change at the aggregate level takes the Harrod-neutral form, as in (4.2), and not for example the

[1] As is usual in simple macroeconomic models, in both (4.1) and (4.2) it is simplifying assumed that technological progress is *disembodied.* This means that new technical and organizational knowledge increases the combined productivity of capital and workers independently of when the first were constructed and the latter educated, cf. Chapter 2.2.

form $Y_t = F(X_tK_t, T_tL_t)$ (where both X and T are changing over time, at least one of them growing), is the following. You want the long-run properties of the model to comply with Kaldor's list of "stylized facts" (Kaldor 1961) concerning the long-run evolution of certain "Great Ratios" of industrialized economies. Abstracting from short-run fluctuations, Kaldor's "stylized facts" are:

1. K/L and Y/L are growing over time and have roughly constant growth rates;

2. the output-capital ratio, Y/K, the income share of labor, wL/Y, and the economy-wide rate of return to capital, $(Y - wL - \delta K)/K$,[2] are roughly constant over time;

3. the growth rate of Y/L can vary substantially across countries for quite long time.

Ignoring the conceptual difference between the path of Y/L and that of Y *per capita* (a difference not so important in this context), the figures 4.1 and 4.2 illustrate Kaldor's "fact 1" about the long-run property of the Y/L path for the more developed countries. Japan had an extraordinarily high growth rate of GDP per capita for a couple of decades after World War II, usually explained by fast technology transfer from the most developed countries (the catching-up process which can only last until the technology gap is eliminated). Fig. 4.3 gives rough support for a part of Kaldor's "fact 2", namely the claim about long-run constancy of the labor income share of national income. "Fact 3" about large diversity across countries regarding the growth rate of Y/L over long time intervals is well documented empirically.[3]

It is fair to add, however, that the claimed regularities 1 and 2 do not fit all developed countries equally well. While Solow's famous growth model (Solow, 1956) can be seen as the first successful attempt at building a model consistent with Kaldor's "stylized facts", Solow himself once remarked about them: "There is no doubt that they are stylized, though it is possible to question whether they are facts" (Solow, 1970). Recently, several empiricists (see Literature notes) have questioned the methods which standard national income accounting applies to separate the income of entrepreneurs, sole proprietors, and unincorporated businesses into labor and capital income. It is claimed that these methods obscure a tendency of the labor income share to fall in recent decades.

[2] In this formula w is the real wage and δ is the capital depreciation rate. Land is ignored. For countries where land is a quantitatively important production factor, the denominator should be replaced by $K + p_J J$, where p_J is the real price of land, J.

[3] For a summary, see Pritchett (1997).

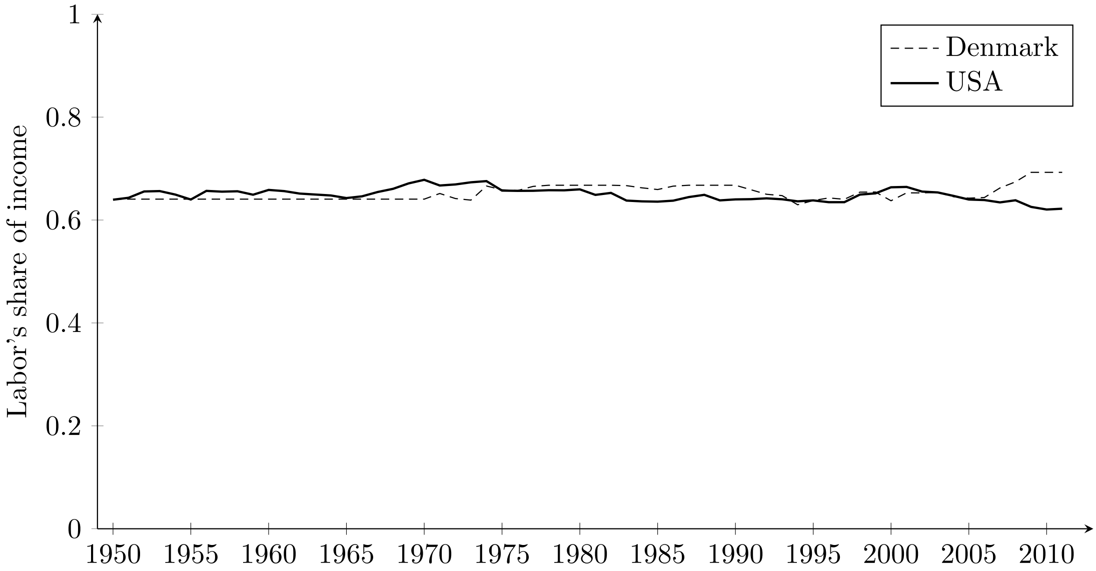


Figure 4.3: Labor's share of GDP in USA (1950-2011) and Denmark (1970-2011). Source: Feenstra, Inklaar and Timmer (2013), www.ggdc.net/pwt.

Notwithstanding these ambiguities, it is definitely a fact that many long-run models are constructed so a to comply with Kaldor's stylized facts. Let us briefly take a look at the Solow model (in discrete time) and check its consistency with Kaldor's "stylized facts". The point of departure of the Solow model and many other growth models is the *dynamic resource constraint for a closed economy*:

$$K_{t+1} - K_t = I_t - \delta K_t = S_t - \delta K_t \equiv Y_t - C_t - \delta K_t, \qquad K_0 > 0 \text{ given}, \tag{4.3}$$

where I_t is gross investment, which in a closed economy equals gross saving, $S_t \equiv Y_t - C_t$; δ is a constant capital depreciation rate, $0 \leq \delta \leq 1$.

The Solow model and Kaldor's stylized facts

As is well-known, the Solow model postulates a constant aggregate saving-income *ratio*, $\hat{s}$, so that $S_t = \hat{s}Y_t$, $0 < \hat{s} < 1$.[4] Further, the model assumes that the aggregate production function is neoclassical and features Harrod-neutral technological progress. So, let F in (4.2) be Solow's production function. To this Solow adds assumptions of CRS and exogenous geometric growth in both the technology level T and the labour force L, i.e., $T_t = T_0(1+g)^t$, $g \geq 0$, and $L_t = L_0(1+n)^t$, $n > -1$. In view of CRS, we have $Y = F(K, AL) = TLF(\tilde{k}, 1) \equiv TLf(\tilde{k})$, where $\tilde{k} \equiv K/(TL)$ is the *effective capital-labor ratio* while $f' > 0$ and $f'' < 0$.

[4] Note that $\hat{s}$ is a *ratio* while the s in the Diamond model stands for the *saving* per young.

Substituting $S_t = \hat{s}Y_t$ into $K_{t+1} - K_t = S_t - \delta K_t$, dividing through by $T_t(1+g)L_t(1+n)$ and rearranging gives the "law of motion" of the Solow economy:

$$\tilde{k}_{t+1} = \frac{\hat{s}f(\tilde{k}_t) + (1-\delta)\tilde{k}_t}{(1+g)(1+n)} \equiv \varphi(\tilde{k}_t). \tag{4.4}$$

Defining $G \equiv (1+g)(1+n)$, we have $\varphi'(\tilde{k}) = (\hat{s}f'(\tilde{k}) + 1 - \delta)/G > 0$ and $\varphi''(\tilde{k}) = \hat{s}f''(\tilde{k})/G < 0$. If $G > 1-\delta$ and f satisfies the Inada conditions $\lim_{\tilde{k}\to 0} f'(\tilde{k}) = \infty$ and $\lim_{\tilde{k}\to\infty} f'(\tilde{k}) = 0$, there is a unique and globally asymptotically stable steady state $\tilde{k}^* > 0$. The transition diagram looks entirely as in Fig. 3.4 of the previous chapter (ignoring the tildes).[5] The convergence of $\tilde{k}$ to $\tilde{k}^*$ implies that in the long run we have $K/L = \tilde{k}^*T$ and $Y/L = f(\tilde{k}^*)T$. Both K/L and Y/L are consequently growing at the same constant rate as T, the rate g. And constancy of $\tilde{k}$ implies that $Y/K = f(\tilde{k})/\tilde{k}$ is constant and so is the labor income share, $wL/Y = (f(\tilde{k}) - \tilde{k}f'(\tilde{k}))/f(\tilde{k})$, and hence also the net rate of return, $(1-wL/Y)Y/K - \delta$.

It follows that the Solow model complies with the stylized facts 1 and 2 above. Many different models aim at doing that. What these models must then have *in common* is a capability of generating *balanced growth*.

Balanced growth

With K_t, Y_t, and C_t denoting aggregate capital, output, and consumption as above, we define a balanced growth path the following way:

DEFINITION 1 A *balanced growth path*, BGP, is a path $\{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ along which the variables K_t, Y_t, and C_t are positive and grow at constant rates (not necessarily positive).

At least for a closed economy there is a general equivalence relationship between balanced growth and constancy of certain key ratios like Y/K and C/Y. This relationship is an implication of accounting based on the above aggregate dynamic resource constraint (4.3).

For an arbitrary variable $x_t \in \mathbb{R}_{++}$, we define $\Delta x_t \equiv x_t - x_{t-1}$. Whenever $x_{t-1} > 0$, the *growth rate* of x from $t-1$ to t, denoted $g_x(t)$, is defined by $g_x(t) \equiv \Delta x_t / x_{t-1}$. When there is no risk of confusion, we suppress the explicit dating and write $g_x \equiv \Delta x / x$.

PROPOSITION 1 *(the balanced growth equivalence theorem).* Let $P \equiv \{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ be a path along which K_t, Y_t, C_t, and $S_t \equiv Y_t - C_t$ are positive for all $t = 0, 1, 2, \ldots$.

[5] What makes the Solow model so easily tractable compared to the Diamond OLG model is the constant saving-income ratio which makes the transition function essentially dependent only on the production function in intensive form. Owing to dimishing marginal productivity of capital, this is a strictly concave function. Anyway, the Solow model emerges as a special case of the Diamond model, see Exercise IV.??.

Then, given the dynamic resource constraint for a closed economy, (4.3), the following holds:
(i) If P is a BGP, then $g_Y = g_K = g_C$ and the ratios Y/K and C/Y are constant.
(ii) If Y/K and C/Y are constant, then P is a BGP with $g_Y = g_K = g_C$, i.e., not only is balanced growth present but the constant growth rates of Y, K, and C are the same.

Proof Consider a path $\{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ along which K, Y, C, and $S_t \equiv Y - C_t$ are positive for all $t = 0, 1, 2, \ldots$.

(i) Suppose the path is a balanced growth path. Then, by definition, g_Y, g_K, and g_C are constant. Hence, by (4.3), $S/K = g_K + \delta$ must be constant, implying[6]

$$g_S = g_K. \tag{*}$$

By (4.3), $Y \equiv C + S$, and so

$$\begin{aligned} g_Y &= \frac{\Delta Y}{Y} = \frac{\Delta C}{Y} + \frac{\Delta S}{Y} = \frac{C}{Y}g_C + \frac{S}{Y}g_S = \frac{C}{Y}g_C + \frac{S}{Y}g_K \qquad \text{(by (*))} \\ &= \frac{C}{Y}g_C + \frac{Y-C}{Y}g_K = \frac{C}{Y}(g_C - g_K) + g_K. \qquad (**) \end{aligned}$$

Let us provisionally assume that $g_C \neq g_K$. Then (**) gives

$$\frac{C}{Y} = \frac{g_Y - g_K}{g_C - g_K}, \tag{***}$$

a constant since g_Y, g_K, and g_C are constant. Constancy of C/Y requires $g_C = g_Y$, hence, by (***), $C/Y = 1$, i.e., $C = Y$. In view of $Y \equiv C + S$, however, this implication contradicts the given condition that $S > 0$. Hence, our provisional assumption and its implication (***) are falsified. Instead we have $g_C = g_K$. By (**), this implies $g_Y = g_K = g_C$, but now without the condition $C/Y = 1$ being implied. It follows that Y/K and C/Y are constant.

(ii) Suppose Y/K and C/Y are positive constants. Applying that the ratio between two variables is constant if and only if the variables have the same (not necessarily constant or positive) growth rate, we can conclude that $g_Y = g_K = g_C$. By constancy of C/Y follows that $S/Y \equiv 1 - C/Y$ is constant. So $g_S = g_Y = g_K$, which in turn implies that S/K is constant. By (4.3),

$$\frac{S}{K} = \frac{\Delta K + \delta K}{K} = g_K + \delta,$$

[6] The ratio between two positive variables is constant if and only if the variables have the same growth rate (not necessarily constant or positive). For this and similar simple growth-arithmetic rules, see Appendix A.

so that also g_K is constant. This, together with constancy of Y/K and C/Y, implies that also g_Y and g_C are constant. □

Remark. It is part (i) of the proposition which requires the assumption $S > 0$ for all $t \geq 0$. If $S = 0$, we would have $g_K = -\delta$ and $C \equiv Y - S = Y$, hence $g_C = g_Y$ for all $t \geq 0$. Then there would be balanced growth if the common value of C and Y had a constant growth rate. This growth rate, however, could easily differ from that of K. Suppose $Y = AK^{\alpha}L^{1-\alpha}$, $0 < \alpha < 1$, $g_A = \gamma$ and $g_L = n$, where γ and n are constants. By the product and power function rule (see Appendix A), we would then have $1 + g_C = 1 + g_Y = (1+\gamma)(1-\delta)^{\alpha}(1+n)^{1-\alpha}$, which could easily be larger than 1 and thereby different from $1 + g_K = 1 - \delta \leq 1$ so that (i) no longer holds. *Example:* If $\delta = n = 0 < \gamma$, then $1 + g_Y = 1 + \gamma > 1 = 1 + g_K$.

It is part (ii) of the proposition which requires the assumption of a closed economy. In an open economy we do not necessarily have $I = S$, hence constancy of S/K no longer implies constancy of $g_K = I/K - \delta$. □

For many long-run closed-economy models, including the Diamond OLG model, it holds that if and only if the dynamic system implied by the model is in a steady state, will the economy feature balanced growth, cf. Proposition 4 below. There *exist* cases, however, where this equivalence between steady state and balanced growth does not hold (some open economy models and some models with *embodied* technological change). Hence, we shall maintain a distinction between the two concepts.

Note that Proposition 1 pertains to *any* model for which (4.3) is valid. No assumption about market form and economic agents' behavior are involved. And except for the assumed constancy of the capital depreciation rate δ, no assumption about the technology is involved, not even that constant returns to scale is present.

Proposition 1 suggests that if one accepts Kaldor's stylized facts as a rough description of more than a century's growth experience and therefore wants the model to be consistent with them, one should construct the model so that it can generate balanced growth.

Balanced growth requires Harrod-neutrality

Our next proposition states that for a model to be capable of generating balanced growth, technological progress *must* take the Harrod-neutral form (i.e., be labor-augmenting). Also this proposition holds in a fairly general setting, but not as general as that of Proposition 1. Constant returns to scale and a constant growth rate in the labor force, two aspects about which Proposition 1 is silent, will now have a role to play.[7]

[7] On the other hand we do *not* imply that CRS is *always* necessary for a balanced growth path (see Exercise 4.??).

Consider an aggregate production function

$$Y_t = \tilde{F}(K_t, BL_t, t), \qquad B > 0,\ \tilde{F}_2' \geq 0, \tilde{F}_3' > 0, \tag{4.5}$$

where B is a constant that depends on measurement units, and the function $\tilde{F}$ is homogeneous of degree one with respect to the first two arguments (CRS) and is non-decreasing in its second argument and increasing in the third, time. The latter property represents technological progress: as time proceeds, unchanged inputs of capital and labor result in more and more output. Note that $\tilde{F}$ need not be neoclassical.

Let the labor force change at a constant rate:

$$L_t = L_0(1+n)^t, \qquad n > -1, \tag{4.6}$$

where $L_0 > 0$. The Japanese economist Hirofumi Uzawa (1928-) is famous for several contributions, not least his balanced growth theorem (Uzawa 1961).

PROPOSITION 2 *(Uzawa's balanced growth theorem).* Let $P \equiv \{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ be a path along which Y_t, K_t, C_t, and $S_t \equiv Y_t - C_t$ are positive for all $t = 0, 1, 2, \ldots$, and satisfy the dynamic resource constraint for a closed economy, (4.3), given the production function (4.5) and the labor force (4.6). Then:
(i) A *necessary* condition for the path P to be a BGP is that along P it holds that

$$Y_t = \tilde{F}(K_t, T_t L_t, 0), \tag{4.7}$$

where $T_t = T_0(1+g)^t$ with $T_0 = B$ and $1 + g \equiv (1 + g_Y)/(1 + n) > 1$, g_Y being the constant growth rate of output along the BGP.
(ii) Assume $(1 + g)(1 + n) > 1 - \delta$. Then, for any $g \geq 0$ such that there is a $q > (1 + g)(1 + n) - (1 - \delta)$ with the property that the production function $\tilde{F}$ in (4.5) allows an output-capital ratio equal to q at $t = 0$ (i.e., $\tilde{F}(1, \tilde{k}^{-1}, 0) = q$ for some real number $\tilde{k} > 0$), a *sufficient* condition for $\tilde{F}$ to be compatible with a BGP with output-capital ratio equal to q is that $\tilde{F}$ can be written as in (4.7) with $T_t = B(1 + g)^t$.

Proof (i) Suppose the given path $\{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ is a BGP. By definition, g_K and g_Y are then constant so that $K_t = K_0(1 + g_K)^t$ and $Y_t = Y_0(1 + g_Y)^t$. With $t = 0$ in (4.5) we then have

$$Y_t(1 + g_Y)^{-t} = Y_0 = \tilde{F}(K_0, BL_0, 0) = \tilde{F}(K_t(1 + g_K)^{-t}, BL_t(1 + n)^{-t}, 0). \tag{4.8}$$

In view of the assumption that $S_t \equiv Y_t - C_t > 0$, we know from (i) of Proposition 1, that Y/K is constant so that $g_Y = g_K$. By CRS, (4.8) then implies

$$\tilde{F}(K_t, B(1 + g_Y)^t(1 + n)^{-t}L_t, 0) = (1 + g_Y)^t Y_0 = Y_t.$$

As $1+g \equiv (1+g_Y)/(1+n)$, this implies

$$Y_t = \tilde{F}(K_t, B(1+g)^t L_t, 0) = \tilde{F}(K_t, BL_t, t),$$

where the last equality comes from combining the first equality with (4.5). Now, the first equality shows that (4.7) holds for $T_t = B(1+g)^t = T_0(1+g)^t$. By $\tilde{F}_3'$ $(= \partial\tilde{F}/\partial t) > 0$ follows that for K_t and L_t fixed over time, Y_t is rising over time. For this to be consistent with the first equality, we must have $g > 0$.

(ii) See Appendix B. □

The form (4.7) indicates that along a BGP, technological progress must be Harrod-neutral. Moreover, by defining a new CRS production function F by $F(K_t, T_tL_t) \equiv \tilde{F}(K_t, T_tL_t, 0)$, we see that (i) of the proposition implies that at least along the BGP, we can rewrite the original production function this way:

$$Y_t = \tilde{F}(K_t, BL_t, t) = \tilde{F}(K_t, T_tL_t, 0) \equiv F(K_t, T_tL_t). \tag{4.9}$$

where F has CRS, and $T_t = T_0(1+g)^t$, with $T_0 = B$ and $1+g \equiv (1+g_Y)/(1+n)$.

What is the intuition behind the Uzawa result that for balanced growth to be possible, technological progress must at the aggregate level have the purely labor-augmenting form? We may first note that there is an asymmetry between capital and labor. Capital is an accumulated amount of non-consumed output and has thus at least a "tendency" to inherit the trend in output. In contrast, labor is a non-produced production factor. The labor force grows in an exogenous way and does *not* inherit the trend in output. Indeed, the ratio L_t/Y_t is free to adjust as t proceeds.

More specifically, consider the point of departure, the original production function (4.5). Because of CRS, it must satisfy

$$1 = \tilde{F}(\frac{K_t}{Y_t}, \frac{BL_t}{Y_t}, t). \tag{4.10}$$

We know from Proposition 1 that along a BGP, K_t/Y_t is constant. The assumption $\tilde{F}_3'$ $(= \partial\tilde{F}/\partial t) > 0$ implies that technological progress is present. Along a BGP, this progress must manifest itself in the form of a compensating change in L_t/Y_t in (4.10) as t proceeds, because otherwise the right-hand side of (4.10) would increase, which would contradict the constancy of the left-hand side. As we have in (4.5) assumed $\partial\tilde{F}/\partial L \geq 0$, the needed change in L_t/Y_t is a *fall*. The fall in L_t/Y_t must exactly offset the effect on $\tilde{F}$ of the rising t, when there is a fixed capital-output ratio and the left-hand side of (4.10) remains unchanged. It follows that along the considered BGP, L_t/Y_t is a decreasing function of t. The inverse, Y_t/L_t, is thus an *increasing* function of t. If we denote this function T_t, we end up with (4.9).

The generality of Uzawa's theorem is noteworthy. Like Proposition 1, Uzawa's theorem is about technically feasible paths, while economic institutions, market forms, and agents' behavior are not involved. The theorem presupposes CRS, but does not need that the technology has neoclassical properties not to speak of satisfying the Inada conditions. And the theorem holds for exogenous as well as endogenous technological progress.

A simple implication of the theorem is the following. Let y_t denote "labor productivity" in the sense of Y_t/L_t, k_t denote the capital-labor ratio, K_t/L_t, and c_t the consumption-labor ratio, C_t/L_t. We have:

COROLLARY Along a BGP with positive gross saving and the technology level T growing at a constant rate $g \geq 0$, output grows at the rate $(1+g)(1+n)-1$ ($\approx g+n$ for g and n "small") while labor productivity, y, capital-labor ratio, k, and consumption-labor ratio, c, all grow at the rate g.

Proof That $g_Y = (1+g)(1+n)-1$ follows from (i) of Proposition 2. As to g_y we have

$$y_t \equiv \frac{Y_t}{L_t} = \frac{Y_0(1+g_Y)^t}{L_0(1+n)^t} = y_0(1+g)^t,$$

since $1+g = (1+g_Y)/(1+n)$. This shows that y grows at the rate g. Moreover, $y/k = Y/K$, which is constant along a BGP, by (i) of Proposition 1. Hence k grows at the same rate as y. Finally, also $c/y \equiv C/Y$ is constant along a BGP, implying that also c grows at the same rate as y. □

Factor income shares

There is one facet of Kaldor's stylized facts which we have not yet related to Harrod-neutral technological progress, namely the claimed long-run "approximate" constancy of both the income share of labor and the rate of return on capital. It turns out that, if we assume (a) neoclassical technology, (b) profit maximizing firms, and (c) perfect competition in the output and factor markets, then these constancies are inherent in the combination of constant returns to scale and balanced growth.

To see this, let the aggregate production function be

$$Y_t = F(K_t, T_t L_t), \tag{4.11}$$

where F is neoclassical and has CRS. In view of perfect competition, the representative firm chooses inputs such that

$$\frac{\partial Y_t}{\partial K_t} = F_1(K_t, T_t L_t) = r_t + \delta, \qquad \text{and,} \tag{4.12}$$

$$\frac{\partial Y_t}{\partial L_t} = F_2(K_t, T_t L_t) T_t = w_t, \tag{4.13}$$

where the right-hand sides indicate the factor prices, r_t being the interest rate, δ the depreciation rate, and w_t the real wage.

In equilibrium the labor income share will be

$$\frac{w_t L_t}{Y_t} = \frac{\frac{\partial Y_t}{\partial L_t} L_t}{Y_t} = \frac{F_2(K_t, T_t L_t) T_t L_t}{Y_t}. \tag{4.14}$$

Since land as a production factor is ignored, gross capital income equals non-labor income, $Y_t - w_t L_t$. Denoting the gross capital income share by α_t, we thus have

$$\begin{aligned} \alpha_t &= \frac{Y_t - w_t L_t}{Y_t} = \frac{F(K_t, T_t L_t) - F_2(K_t, T_t L_t) T_t L_t}{Y_t} \\ &= \frac{F_1(K_t, T_t L_t) K_t}{Y_t} = \frac{\frac{\partial Y_t}{\partial K_t} K_t}{Y_t} = (r_t + \delta) \frac{K_t}{Y_t}, \end{aligned} \tag{4.15}$$

where we have used (4.13), Euler's theorem,[8] and then (4.12). Finally, when the capital good is nothing but a non-consumed output good, it has price equal to 1, and so the economy-wide rate of return on capital can be written

$$\frac{Y_t - w_t L_t - \delta K_t}{1 \cdot K_t} = \frac{Y_t - w_t L_t}{Y_t} \cdot \frac{Y_t}{K_t} - \delta = \alpha_t \cdot \frac{Y_t}{K_t} - \delta = r_t, \tag{4.16}$$

where the last equality comes from (4.15).

PROPOSITION 3 *(factor income shares under perfect competition)* Let the dynamic resource constraint for a closed economy be given as in (4.3). Assume F is neoclassical with CRS, and that the economy is competitive. Let the path $P = \{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ be BGP with positive gross saving. Then, along the path P:
(i) The gross capital income share equals some constant $\alpha \in (0, 1)$, and the labor income then equals $1 - \alpha$.
(ii) The rate of return on capital is $\alpha q - \delta$, where q is the constant output-capital ratio along the BGP.

Proof In view of CRS, $Y_t = F(K_t, T_t L_t) = T_t L_t F(\tilde{k}_t, 1) \equiv T_t L_t f(\tilde{k}_t)$, where $\tilde{k}_t \equiv K_t/(T_t L_t)$, and $f' > 0$, $f'' < 0$. From Proposition 1 follows that along the given path P, which is a BGP, Y_t/K_t is some constant, say q, equal to $f(\tilde{k}_t)/\tilde{k}_t$. Hence, $\tilde{k}_t$ is constant, say equal to $\tilde{k}^*$. Consequently, along P, $\partial Y_t/\partial K_t = f'(\tilde{k}^*) = r_t + \delta$. From this follows that r_t is a constant, r. (i) From (4.15) now follows that $\alpha_t = f'(\tilde{k}^*)/q \equiv \alpha$. Moreover, $0 < \alpha < 1$, since $0 < \alpha$ is implied by $f' > 0$, and $\alpha < 1$ is implied by the fact that $f'(\tilde{k}^*) < f(\tilde{k}^*)/\tilde{k}^* = Y/K = q$, where "<" is

[8] Indeed, from Euler's theorem follows that $F_1 K + F_2 T L = F(K, TL)$, when F is homogeneous of degree one.

due to $f'' < 0$ and $f(0) \geq 0$ (draw the graph of $f(\tilde{k})$). By the first equality in (4.15), the labor income share can be written $w_t L_t / Y_t = 1 - \alpha_t = 1 - \alpha$. (ii) Consequently, by (4.16), the rate of return on capital equals r_t $(= r) = \alpha q - \delta$. □

What this proposition amounts to is that a BGP in this economy exhibits both the first and the second "Kaldor fact" (point 1 and 2, respectively, in the list at the beginning of the chapter).

Although the proposition implies constancy of the factor income shares under balanced growth, it does not *determine* them. The proposition expresses the factor income shares in terms of the unknown constants α and q. These constants will generally depend on the effective capital-labor ratio in steady state, $\tilde{k}^*$, which will generally be an unknown as long as we have not formulated a theory of saving. This takes us back to Diamond's OLG model which provides such a theory.

4.2 The Diamond OLG model with Harrod-neutral technological progress

Recall from the previous chapter that in the Diamond OLG model people live in two periods, as young and as old. Only the young work and each young supplies one unit of labor inelastically. The period utility function, $u(c)$, satisfies the No Fast Assumption. The saving function of the young is $s_t = s(w_t, r_{t+1})$. We now include Harrod-neutral technological progress in the Diamond model.

Let (4.11) be the aggregate production function in the economy and assume, as before, that F is neoclassical with CRS. The technology level T_t grows at a constant exogenous rate:

$$T_t = T_0(1+g)^t, \qquad g \geq 0. \tag{4.17}$$

The initial level of technology, T_0, is historically given. The employment level L_t equals the number of young and thus grows at the constant exogenous rate $n > -1$.

Suppressing for a while the explicit dating of the variables, in view of CRS with respect to K and TL, we have

$$\tilde{y} \equiv \frac{Y}{TL} = F(\frac{K}{TL}, 1) = F(\tilde{k}, 1) \equiv f(\tilde{k}), \qquad f' > 0, f'' < 0,$$

where TL is *labor input in efficiency units;* $\tilde{k} \equiv K/(TL)$ is known as the *effective* or *technology-corrected capital-labor ratio* - also sometimes just called the "capital intensity". There is perfect competition in all markets. In each period the representative firm maximizes profit, $\Pi = F(K, TL) - \hat{r}K - wL$. Given the constant capital depreciation rate $\delta \in [0, 1]$, this leads to the first-order conditions

$$\frac{\partial Y}{\partial K} = \frac{\partial \left[TLf(\tilde{k})\right]}{\partial K} = f'(\tilde{k}) = r + \delta, \tag{4.18}$$

and

$$\frac{\partial Y}{\partial L} = \frac{\partial \left[TLf(\tilde{k})\right]}{\partial L} = \left[f(\tilde{k}) - f'(\tilde{k})\tilde{k}\right] T = w. \tag{4.19}$$

In view of $f'' < 0$, a $\tilde{k}$ satisfying (4.18) is unique. We let its value in period t be denoted $\tilde{k}_t^d$. Assuming equilibrium in the factor markets, this desired effective capital-labor ratio equals the effective capital-labor ratio from the supply side, $\tilde{k}_t \equiv K_t/(T_t L_t) \equiv k_t/T_t$, which is predetermined in every period. The equilibrium interest rate and real wage in period t are thus determined by

$$r_t = f'(\tilde{k}_t) - \delta \equiv r(\tilde{k}_t), \quad \text{where } r'(\tilde{k}_t) = f''(\tilde{k}_t) < 0, \tag{4.20}$$

$$w_t = \left[f(\tilde{k}_t) - f'(\tilde{k}_t)\tilde{k}\right] T_t \equiv \tilde{w}(\tilde{k}_t)T_t, \quad \text{where } \tilde{w}'(\tilde{k}_t) = -\tilde{k}_t f''(\tilde{k}_t) > 0. \tag{4.21}$$

Here, $\tilde{w}(\tilde{k}_t) = w_t/T_t$ is known as the *technology-corrected real wage.*

The equilibrium path

The aggregate capital stock at the beginning of period $t + 1$ must still be owned by the old generation in that period and thus equal the aggregate saving these people had as young in the previous period. Hence, as before, $K_{t+1} = s_t L_t = s(w_t, r_{t+1})L_t$. In view of $K_{t+1} \equiv \tilde{k}_{t+1}T_{t+1}L_{t+1} = \tilde{k}_{t+1}T_t(1+g)L_t(1+n)$, together with (4.20) and (4.21), we get

$$\tilde{k}_{t+1} = \frac{s(\tilde{w}(\tilde{k}_t)T_t, r(\tilde{k}_{t+1}))}{T_t(1 + g)(1 + n)}. \tag{4.22}$$

This is the general version of the law of motion of the Diamond OLG model with Harrod-neutral technological progress.

For the model to comply with Kaldor's "stylized facts", the model should be capable of generating balanced growth. Essentially, this capability is equivalent to being able to generate a steady state. In the presence of technological progress this latter capability requires a restriction on the lifetime utility function, U. Indeed, we see from (4.22) that the model is consistent with existence of a steady state only if the time-dependent technology level, T_t, in the numerator and denominator cancels out. This requires that the saving function is homogeneous of degree one in its first argument such that $s(\tilde{w}(\tilde{k}_t)T_t, r(\tilde{k}_{t+1})) = s(\tilde{w}(\tilde{k}_t), r(\tilde{k}_{t+1}))T_t$. In turn this is so if and only if the lifetime utility function of the young is *homothetic.* So, in addition to the No Fast Assumption from Chapter 3, we impose the Homotheticity Assumption:

$$\text{the lifetime utility function } U \text{ is homothetic.} \tag{A4}$$

This property entails that if the value of the "endowment", here the human wealth w_t, is multiplied by a $\lambda > 0$, then the chosen c_{1t} and c_{2t+1} are also multiplied by this factor λ (see Appendix C). It then follows that s_t is multiplied by λ as well. Letting $\lambda = 1/(\tilde{w}(\tilde{k}_t)T_t)$, (A4) thus allows us to write

$$s_t = s(1, r(\tilde{k}_{t+1}))\tilde{w}(\tilde{k}_t)T_t \equiv \hat{s}(r(\tilde{k}_{t+1}))\tilde{w}(\tilde{k}_t)T_t, \tag{4.23}$$

where $\hat{s}(r(\tilde{k}_{t+1}))$ is the saving-wealth *ratio* of the young. The distinctive feature is that the homothetic lifetime utility function U allows a decomposition of the young's saving into two factors, where one is the saving-wealth ratio, which depends only on the interest rate, and the other is the human wealth. By (4.22), the law of motion of the economy reduces to

$$\tilde{k}_{t+1} = \frac{\hat{s}(r(\tilde{k}_{t+1}))}{(1+g)(1+n)}\tilde{w}(\tilde{k}_t). \tag{4.24}$$

The equilibrium path of the economy can be analyzed in a similar way as in the case of no technological progress. In the assumptions (A2) and (A3) from Chapter 3 we replace k by $\tilde{k}$ and $1 + n$ by $(1 + g)(1 + n)$. As a generalization of Proposition 4 from Chapter 3, these generalized versions of (A2) and (A3), together with the No Fast Assumption (A1) and the Homotheticity Assumption (A4), guarantee that k_t over time converges to some steady state value $\tilde{k}^* > 0$.

Let an economy that can be described by the Diamond model be called a *Diamond economy.* Our conclusion is then that a Diamond economy will sooner or later settle down in a steady state. The convergence of $\tilde{k}$ implies convergence of many key variables, for instance the interest rate and the technology-corrected real wage. In view of (4.20) and (4.21), respectively, we get, for $t \to \infty$,

$$\begin{aligned} r_t &= f'(\tilde{k}_t) - \delta \to f'(\tilde{k}^*) - \delta \equiv r^*, \quad \text{and} \\ \frac{w_t}{T_t} &= f(\tilde{k}_t) - \tilde{k}_t f'(\tilde{k}_t) \to f(k^*) - k^* f'(k^*) \equiv \tilde{w}^*. \end{aligned}$$

Moreover, for instance the labor income share converges to a constant:

$$\frac{w_t L_t}{Y_t} = \frac{w_t/T_t}{Y_t/(T_t L_t)} = \frac{f(\tilde{k}_t) - \tilde{k}_t f'(\tilde{k}_t)}{f(\tilde{k}_t)} \to 1 - \frac{k^* f'(k^*)}{f(k^*)} \equiv 1 - \alpha^* \text{ for } t \to \infty.$$

The prediction from the model is thus that a Diamond economy will in the long run behave in accordance with Kaldor's stylized facts. The background for this is that convergence to a steady state is, in this and many other models, equivalent to "convergence" to a BGP. This equivalence follows from:

PROPOSITION 4 Consider a Diamond economy with Harrod-neutral technological progress at the constant rate $g \geq 0$ and positive gross saving for all t.
(i) If the economy features balanced growth, then it is in a steady state.
(ii) If the economy is in a steady state, then it features balanced growth.

Proof (i) Suppose the economy features balanced growth. Then, by Proposition 1, Y/K is constant. As $Y/K = \tilde{y}/\tilde{k} = f(\tilde{k})/\tilde{k}$, also $\tilde{k}$ is constant. Thereby the economy is in a steady state. (ii) Suppose the economy is in a steady state, i.e., for some $\tilde{k}^* > 0$, (4.24) holds for $\tilde{k}_t = \tilde{k}_{t+1} = \tilde{k}^*$. The constancy of $\tilde{k} \equiv K/(TL)$ and $\tilde{y} \equiv Y/(TL) = f(\tilde{k})$ implies that both g_K and g_Y equal the constant g_{TL} $= (1+g)(1+n) - 1 > 0$. As $S \equiv Y - C$, constancy of g_K implies constancy of $S/K = (\Delta K + \delta K)/K = g_K + \delta$, so that also S grows at the rate g_K and thereby at the same rate as output. Hence S/Y, and thereby also $C/Y \equiv 1 - S/Y$, is constant. Hence, also C grows at the constant rate g_Y. All criteria for a BGP are thus satisfied. □

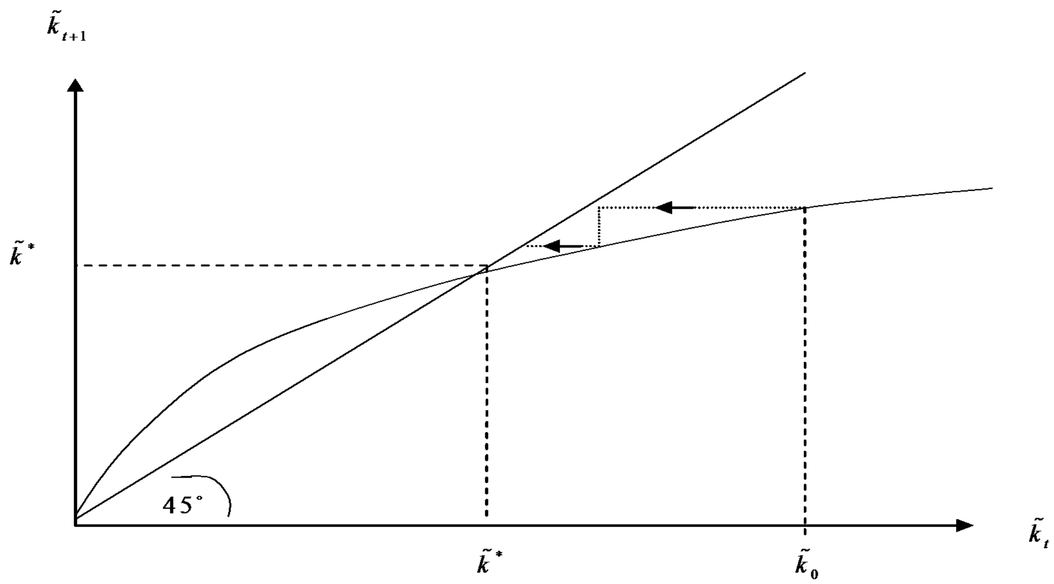


Figure 4.4: Transition curve for a well-behaved Diamond OLG model with Harrod-neutral technical progress.

Let us portray the dynamics by a transition diagram. Fig. 4.4 shows a "well-behaved" case in the sense that there is only one steady state. In the figure the initial effective capital-labor ratio, $\tilde{k}_0$, is assumed to be relatively large. This need not be interpreted as if the economy is highly developed and has a high initial capital-labor ratio, K_0/L_0. Indeed, the reason that $\tilde{k}_0 \equiv K_0/(T_0L_0)$ is large relative to its steady-steady value may be that the economy is "backward" in the sense of having a relatively low initial level of technology. Growing at a given rate g, the technology will in this situation grow faster than the capital-labor ratio, K/L, so that the effective capital-labor ratio declines over time. The process continues until the steady state is essentially reached with a real interest rate $r^* = f'(\tilde{k}^*) - \delta$. This is to remind ourselves that from an empirical point of view, the adjustment towards a steady state can be from above as well as from below.

The output growth rate in steady state, $(1+g)(1+n)-1$, is sometimes called the "natural rate of growth". Since $(1+g)(1+n)-1 = g+n+gn \approx g+n$ for g and n "small", the natural rate of growth approximately equals the sum of the rate of technological progress and the growth rate of the labor force.

Warning: When measured on an *annual* basis, the growth rates of technology and labor force, $\bar{g}$ and $\bar{n}$, do indeed tend to be "small", say $\bar{g} = 0.02$ and $\bar{n} = 0.005$, so that $\bar{g}+\bar{n}+\bar{g}\bar{n} = 0.0251 \approx 0.0250 = \bar{g}+\bar{n}$. But in the context of models like Diamond's, the period length is, say, 30 years. Then the corresponding g and n will satisfy the equations $1+g = (1+\bar{g})^{30} = 1.02^{30} = 1.8114$ and $1+n = (1+\bar{n})^{30} = 1.005^{30} = 1.1614$, respectively. We get $g+n = 0.973$, which is about 10 percent smaller than the true output growth rate over 30 years, which is $g+n+gn = 1.104$.

We end our account of Diamond's OLG model with some remarks on a popular special case of a homothetic utility function.

Example: CRRA utility

An example of a homothetic lifetime utility function is obtained by letting the period utility function take the CRRA form introduced in the previous chapter. Then

$$U(c_1, c_2) = \frac{c_1^{1-\theta}-1}{1-\theta} + (1+\rho)^{-1}\frac{c_2^{1-\theta}-1}{1-\theta}, \qquad \theta > 0. \tag{4.25}$$

Recall that the CRRA utility function with parameter θ has the property that the (absolute) elasticity of marginal utility of consumption equals the constant $\theta > 0$ for all $c > 0$. Up to a positive linear transformation it is, in fact, the only period utility function having this property. A proof that the utility function (4.25) is indeed homothetic is given in Appendix C.

One of the reasons that the CRRA function is popular in macroeconomics is that in *representative* agent models, the period utility function *must* have this form to obtain consistency with balanced growth and Kaldor's stylized facts (this is shown in Chapter 7). In contrast, a model with heterogeneous agents, like the

Diamond model, does not need CRRA utility to comply with the Kaldor facts. CRRA utility is just a convenient special case leading to homothetic lifetime utility. And *this* is what is needed for a BGP to exist and thereby for compatibility with Kaldor's stylized facts.

Given the CRRA assumption in (4.25), the saving-wealth ratio of the young becomes

$$\hat{s}(r) = \frac{1}{1 + (1+\rho)\left(\frac{1+r}{1+\rho}\right)^{(\theta-1)/\theta}}. \tag{4.26}$$

It follows that $\hat{s}'(r) \gtreqless 0$ for $\theta \lesseqgtr 1$.

When $\theta = 1$ (the case $u(c) = \ln c$), $\hat{s}(r) = 1/(2+\rho) \equiv \hat{s}$, a constant, and the law of motion (4.24) thus simplifies to

$$\tilde{k}_{t+1} = \frac{1}{(1+g)(1+n)(2+\rho)}\tilde{w}(\tilde{k}_t).$$

We see that in the $\theta = 1$ case, whatever the production function, $\tilde{k}_{t+1}$ enters only at the left-hand side of the fundamental difference equation, which thereby reduces to a simple transition function. Since $\tilde{w}'(\tilde{k}) > 0$, the transition curve is positively sloped everywhere. If the production function is Cobb-Douglas, $Y_t = K_t^{\alpha}(T_tL_t)^{1-\alpha}$, then $\tilde{w}(\tilde{k}_t) = (1-\alpha)\tilde{k}_t^{\alpha}$. Combining this with $\theta = 1$ yields a "well-behaved" Diamond model (thus having a unique and globally asymptotically stable steady state), cf. Fig. 4.4 above. In fact, as noted in Chapter 3, in combination with Cobb-Douglas technology, CRRA utility results in "well-behavedness" whatever the value of $\theta > 0$.

4.3 The golden rule under Harrod-neutral technological progress

Given that there is technological progress, consumption per unit of labor is likely to grow over time. Therefore the definition of the golden-rule capital-labor ratio from Chapter 3 has to be generalized to cover the case of growing consumption per unit of labor. To allow existence of steady states and BGPs, we maintain the assumption that technological progress is Harrod-neutral, that is, we maintain the production function (4.11) where the technology level, T, grows at a constant rate $g > 0$. We also maintain the assumption that the labor force, L_t, is fully employed and grows at a constant rate, $n \geq 0$.

Since we need not have a Diamond economy in mind, we can consider an arbitrary period length. It could be one year for instance. Consumption per unit of labor is

$$\begin{aligned} c_t &\equiv \frac{C_t}{L_t} = \frac{F(K_t, T_t L_t) - S_t}{L_t} = \frac{f(\tilde{k}_t)T_t L_t - (K_{t+1} - K_t + \delta K_t)}{L_t} \\ &= f(\tilde{k}_t)T_t - (1+g)T_t(1+n)\tilde{k}_{t+1} + (1-\delta)T_t\tilde{k}_t \\ &= \left[f(\tilde{k}_t) + (1-\delta)\tilde{k}_t - (1+g)(1+n)\tilde{k}_{t+1}\right] T_t. \end{aligned}$$

DEFINITION 2 The golden-rule capital intensity, $\tilde{k}_{GR}$, is that level of $\tilde{k} \equiv K/(TL)$ which gives the highest sustainable path for consumption per unit of labor in the economy.

To comply with the sustainability requirement, let us consider a steady state. So $\tilde{k}_{t+1} = \tilde{k}_t = \tilde{k}$ and therefore

$$c_t = \left[f(\tilde{k}) + (1-\delta)\tilde{k} - (1+g)(1+n)\tilde{k}\right] T_t \equiv \tilde{c}(\tilde{k})T_t, \tag{4.27}$$

where $\tilde{c}(\tilde{k})$ is the "technology-corrected" level of consumption per unit of labor in the steady state. We see that in steady state, consumption per unit of labor will grow at the same rate as the technology. Thus, $\ln c_t = \ln \tilde{c}(\tilde{k}) + \ln T_0 + t\ln(1+g)$. Fig. 4.5 illustrates.

Since the evolution of the technology level T_t in (4.27) is exogenous, the highest possible path of c_t is found by maximizing $\tilde{c}(\tilde{k})$. This gives the first-order condition

$$\tilde{c}'(\tilde{k}) = f'(\tilde{k}) + (1-\delta) - (1+g)(1+n) = 0. \tag{4.28}$$

When $n \geq 0$, we have $(1+g)(1+n) - (1-\delta) > 0$ in view of $g > 0$. Then, by continuity, the equation (4.28) necessarily has a unique solution in $\tilde{k} > 0$, if the production function satisfies the condition

$$\lim_{\tilde{k}\to 0} f'(\tilde{k}) > (1+g)(1+n) - (1-\delta) > \lim_{\tilde{k}\to\infty} f'(\tilde{k}),$$

which we assume. This is a milder condition than the Inada conditions. Considering the second-order condition $\tilde{c}''(\tilde{k}) = f''(\tilde{k}) < 0$, the $\tilde{k}$ satisfying (4.28) does indeed maximize $\tilde{c}(\tilde{k})$. By definition, this $\tilde{k}$ is the golden-rule capital intensity, $\tilde{k}_{GR}$. Thus

$$f'(\tilde{k}_{GR}) - \delta = (1+g)(1+n) - 1 \approx g + n, \tag{4.29}$$

where the right-hand side is the "natural rate of growth". This says that the golden-rule capital intensity is that level of the capital intensity at which the net marginal productivity of capital equals the output growth rate in steady state.

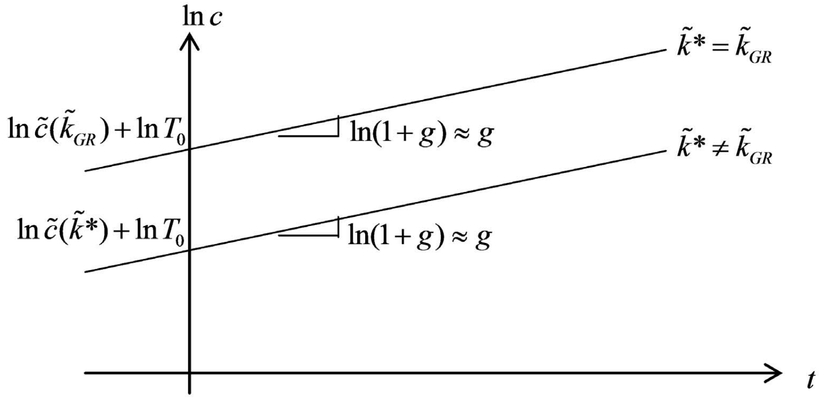


Figure 4.5: The highest sustainable path of consumption is where $\tilde{k}^* = \tilde{k}_{GR}$.

Has dynamic inefficiency been a problem in practice? As in the Diamond model without technological progress, it is theoretically possible that the economy ends up in a steady state with $\tilde{k}^* > \tilde{k}_{GR}$.[9] If this happens, the economy is dynamically inefficient and $r^* < (1+g)(1+n) - 1 \approx g + n$. To check whether dynamic inefficiency is a realistic outcome in an industrialized economy or not, we should compare the observed average GDP growth rate over a long stretch of time to the average real interest rate or rate of return in the economy. For the period after the Second World War the average GDP growth rate ($\approx g + n$) in Western countries is typically about 3 percent per year. But what interest rate should one choose? In simple macro models, like the Diamond model, there is no uncertainty and no need for money to carry out trades. In such models all assets earn the same rate of return, r, in equilibrium. In the real world there is a spectrum of interest rates, reflecting the different risk and liquidity properties of the different assets. The expected real rate of return on a short-term government bond is typically less than 3 percent per year (a relatively safe and liquid asset). This is much lower than the expected real rate of return on corporate stock, say 10 percent per year. Our model cannot tell which rate of return we should choose, but the conclusion hinges on that choice.

Abel et al. (1989) study the problem on the basis of a model with *uncertainty*. They show that a sufficient condition for dynamic efficiency is that gross investment, I, does not exceed the gross capital income in the long run, that is $I \leq Y - wL$. They find that for the U.S. and six other major OECD nations this seems to hold. Indeed, for the period 1929-85 the U.S. has, on average, $I/Y = 0.15$ and $(Y - wL)/Y = 0.29$. A similar difference is found for other industrialized countries, suggesting that they are dynamically efficient. At least in these countries,

[9] The proof is analogue to that in Chapter 3 for the case $g = 0$.

therefore, the potential coordination failure laid bare by OLG models does not seem to have been operative in practice.

4.4 The functional income distribution

By the *functional income distribution* is meant the distribution of national income on the different basic income categories: income to providers of labor, capital, and land (including other natural resources). Theory of the functional income distribution is thus theory about the determination and evolution of factor income shares.

The simplest theory about the functional income distribution is the *neoclassical theory* of the functional income distribution. It relies on competitive markets and an aggregate production function,

$$Y = F(K, L, J),$$

where K and L have the usual meaning, but the new symbol J measures the input of land. The production function is assumed to be neoclassical with CRS. When the representative firm maximizes profit and the factor markets clear, the equilibrium factor prices must satisfy:

$$\hat{r} = F_K(K, L, J), \quad w = F_L(K, L, J), \quad z = F_J(K, L, J),$$

where z denotes land rent (the charge for the use of land per unit of land), and, as usual, $\hat{r}$ is the rental rate per unit of capital, and w is the real wage per unit of labor. If in the given period, the supply of three production factors is predetermined, the three equations *determine* the three factor prices, and the factor income shares are determined as

$$\frac{\hat{r}K}{Y} = \frac{F_K(K, L, J)K}{F(K, L, J)}, \quad \frac{wL}{Y} = \frac{F_L(K, L, J)L}{F(K, L, J)}, \quad \frac{zJ}{Y} = \frac{F_J(K, L, J)J}{F(K, L, J)}.$$

The theory is also called the *marginal productivity theory* of the functional income distribution.

In advanced economies the role of land is relatively minor.[10] In fact many theoretical models completely ignore land. Below we follow that tradition, while considering the question: How is the direction of movement of the labor and capital income shares determined during the adjustment process from arbitrary initial conditions toward steady state?

First we consider the case of no technological progress.

[10]In 1750 land rent made up 20 percent of national income in England, in 1850 8 percent, and in 2010 less than 0.1 percent (Jones and Vollrath, 2013). The approximative numbers often used for the labor income share and capital income share in advanced economies are 2/3 and 1/3, respectively.

How the labor income share depends on the capital-labor ratio

Ignoring, to begin with, technological progress, we write aggregate output as $Y = F(K, L)$, where F is neoclassical with CRS. From Euler's theorem follows that $F(K, L) = F_1K + F_2L = f'(k)K + (f(k) - kf'(k))L$, where $k \equiv K/L$ and f is the production function in intensive form. In equilibrium under perfect competition we have

$$Y = \hat{r}K + wL,$$

where $\hat{r} = r + \delta = f'(k) \equiv \hat{r}(k)$ and $w = f(k) - kf'(k) \equiv w(k)$.

The *labor income share* is

$$\frac{wL}{Y} = \frac{f(k) - kf'(k)}{f(k)} \equiv \frac{w(k)}{f(k)} \equiv SL(k) = \frac{wL}{\hat{r}K + wL} = \frac{\frac{w/\hat{r}}{k}}{1 + \frac{w/\hat{r}}{k}}, \tag{4.30}$$

where the function $SL(\cdot)$ is the income share of labor function, $w/\hat{r}$ is the *factor price ratio*, and $(w/\hat{r})/k = w/(\hat{r}k)$ is the *factor income ratio*. As $\hat{r}'(k) = f''(k) < 0$ and $w'(k) = -kf''(k) > 0$, the relative factor price $w/\hat{r}$ is an increasing function of k.

Suppose that capital tends to grow faster than labor so that k rises over time. Unless the production function is Cobb-Douglas, this will under perfect competition affect the labor income share. But apriori it is not obvious in what direction. By (4.30) we see that the labor income share moves in the same direction as the factor *income* ratio, $(w/\hat{r})/k$. The latter goes up (down) depending on whether the percentage rise in the factor price ratio $w/\hat{r}$ is greater (smaller) than the percentage rise in k. So, if we let $\mathrm{E}\ell_x g(x)$ denote the elasticity of a function $g(x)$ w.r.t. x, that is, $xg'(x)/g(x))$, we can only say that

$$SL'(k) \gtreqless 0 \text{ for } \mathrm{E}\ell_k \frac{w}{\hat{r}} \gtreqless 1, \tag{4.31}$$

respectively. In words: if the production function is such that the ratio of the marginal productivities of the two production factors is strongly (weakly) sensitive to the capital-labor ratio, then the labor income share rises (falls) along with a rise in K/L.

Usually, however, the inverse elasticity is considered, namely $\mathrm{E}\ell_{w/\hat{r}}k$ $(= 1/\mathrm{E}\ell_k \frac{w}{\hat{r}})$. This elasticity indicates how sensitive the cost minimizing capital-labor ratio, k, is to a given factor price ratio $w/\hat{r}$. Under perfect competition $\mathrm{E}\ell_{w/\hat{r}}k$ coincides with what is known as the *elasticity of factor substitution* (for a general definition, see below). The latter is often denoted σ. In the CRS case, σ will be a function of only k so that we can write $\mathrm{E}\ell_{w/\hat{r}}k = \sigma(k)$. By (4.31), we therefore have

$$SL'(k) \gtreqless 0 \text{ for } \sigma(k) \lesseqgtr 1,$$

respectively.

The size of the elasticity of factor substitution is a property of the production function, hence of the technology. In special cases the elasticity of factor substitution is a constant, i.e., independent of k. For instance, if F is Cobb-Douglas, i.e., $Y = K^{\alpha}L^{1-\alpha}, 0 < \alpha < 1$, we have $\sigma(k) \equiv 1$, as we will see in Section 4.6. In this case variation in k does not change the labor income share under perfect competition. Empirically there is not agreement about the "normal" size of the elasticity of factor substitution for industrialized economies, but the bulk of studies seems to conclude with $\sigma(k) < 1$ (see below).

Adding Harrod-neutral technical progress We now add Harrod-neutral technical progress. We write aggregate output as $Y = F(K, TL)$, where F is neoclassical with CRS, and $T = T_t = T_0(1+g)^t$, $g \geq 0$. Then the labor income share is

$$\frac{wL}{Y} = \frac{w/T}{Y/(TL)} \equiv \frac{\tilde{w}}{\tilde{y}}.$$

The above formulas still hold if we replace k by $\tilde{k} \equiv K/(TL)$ and w by $\tilde{w} \equiv w/T$. We get

$$SL'(\tilde{k}) \gtreqless 0 \text{ for } \sigma(\tilde{k}) \lesseqgtr 1,$$

respectively. We see that if $\sigma(\tilde{k}) < 1$ in the relevant range for $\tilde{k}$, then market forces tend to *increase* the income share of the factor that is becoming relatively more scarce. This factor is efficiency-adjusted labor, TL, if $\tilde{k}$ is increasing, which $\tilde{k}$ will be during the transitional dynamics in a well-behaved Diamond model if $\tilde{k}_0 < \tilde{k}^*$. And if instead $\sigma(\tilde{k}) > 1$ in the relevant range for $\tilde{k}$, then market forces tend to *decrease* the income share of the factor that is becoming relatively more scarce. This factor is K, if $\tilde{k}$ is decreasing, which $\tilde{k}$ will be during the transitional dynamics in a well-behaved Diamond model if $\tilde{k}_0 > \tilde{k}^*$, cf. Fig. 4 above. Note that, given the production function in intensive form, f, the elasticity of substitution between capital and labor does not depend on whether $g = 0$ or $g > 0$, but only on the function f itself and the level of $K/(TL)$. This follows from Section 4.6.

While k empirically is clearly growing, $\tilde{k} \equiv k/T$ is not necessarily so because also T is increasing. Indeed, according to Kaldor's "stylized facts", apart from short- and medium-term fluctuations, $\tilde{k}$ − and therefore also $\hat{r}$ and the labor income share − tend to be more or less constant over time. This can happen whatever the sign of $\sigma(\tilde{k}^*) - 1$, where $\tilde{k}^*$ is the long-run value of the effective capital-labor ratio $\tilde{k}$.

As alluded to earlier, there are empiricists who reject Kaldor's "facts" as a general tendency. For instance Piketty (2014) essentially claims that in the very long run the effective capital-labor ratio $\tilde{k}$ has an upward trend, temporarily braked by two world wars and the Great Depression in the 1930s. If so, the sign of $\sigma(\tilde{k}) - 1$ becomes decisive for in what direction wL/Y will move. Piketty interprets the econometric literature as favoring $\sigma(\tilde{k}) > 1$, which means there should be downward pressure on wL/Y. This particular source behind a falling wL/Y can be questioned, however. Indeed, $\sigma(\tilde{k}) > 1$ contradicts the more general empirical view.

Immigration*

The phenomenon of migration provides another example that illustrates how the size of $\sigma(\tilde{k})$ matters. Consider a competitive economy with perfect competition, a given aggregate capital stock K, and a given technology level T (entering the production function in the labor-augmenting way as above). Suppose that due to immigration an upward shift in aggregate labor supply, L, occurs. Full employment is maintained by the needed downward real wage adjustment. Given the present model, in what direction will aggregate labor income $wL = \tilde{w}(\tilde{k})TL$ then change? The effect of the larger L is to some extent offset by a lower w brought about by the lower effective capital-labor ratio. Indeed, in view of $d\tilde{w}/d\tilde{k} = -\tilde{k}f''(\tilde{k}) > 0$, we have $\tilde{k} \downarrow$ implies $w \downarrow$ for fixed T. So we cannot apriori sign the change in wL. The following relationship can be shown (Exercise ??), however:

$$\frac{\partial(wL)}{\partial L} = (1 - \frac{\alpha(\tilde{k})}{\sigma(\tilde{k})})w \gtreqless 0 \text{ for } \sigma(\tilde{k}) \gtreqless \alpha(\tilde{k}), \tag{4.32}$$

respectively, where $a(\tilde{k}) \equiv \tilde{k}f'(\tilde{k})/f(\tilde{k})$ is the output elasticity w.r.t. capital which under perfect competition equals the gross capital income share. It follows that the larger L will not be fully offset by the lower w as long as the elasticity of factor substitution, $\sigma(\tilde{k})$, exceeds the gross capital income share, $\alpha(\tilde{k})$. This condition seems confirmed by most of the empirical evidence, see next section.

The next section describes the concept of elasticity of factor substitution at a more general level. The subsequent section introduces the special case known as the CES production function.

4.5 The elasticity of factor substitution

We shall here discuss the concept of elasticity of factor substitution at a more general level. Fig. 4.6 depicts an isoquant, $F(K, L) = \bar{Y}$, for a given neoclassical production function, $F(K, L)$, which need not have CRS. Let MRS denote the marginal rate of substitution of K for L, i.e.,

$$MRS = -\frac{dK}{dL}_{|Y=\bar{Y}} = F_L(K,L)/F_K(K,L).$$

MRS thus measures how much extra of K is needed to compensate for a reduction in L by one unit?[11] At a given point (K,L) on the isoquant curve, MRS is given by the absolute value of the slope of the tangent to the isoquant at that point. This tangent coincides with that isocost line which, given the factor prices, has minimal intercept with the vertical axis while at the same time touching the isoquant. In view of $F(\cdot)$ being neoclassical, the isoquants are by definition strictly convex to the origin. Consequently, MRS is rising along the curve when L decreases and thereby K increases. Conversely, we can let MRS be the independent variable and consider the corresponding point on the indifference curve, and thereby the ratio K/L, as a function of MRS. If we let MRS rise along the given isoquant, the corresponding value of the ratio K/L will also rise.

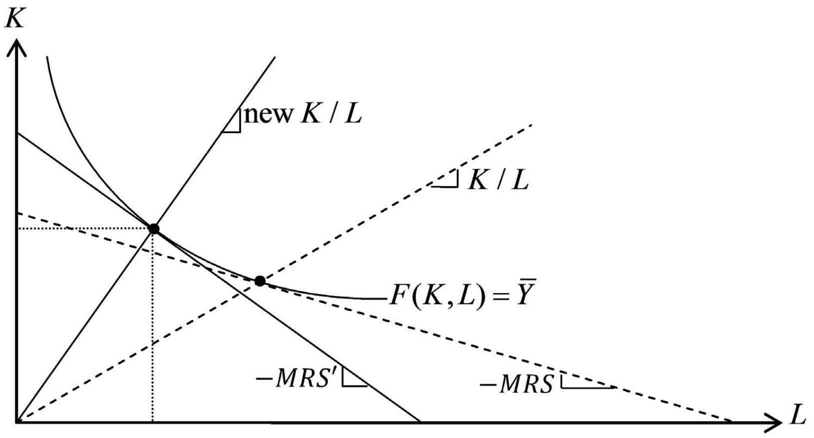


Figure 4.6: Substitution of capital for labor as the marginal rate of substitution increases from MRS to MRS'.

The *elasticity of substitution* between capital and labor, denoted $\hat{\sigma}(K,L)$, measures how sensitive K/L is vis-a-vis a rise in MRS. More precisely, $\hat{\sigma}(K,L)$ is defined as the elasticity of the ratio K/L with respect to MRS when moving along a given isoquant, evaluated at the point (K,L). Thus,

$$\hat{\sigma}(K,L) \equiv E\ell_{MRS}K/L = \frac{MRS}{K/L}\frac{d(K/L)}{dMRS}_{|Y=\bar{Y}} \approx \frac{\frac{\Delta(K/L)}{K/L}}{\frac{\Delta MRS}{MRS}}_{|Y=\bar{Y}}. \tag{4.33}$$

[11] When there is no risk of confusion as to what is up and what is down, we use MRS as a shorthand for the more precise notation MRS_{KL}.

Although the elasticity of factor substitution is a characteristic of the technology as such and is here defined without reference to markets and factor prices, it helps the intuition to refer to factor prices. At a cost-minimizing point, MRS equals the factor price ratio $w/\hat{r}$. Thus, the *elasticity of factor substitution* will under cost minimization coincide with *the percentage increase in the ratio of the cost-minimizing factor ratio induced by a one percentage increase in the inverse factor price ratio, holding the output level unchanged.*[12] The elasticity of factor substitution is thus a positive number and reflects how sensitive the capital-labor ratio K/L is under cost minimization to a one percentage increase in the factor price ratio $w/\hat{r}$ for a given output level. The less curvature the isoquant has, the greater is the elasticity of factor substitution. In an analogue way, in consumer theory one considers the elasticity of substitution between two consumption goods or between consumption today and consumption tomorrow, cf. Chapter 3. In that context the role of the given isoquant is taken over by an indifference curve. That is also the case when we consider the intertemporal elasticity of substitution in labor supply, cf. the next chapter.

Calculating the elasticity of substitution between K and L at the point (K, L), we get

$$\hat{\sigma}(K, L) = -\frac{F_K F_L (F_K K + F_L L)}{KL\left[(F_L)^2 F_{KK} - 2F_K F_L F_{KL} + (F_K)^2 F_{LL}\right]}, \tag{4.34}$$

where all the derivatives are evaluated at the point (K, L). When $F(K, L)$ has CRS, the formula (4.34) simplifies to

$$\hat{\sigma}(K, L) = \frac{F_K(K, L) F_L(K, L)}{F_{KL}(K, L) F(K, L)} = -\frac{f'(k)\left(f(k) - f'(k)k\right)}{f''(k)kf(k)} \equiv \sigma(k), \tag{4.35}$$

where $k \equiv K/L$.[13] We see that under CRS, the elasticity of substitution depends only on the capital-labor ratio k, not on the output level.

There is an alternative way of interpreting the substitution elasticity formula (4.33). This is based on the fact that any elasticity of a function $y = \varphi(x)$ can be written as a "double-log derivative": $\mathrm{E}\ell_x y \equiv (x/y)dy/dx = d\ln y/d\ln x$.[14] So, we can rewrite (4.33) as $\hat{\sigma}(K, L) = d\ln(K/L)/d\ln MRS$, which is a simple derivative when the data for K/L and MRS are given in logs.

[12] This characterization is equivalent to interpreting the elasticity of substitution as the percentage *decrease* in the factor ratio (when moving along a given isoquant) induced by a one-percentage *increase* in the *corresponding* factor price ratio.

[13] The formulas (4.34) and (4.35) are derived in Appendix D.

[14] To see this, let $X = \ln x$ and $Y = \ln y$. Then, by the chain rule,

$$\frac{d\ln y}{d\ln x} = \frac{dY}{dX} = \frac{dY}{dy}\frac{dy}{dx}\frac{dx}{dX} = \frac{1}{y}\frac{dy}{dx}e^X = \frac{x}{y}\frac{dy}{dx} = \mathrm{E}\ell_x y.$$

We will now consider the case where the elasticity of substitution is independent also of the capital-labor ratio.

4.6 The CES production function

It can be shown[15] that if a neoclassical production function with CRS has a constant elasticity of factor substitution different from one, it must be of the form

$$Y = A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{1}{\beta}}, \tag{4.36}$$

where A, α, and β are parameters satisfying $A > 0$, $0 < \alpha < 1$, and $\beta < 1$, $\beta \neq 0$. This function has been used intensively in empirical studies and is called a *CES production function* (CES for Constant Elasticity of Substitution). For a given choice of measurement units, the parameter A reflects efficiency (or what is known as *total factor productivity*) and is thus called the *efficiency parameter.* The parameters α and β are called the *distribution parameter* and the *substitution parameter,* respectively. The restriction $\beta < 1$ ensures that the isoquants are strictly convex to the origin. Note that if $\beta < 0$, the right-hand side of (4.36) is not defined when either K or L (or both) equal 0. We can circumvent this problem by extending the domain of the CES function and assign the function value 0 to these points when $\beta < 0$. Continuity is maintained in the extended domain (see Appendix E).

By taking partial derivatives in (4.36) and substituting back we get

$$\frac{\partial Y}{\partial K} = \alpha A^{\beta}\left(\frac{Y}{K}\right)^{1-\beta} \quad \text{and} \quad \frac{\partial Y}{\partial L} = (1-\alpha)A^{\beta}\left(\frac{Y}{L}\right)^{1-\beta}, \tag{4.37}$$

where $Y/K = A\left[\alpha + (1-\alpha)k^{-\beta}\right]^{\frac{1}{\beta}}$ and $Y/L = A\left[\alpha k^{\beta} + 1 - \alpha\right]^{\frac{1}{\beta}}$.[16] The marginal rate of substitution of K for L therefore is

$$MRS = \frac{\partial Y/\partial L}{\partial Y/\partial K} = \frac{1-\alpha}{\alpha}k^{1-\beta} > 0.$$

Consequently,

$$\frac{dMRS}{dk} = \frac{1-\alpha}{\alpha}(1-\beta)k^{-\beta},$$

where the inverse of the right-hand side is the value of $dk/dMRS$. Substituting these expressions into (4.33) gives

$$\hat{\sigma}(K,L)\ (= \sigma(k)) = \frac{1}{1-\beta} \equiv \sigma, \tag{4.38}$$

[15] See, e.g., Arrow et al. (1961).

[16] The calculations are slightly simplified if we start from the transformation $Y^{\beta} = A^{\beta}\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]$.

confirming the constancy of the elasticity of substitution. Since $\beta < 1$, $\sigma > 0$ always. A higher substitution parameter, β, results in a higher elasticity of factor substitution, σ. And $\sigma \lesseqgtr 1$ for $\beta \lesseqgtr 0$, respectively.

Since $\beta = 0$ is not allowed in (4.36), at first sight we cannot get $\sigma = 1$ from this formula. Yet, $\sigma = 1$ can be introduced as the *limiting* case of (4.36) when $\beta \rightarrow 0$, which turns out to be the Cobb-Douglas function. Indeed, one can show[17] that, for fixed K and L,

$$A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{1}{\beta}} \rightarrow AK^{\alpha}L^{1-\alpha}, \text{ for } \beta \rightarrow 0 \text{ (so that } \sigma \rightarrow 1).$$

By a similar procedure as above we find that a Cobb-Douglas function always has elasticity of substitution equal to 1; this is exactly the value taken by σ in (4.38) when $\beta = 0$. In addition, the Cobb-Douglas function is the *only* production function that has unit elasticity of substitution whatever the capital-labor ratio.

Another interesting limiting case of the CES function appears when, for fixed K and L, we let $\beta \rightarrow -\infty$ so that $\sigma \rightarrow 0$. We get

$$A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{1}{\beta}} \rightarrow A\min(K, L), \text{ for } \beta \rightarrow -\infty \text{ (so that } \sigma \rightarrow 0). \quad (4.39)$$

So in this case the CES function approaches a Leontief production function, the isoquants of which form a right angle, cf. Fig. 4.7. In the limit there is *no* possibility of substitution between capital and labor. In accordance with this the elasticity of substitution calculated from (4.38) approaches zero when β goes to $-\infty$.

Finally, let us consider the "opposite" transition. For fixed K and L we let the substitution parameter rise towards 1 and get

$$A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{1}{\beta}} \rightarrow A\left[\alpha K + (1-\alpha)L\right], \text{ for } \beta \rightarrow 1 \text{ (so that } \sigma \rightarrow \infty).$$

Here the elasticity of substitution calculated from (4.38) tends to ∞ and the isoquants tend to straight lines with slope $-(1-\alpha)/\alpha$. In the limit, the production function thus becomes linear and capital and labor become *perfect substitutes.*

Fig. 4.7 depicts isoquants for alternative CES production functions and their limiting cases. In the Cobb-Douglas case, $\sigma = 1$, the horizontal and vertical asymptotes of the isoquant coincide with the coordinate axes. When $\sigma < 1$, the horizontal and vertical asymptotes of the isoquant belong to the interior of the positive quadrant. This implies that both capital and labor are essential inputs. When $\sigma > 1$, the isoquant terminates in points *on* the coordinate axes. Then neither capital, nor labor are essential inputs. Empirically there is not complete agreement about the "normal" size of the elasticity of factor substitution for

[17] Proofs of this and the further claims below are in Appendix E.

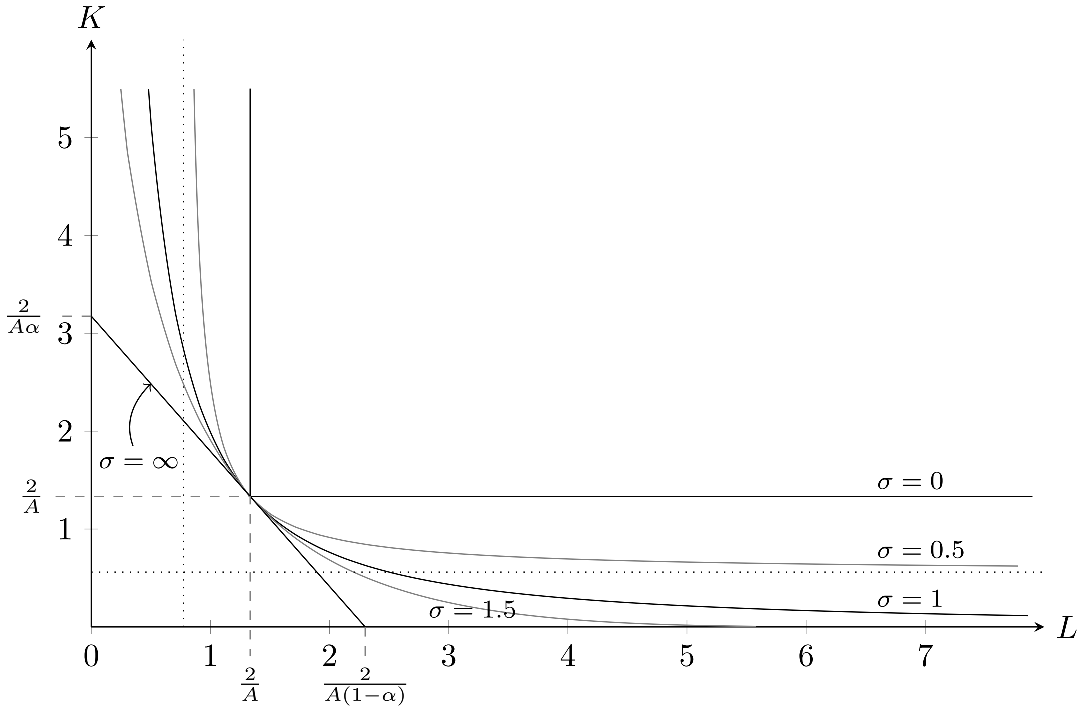


Figure 4.7: Isoquants for the CES function for alternative values of σ ($A = 1.5$, $\bar{Y} = 2$, and $\alpha = 0.42$).

industrialized economies. The elasticity also differs across the production sectors. A thorough econometric study (Antràs, 2004) of U.S. data indicate the aggregate elasticity of substitution to be in the interval $(0.5, 1.0)$. The survey by Chirinko (2008) concludes with the interval $(0.4, 0.6)$. Starting from micro data, a recent study by Oberfield and Raval (2014) finds that the elasticity of factor substitution for the US manufacturing sector as a whole has been stable since 1970 at about 0.7.

The CES production function in intensive form

Dividing through by L on both sides of (4.36), we obtain the CES production function in intensive form,

$$y \equiv \frac{Y}{L} = A(\alpha k^{\beta} + 1 - \alpha)^{\frac{1}{\beta}}, \tag{4.40}$$

where $k \equiv K/L$. The marginal productivity of capital can be written

$$MPK = \frac{dy}{dk} = \alpha A \left[\alpha + (1-\alpha)k^{-\beta}\right]^{\frac{1-\beta}{\beta}} = \alpha A^{\beta} \left(\frac{y}{k}\right)^{1-\beta},$$

which of course equals $\partial Y/\partial K$ in (4.37). We see that the CES function violates either the lower or the upper Inada condition for MPK, depending on the sign of β. Indeed, when $\beta < 0$ (i.e., $\sigma < 1$), then for $k \to 0$ both y/k and dy/dk approach an upper bound equal to $A\alpha^{1/\beta} < \infty$, thus violating the *lower* Inada condition for MPK (see the left-hand panel of Fig. 4.8). It is also noteworthy that in this case, for $k \to \infty$, y approaches an upper bound equal to $A(1-\alpha)^{1/\beta} < \infty$. These features reflect the low degree of substitutability when $\beta < 0$.

When instead $\beta > 0$, there is a high degree of substitutability ($\sigma > 1$). Then, for $k \to \infty$ both y/k and $dy/dk \to A\alpha^{1/\beta} > 0$, thus violating the *upper* Inada condition for MPK (see right-hand panel of Fig. 4.8). It is also noteworthy that for $k \to 0$, y approaches a positive lower bound equal to $A(1-\alpha)^{1/\beta} > 0$. Thus, when $\sigma > 1$, capital is not essential. At the same time $dy/dk \to \infty$ for $k \to 0$ (so the lower Inada condition for the marginal productivity of capital holds). Details are in Appendix E.

The marginal productivity of labor is

$$MPL = \frac{\partial Y}{\partial L} = (1-\alpha)A^{\beta}y^{1-\beta} = (1-\alpha)A(\alpha k^{\beta} + 1 - \alpha)^{(1-\beta)/\beta} \equiv w(k),$$

from (4.37). Under perfect competition, the equilibrium labor income share is thus

$$\frac{wL}{Y} = \frac{(1-\alpha)(\alpha k^{\beta} + 1 - \alpha)^{1/\beta - 1}}{(\alpha k^{\beta} + 1 - \alpha)^{\frac{1}{\beta}}} = \frac{1-\alpha}{\alpha k^{\beta} + 1 - \alpha}.$$

Since (4.36) is symmetric in K and L, we get a series of symmetric results by considering output per unit of capital as $x \equiv Y/K = A\left[\alpha + (1-\alpha)(L/K)^{\beta}\right]^{1/\beta}$. In total, therefore, when there is low substitutability ($\sigma < 1$), for fixed input of either of the production factors, there is an upper bound for how much an unlimited input of the other production factor can increase output. And when there is high substitutability ($\sigma > 1$), there is no such bound and an unlimited input of either production factor take output to infinity.

The Cobb-Douglas case, i.e., the limiting case for $\beta \to 0$, constitutes in several respects an intermediate case in that *all* four Inada conditions are satisfied and we have $y \to 0$ for $k \to 0$, and $y \to \infty$ for $k \to \infty$.

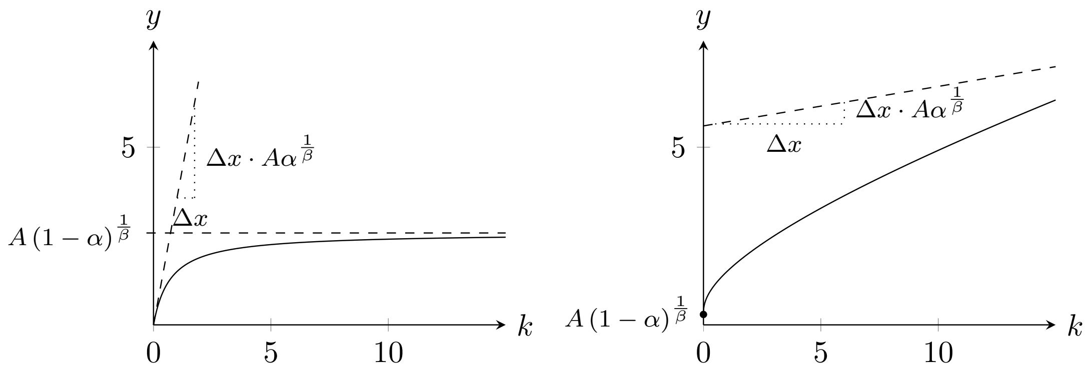
Returning to the general CES function, in case of Harrod-neutral technological progress, (4.36) and (4.40) are replaced by

$$Y = A\left[\alpha K^{\beta} + (1-\alpha)(TL)^{\beta}\right]^{\frac{1}{\beta}}$$

and

$$\tilde{y} \equiv \frac{Y}{TL} = A(\alpha \tilde{k}^{\beta} + 1 - \alpha)^{\frac{1}{\beta}},$$

respectively, where T is the technology level, and $\tilde{k} \equiv K/(TL)$.



a) The case of $\sigma < 1$. a) The case of $\sigma > 1$.

Figure 4.8: The CES production function in intensive form, $\sigma = 1/(1-\beta)$, $\beta < 1$.

Generalizations*

The CES production functions considered above have CRS. By adding an elasticity of scale parameter, γ, we get the generalized form (the case without technological progress):

$$Y = A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{\gamma}{\beta}}, \qquad \gamma > 0. \tag{4.41}$$

In this form the CES function is homogeneous of degree γ. For $0 < \gamma < 1$, there are DRS, for $\gamma = 1$ CRS, and for $\gamma > 1$ IRS. If $\gamma \neq 1$, it may be convenient to consider $Q \equiv Y^{1/\gamma} = A^{1/\gamma}\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{1/\beta}$ and $q \equiv Q/L = A^{1/\gamma}(\alpha k^{\beta} + 1 - \alpha)^{1/\beta}$.

The elasticity of substitution between K and L is $\sigma = 1/(1-\beta)$ whatever the value of γ. So including the limiting cases as well as non-constant returns to scale in the "family" of production functions with constant elasticity of substitution, we have the simple classification displayed in Table 4.1.

Table 4.1 The family of production functions with constant elasticity of substitution.

$\sigma = 0$	$0 < \sigma < 1$	$\sigma = 1$	$\sigma > 1$
Leontief	CES	Cobb-Douglas	CES

Note that only for $\gamma \leq 1$ is (4.41) a *neoclassical* production function. This is because, when $\gamma > 1$, the conditions $F_{KK} < 0$ and $F_{NN} < 0$ do not hold everywhere.

We may generalize further by assuming there are n inputs, in the amounts $X_1, X_2, ..., X_n$. Then the CES production function takes the form

$$Y = A\left[\alpha_1 X_1{}^{\beta} + \alpha_2 X_2{}^{\beta} + ...\alpha_n X_n{}^{\beta}\right]^{\frac{\gamma}{\beta}}, \ \alpha_i > 0 \text{ for all } i, \sum_i \alpha_i = 1, \gamma > 0. \quad (4.42)$$

In analogy with (4.33), for an n-factor production function the *partial elasticity of substitution* between factor i and factor j is defined as

$$\sigma_{ij} = \frac{MRS_{ij}}{X_i/X_j} \frac{d(X_i/X_j)}{dMRS_{ij}}_{|Y=\bar{Y}},$$

where it is understood that not only the output level but also all X_k, $k \neq i$, j, are kept constant. Note that $\sigma_{ji} = \sigma_{ij}$. In the CES case considered in (4.42), all the partial elasticities of substitution take the same value, $1/(1-\beta)$.

4.7 Concluding remarks

(incomplete)

When speaking of "sustained growth" in variables like K, Y, and C, we do not mean growth in a narrow physical sense. Given limited natural resources (matter and energy), sustained exponential growth in a physical sense is not possible. But sustained exponential growth in terms of economic value is not ruled out. We should for instance understand K broadly as "produced means of production" of *rising quality* and *falling material intensity* (think of the rising efficiency of microprocessors). Similarly, C must be seen as a composite of consumer goods and services with declining material intensity over time. This accords with the empirical fact that as income rises, the share of consumption expenditures devoted to agricultural and industrial products declines and the share devoted to services, hobbies, and amusement increases. Although "economic development" is perhaps a more appropriate term (suggesting qualitative and structural change), we will in this book retain standard terminology and speak of "economic growth".

A further remark about terminology. In the branch of economics called economic growth theory, the term "economic growth" is used primarily for growth of *productivity* and *income per capita* rather than just growth of GDP.

4.8 Literature notes

1. We introduced the assumption that at the macroeconomic level the "direction" of technological progress tends to be Harrod-neutral. Otherwise the model

will not be consistent with Kaldor's stylized facts. The Harrod-neutrality of the "direction" of technological progress is in the present model just an exogenous feature. This raises the question whether there are *mechanisms* tending to generate Harrod-neutrality. Fortunately new growth theory provides clues as to the sources of the speed as well as the direction of technological change. A facet of this theory is that the direction of technological change is linked to the same economic forces as the speed, namely profit incentives. See Acemoglu (2003) and Jones (2006).

2. Recent literature discussing Kaldor's "stylized facts" includes Rognlie (2015), Gollin (2002), Elsby et al. (2013), and Karabarbounis and Neiman (2014). The latter three references conclude with serious scepticism. Attfield and Temple (2010) and others, however, find support for the Kaldor "facts" considering the US and UK based on time-series econometrics. This means an observed evolution roughly obeying balanced growth in terms of *aggregate* variables. *Structural change* is not ruled out by this. *A* changing sectorial composition of the economy is under certain conditions compatible with balanced growth (in a generalized sense) at the aggregate level, cf. the "Kuznets facts" (see Kongsamut et al., 2001, and Acemoglu, 2009).

3. In Section 4.2 we claimed that from an empirical point of view, the adjustment towards a steady state can be from above as well as from below. Indeed, Cho and Graham (1996) find that "on average, countries with a lower income per adult are above their steady-state positions, while countries with a higher income are below their steady-state positions".

4. As to the assessment of whether dynamic inefficiency is - or at least has been - part of reality, in addition to Abel et al. (1989) other useful sources include Ball et al. (1998), Blanchard and Weil (2001), and Barbie, Hagedorn, and Kaul (2004). A survey is given in Weil (2008).

5. In the Diamond OLG model as well as in many other models, a steady state and a balanced growth path imply each other. Indeed, they are two sides of the same process. There *exist* cases, however, where this equivalence does not hold (some open economy models and some models with *embodied* technological change, see Groth et al., 2010). Therefore, it is recommendable always to maintain a terminological distinction between the two concepts.

6. On the declining material intensity of consumer goods and services as technology develops, see Fagnart and Germain (2011).

From here incomplete:

The term "Great Ratios" of the economy was coined by Klein and Kosubud (1961).

La Grandville (1989): normalization of the CES function. La Grandville (2009) contains a lot about analytical aspects linked to the CES production function and the concept of elasticity of factor substitution.

Piketty (2014), Zucman ( ).

According to Summers (2014), Piketty's interpretation of data relevant for estimation of the elasticity of factor substitution relies on conflating gross and net returns to capital. Krusell and Smith (2015) and Ronglie (2015).

Demange and Laroque (1999, 2000) extend Diamond's OLG model to uncertain environments.

For expositions in depth of OLG modeling and dynamics in discrete time, see Azariadis (1993), de la Croix and Michel (2002), and Bewley (2007).

Dynamic inefficiency, see also Burmeister (1980).

Uzawa's theorem: Uzawa (1961), Schlicht (2006).

The way the intuition behind the Uzawa theorem was presented in Section 4.1 draws upon Jones and Scrimgeour (2008).

For more general and flexible production functions applied in econometric work, see, e.g., Nadiri (1982).

Other aspects of life cycle behavior: education. OLG where people live three periods. Also Eggertsson and Mehrotra (2015).

4.9 Appendix

A. Growth and interest arithmetic in discrete time

Let $t = 0, \pm 1, \pm 2, \dots$, and consider the variables z_t, x_t, and y_t, assumed positive for all t. Define $\Delta z_t = z_t - z_{t-1}$ and Δx_t and Δy_t similarly. These Δ's need not be positive. The *growth rate* of x_t from period $t-1$ to period t is defined as the relative rate of increase in x, i.e., $\Delta x_t / x_{t-1} \equiv x_t / x_{t-1}$. And the *growth factor* for x_t from period $t-1$ to period t is defined as $1 + x_t / x_{t-1}$.

As we are here interested not in the time evolution of growth rates, we simplify notation by suppressing the t's. So we write the growth rate of x as $g_x \equiv \Delta x / x_{-1}$ and similarly for y and z.

PRODUCT RULE If $z = xy$, then $1 + g_z = (1 + g_x)(1 + g_y)$ and $g_z \approx g_x + g_y$, when g_x and g_y are "small".

Proof. By definition, $z = xy$, which implies $z_{-1} + \Delta z = (x_{-1} + \Delta x)(y_{-1} + \Delta y)$. Dividing by $z_{-1} = x_{-1} y_{-1}$ gives $1 + \Delta z / z_{-1} = (1 + \Delta x / x_{-1})(1 + \Delta y / y_{-1})$ as claimed. By carrying out the multiplication on the right-hand side of this equation, we get $1 + \Delta z / z_{-1} = 1 + \Delta x / x_{-1} + \Delta y / y_{-1} + (\Delta x / x_{-1})(\Delta y / y_{-1}) \approx 1 + \Delta x / x_{-1} + \Delta y / y_1$ when $\Delta x / x_{-1}$ and $\Delta y / y_{-1}$ are "small". Subtracting 1 on both sides gives the stated approximation. □

So the product of two positive variables will grow at a rate approximately equal to the sum of the growth rates of the two variables.

QUOTIENT RULE If $z = \frac{x}{y}$, then $1 + g_z = \frac{1+g_x}{1+g_y}$ and $g_z \approx g_x - g_y$, when g_x and g_y are "small".

Proof. By interchanging z and x in Product Rule and rearranging, we get $1 + \Delta z/z_{-1} = \frac{1+\Delta x/x_{-1}}{1+\Delta y/y_{-1}}$, as stated in the first part of the claim. Subtracting 1 on both sides gives $\Delta z/z_{-1} = \frac{\Delta x/x_{-1} - \Delta y/y_{-1}}{1+\Delta y/y_{-1}} \approx \Delta x/x_{-1} - \Delta y/y_{-1}$, when $\Delta x/x_{-1}$ and $\Delta y/y_{-1}$ are "small". This proves the stated approximation. □

So the ratio between two positive variables will grow at a rate approximately equal to the excess of the growth rate of the numerator over that of the denominator. An implication of the first part of Claim 2 is: the ratio between two positive variables is constant if and only if the variables have the same growth rate (not necessarily constant or positive).

POWER FUNCTION RULE If $z = x^\alpha$, then $1 + g_z = (1 + g_x)^\alpha$.

Proof. $1 + g_z \equiv z/z_{-1} = (x/x_{-1})^\alpha \equiv (1 + g_x)^\alpha$. □

Given a time series $x_0, x_1, ..., x_n$, by the *average growth rate* per period, or more precisely, the *average compound growth rate,* is meant a g which satisfies $x_n = x_0(1 + g)^n$. The solution for g is $g = (x_n/x_0)^{1/n} - 1$.

Finally, the following approximation may be useful (for intuition) if used with caution:

THE GROWTH FACTOR With n denoting a positive integer above 1 but "not too large", the growth factor $(1 + g)^n$ can be approximated by $1 + ng$ when g is "small". For $g \neq 0$, the approximation error is larger the larger is n.

Proof. (i) We prove the claim by induction. Suppose the claim holds for a fixed $n \geq 2$, i.e., $(1 + g)^n \approx 1 + ng$ for g "small". Then $(1 + g)^{n+1} = (1 + g)^n(1 + g)$ $\approx (1 + ng)(1 + g) = 1 + ng + g + ng^2 \approx 1 + (n + 1)g$ since g "small" implies g^2 "very small" and therefore ng^2 "small" if n is not "too" large. So the claim holds also for $n + 1$. Since $(1 + g)^2 = 1 + 2g + g^2 \approx 1 + 2g$, for g "small", the claim does indeed hold for $n = 2$. □

THE EFFECTIVE ANNUAL RATE OF INTEREST Suppose interest on a loan is charged n times a year at the rate r/n per year. Then the *effective annual interest rate* is $(1 + r/n)^n - 1$.

B. Proof of the sufficiency part of Uzawa's theorem

For convenience we restate the full theorem here:

PROPOSITION 2 *(Uzawa's balanced growth theorem).* Let $P \equiv \{(K_t, Y_t, C_t)\}_{t=0}^{\infty}$ be a path along which Y_t, K_t, C_t, and $S_t \equiv Y_t - C_t$ are positive for all $t = 0, 1, 2, \ldots$, and satisfy the dynamic resource constraint for a closed economy, (4.3), given the production function (4.5) and the labor force (4.6). Then:
(i) A *necessary* condition for the path P to be a BGP is that along P it holds that

$$Y_t = \tilde{F}(K_t, T_t L_t, 0), \qquad (*)$$

where $T_t = T_0(1+g)^t$ with $T_0 = B$ and $1 + g \equiv (1 + g_Y)/(1+n) > 1$, g_Y being the constant growth rate of output along the BGP.
(ii) Assume $(1+g)(1+n) > 1-\delta$. Then, for any $g \geq 0$ such that there is a $q > (1+g)(1+n) - (1-\delta)$ with the property that the production function $\tilde{F}$ in (4.5) allows an output-capital ratio equal to q at $t = 0$ (i.e., $\tilde{F}(1, \tilde{k}^{-1}, 0) = q$ for some real number $\tilde{k} > 0$), a *sufficient* condition for $\tilde{F}$ to be compatible with a BGP with output-capital ratio equal to q is that $\tilde{F}$ can be written as in (4.7) with $T_t = B(1+g)^t$.

Proof (i) See Section 4.1. (ii) Suppose (*) holds with $T_t = B(1+g)^t$. Let $g \geq 0$ be given such that there is a $q > (1+g)(1+n) - (1-\delta) > 0$ with the property that

$$\tilde{F}(1, \tilde{k}^{-1}, 0) = q \qquad (**)$$

for some constant $\tilde{k} > 0$. Our strategy is to prove the claim by construction of a path $P = (Y_t, K_t, C_t)_{t=0}^{\infty}$ which satisfies it. We let P be such that the saving-income ratio is a constant $\hat{s} \equiv [(1+g)(1+n) - (1-\delta)]/q \in (0,1)$, i.e., $Y_t - C_t \equiv S_t = \hat{s}Y_t$ for all $t = 0, 1, 2, \ldots$. Inserting this, together with $Y_t = f(\tilde{k}_t)T_t L_t$, where $f(\tilde{k}_t) \equiv \tilde{F}(\tilde{k}_t, 1, 0)$ and $\tilde{k}_t \equiv K_t/(T_t L_t)$, into (4.3), rearranging gives the Solow equation (4.4), which we may rewrite as

$$\tilde{k}_{t+1} - \tilde{k}_t = \frac{\hat{s} f(\tilde{k}_t) - [(1+g)(1+n) - (1-\delta)]\,\tilde{k}_t}{(1+g)(1+n)}.$$

We see that $\tilde{k}_t$ is constant if and only if $\tilde{k}_t$ satisfies the equation $f(\tilde{k}_t)/\tilde{k}_t = [(1+g)(1+n) - (1-\delta)]/\hat{s} \equiv q$. By (**) and the definition of f, the required value of $\tilde{k}_t$ is $\tilde{k}$, which is thus the steady state for the constructed Solow model. Letting K_0 satisfy $K_0 = \tilde{k}BL_0$, where $B = T_0$, we thus have $\tilde{k}_0 = K_0/(T_0 L_0) = \tilde{k}$. So that the initial value of $\tilde{k}_t$ equals the steady-state value. It follows that $\tilde{k}_t = \tilde{k}$ for all $t = 0, 1, 2, \ldots$, and so $Y_t/K_t = f(\tilde{k}_t)/\tilde{k}_t = f(\tilde{k})/\tilde{k} = q$ for all $t \geq 0$. In addition, $C_t = (1-\hat{s})Y_t$, so that C_t/Y_t is constant along the path P. As both Y/K and C/Y are thus constant along the path P, by (ii) of Proposition 1 follows that P is a BGP. □

It is noteworthy that the proof of the sufficiency part of the theorem is *constructive.* It provides a method for constructing a BGP with a given technology growth rate and a given output-capital ratio.

C. Homothetic utility functions

Generalities A set C in $\mathbb{R}^n$ is called a *cone* if $x \in C$ and $\lambda > 0$ implies $\lambda x \in C$. A function $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ is *homothetic* in the cone C if for all $\mathbf{x}, \mathbf{y} \in C$ and all $\lambda > 0$, $f(\mathbf{x}) = f(\mathbf{y})$ implies $f(\lambda \mathbf{x}) = f(\lambda \mathbf{y})$.

Consider the continuous utility function $U(x_1, x_2)$, defined in $\mathbb{R}^2_+$. Suppose U is *homothetic* and that the consumption bundles (x_1, x_2) and (y_1, y_2) are on the same indifference curve, i.e., $U(x_1, x_2) = U(y_1, y_2)$. Then for any $\lambda > 0$ we have $U(\lambda x_1, \lambda x_2) = U(\lambda y_1, \lambda y_2)$ so that the bundles $(\lambda x_1, \lambda x_2)$ and $(\lambda y_1, \lambda y_2)$ are also on the same indifference curve.

For a continuous utility function $U(x_1, x_2)$, defined in $\mathbb{R}^2_+$ and increasing in each of its arguments (as is our life time utility function in the Diamond model), one can show that U is homothetic if and only if U can be written $U(x_1, x_2) \equiv F(f(x_1, x_2))$ where the function f is homogeneous of degree one and F is an increasing function. The "if" part is easily shown. Indeed, if $U(x_1, x_2) = U(y_1, y_2)$, then $F(f(x_1, x_2)) = F(f(y_1, y_2))$. Since F is increasing, this implies $f(x_1, x_2) = f(y_1, y_2)$. Because f is homogeneous of degree one, if $\lambda > 0$, then $f(\lambda x_1, \lambda x_2) = \lambda f(x_1, x_2)$ and $f(\lambda y_1, \lambda y_2) = \lambda f(y_1, y_2)$ so that $U(\lambda x_1, \lambda x_2) = F(f(\lambda x_1, \lambda x_2)) = F(f(\lambda y_1, \lambda y_2)) = U(\lambda y_1, \lambda y_2)$, which shows that U is homothetic. As to the "only if" part, see Sydsaeter et al. (2002).

Using differentiability of our homothetic time utility function $U(x_1, x_2) \equiv F(f(x_1, x_2))$, we find the marginal rate of substitution of good 2 for good 1 to be

$$MRS = \frac{dx_2}{dx_1}_{|U=\bar{U}} = \frac{\partial U/\partial x_1}{\partial U/\partial x_2} = \frac{F' f_1(x_1, x_2)}{F' f_2(x_1, x_2)} = \frac{f_1(1, \frac{x_2}{x_1})}{f_2(1, \frac{x_2}{x_1})}. \tag{4.43}$$

The last equality is due to Euler's theorem saying that when f is homogeneous of degree 1, then the first-order partial derivatives of f are homogeneous of degree 0. Now, (4.43) implies that for a given MRS, in optimum reflecting a given relative price of the two goods, the same consumption ratio, x_2/x_1, will be chosen whatever the budget. For a given relative price, a rising budget (rising wealth) will change the position of the budget line, but not its slope. So MRS will not change, which implies that the chosen pair (x_1, x_2) will move outward along a given ray in $\mathbb{R}^2_+$. Indeed, as the intercepts with the axes rise proportionately with the budget (the wealth), so will x_1 and x_2.

Proof that the utility function in (4.25) is homothetic In Section 4.2 we claimed that (4.25) is a homothetic utility function. This can be proved in the following way. There are two cases to consider. *Case 1:* $\theta > 0$, $\theta \neq 1$. We rewrite (4.25) as

$$U(c_1, c_2) = \frac{1}{1-\theta}\left[(c_1^{1-\theta} + \beta c_2^{1-\theta})^{1/(1-\theta)}\right]^{1-\theta} - \frac{1+\beta}{1-\theta},$$

where $\beta \equiv (1+\rho)^{-1}$. The function $x = g(c_1, c_2) \equiv (c_1^{1-\theta} + \beta c_2^{1-\theta})^{1/(1-\theta)}$ is homogeneous of degree one and the function $G(x) \equiv (1/(1-\theta))x^{1-\theta} - (1+\beta)/(1-\theta)$ is an increasing function, given $\theta > 0$, $\theta \neq 1$. *Case 2:* $\theta = 1$. Here we start from $U(c_1, c_2) = \ln c_1 + \beta \ln c_2$. This can be written

$$U(c_1, c_2) = (1+\beta)\ln\left[(c_1 c_2^{\beta})^{1/(1+\beta)}\right],$$

where $x = g(c_1, c_2) = (c_1 c_2^{\beta})^{1/(1+\beta)}$ is homogeneous of degree one and $G(x) \equiv (1+\beta)\ln x$ is an increasing function. □

D. General formulas for the elasticity of factor substitution

We here prove (4.34) and (4.35). Given the neoclassical production function $F(K, L)$, the slope of the isoquant $F(K, L) = \bar{Y}$ at the point $(\bar{K}, \bar{L})$ is

$$\frac{dK}{dL}_{|Y=\bar{Y}} = -MRS = -\frac{F_L(\bar{K}, \bar{L})}{F_K(\bar{K}, \bar{L})}. \tag{4.44}$$

We consider this slope as a function of the value of $k \equiv K/L$ as we move along the isoquant. The derivative of this function is

$$\begin{aligned} -\frac{dMRS}{dk}_{|Y=\bar{Y}} &= -\frac{dMRS}{dL}_{|Y=\bar{Y}}\frac{dL}{dk}_{|Y=\bar{Y}} \\ &= -\frac{(F_L)^2 F_{KK} - 2F_K F_L F_{KL} + (F_K)^2 F_{LL}}{F_K^3}\frac{dL}{dk}_{|Y=\bar{Y}} \end{aligned} \tag{4.45}$$

by (2.53) of Chapter 2. In view of $L \equiv K/k$ we have

$$\frac{dL}{dk}_{|Y=\bar{Y}} = \frac{k\frac{dK}{dk}_{|Y=\bar{Y}} - K}{k^2} = \frac{k\frac{dK}{dL}_{|Y=\bar{Y}}\frac{dL}{dk}_{|Y=\bar{Y}} - K}{k^2} = \frac{-kMRS\frac{dL}{dk}_{|Y=\bar{Y}} - K}{k^2}.$$

From this we find

$$\frac{dL}{dk}_{|Y=\bar{Y}} = -\frac{K}{(k+MRS)k},$$

to be substituted into (4.45). Finally, we substitute the inverse of (4.45) together with (4.44) into the definition of the elasticity of factor substitution:

$$\begin{aligned}
\sigma(K,L) &\equiv \frac{MRS}{k}\frac{dk}{dMRS}_{|Y=\bar{Y}} \\
&= -\frac{F_L/F_K}{k}\frac{(k+F_L/F_K)k}{K}\frac{F_K^3}{[(F_L)^2F_{KK}-2F_KF_LF_{KL}+(F_K)^2F_{LL}]} \\
&= -\frac{F_KF_L(F_KK+F_LL)}{KL\left[(F_L)^2F_{KK}-2F_KF_LF_{KL}+(F_K)^2F_{LL}\right]},
\end{aligned}$$

which is the same as (4.34).

Under CRS, this reduces to

$$\begin{aligned}
\sigma(K,L) &= -\frac{F_KF_LF(K,L)}{KL\left[(F_L)^2F_{KK}-2F_KF_LF_{KL}+(F_K)^2F_{LL}\right]} \quad \text{(from (2.54) with } h=1) \\
&= -\frac{F_KF_LF(K,L)}{KLF_{KL}\left[-(F_L)^2L/K-2F_KF_L-(F_K)^2K/L\right]} \quad \text{(from (2.60))} \\
&= \frac{F_KF_LF(K,L)}{F_{KL}(F_LL+F_KK)^2} = \frac{F_KF_L}{F_{KL}F(K,L)}, \quad \text{(using (2.54) with } h=1)
\end{aligned}$$

which proves the first part of (4.35). The second part is an implication of rewriting the formula in terms of the production in intensive form.

E. Properties of the CES production function

The generalized CES production function is

$$Y = A\left[\alpha K^\beta + (1-\alpha)L^\beta\right]^{\frac{\gamma}{\beta}}, \tag{4.46}$$

where A, α, and β are parameters satisfying $A > 0$, $0 < \alpha < 1$, and $\beta < 1$, $\beta \neq 0, \gamma > 0$. If $\gamma < 1$, there is DRS, if $\gamma = 1$, CRS, and if $\gamma > 1$, IRS. The elasticity of substitution is always $\sigma = 1/(1-\beta)$. Throughout below, k means K/L.

The limiting functional forms We claimed in the text that, for fixed $K > 0$ and $L > 0$, (4.46) implies:

$$\begin{aligned}
\lim_{\beta\to 0} Y &= A(K^\alpha L^{1-\alpha})^\gamma = AL^\gamma k^{\alpha\gamma}, &(*)\\
\lim_{\beta\to -\infty} Y &= A\min(K^\gamma, L^\gamma) = AL^\gamma \min(k^\gamma, 1). &(**)
\end{aligned}$$

Proof. Let $q \equiv Y/(AL^\gamma)$. Then $q = (\alpha k^\beta + 1 - \alpha)^{\gamma/\beta}$ so that

$$\ln q = \frac{\gamma \ln(\alpha k^\beta + 1 - \alpha)}{\beta} \equiv \frac{m(\beta)}{\beta}, \tag{4.47}$$

where

$$m'(\beta) = \frac{\gamma \alpha k^\beta \ln k}{\alpha k^\beta + 1 - \alpha} = \frac{\gamma \alpha \ln k}{\alpha + (1-\alpha)k^{-\beta}}. \tag{4.48}$$

Hence, by L'Hôpital's rule for "0/0",

$$\lim_{\beta \to 0} \ln q = \lim_{\beta \to 0} \frac{m'(\beta)}{1} = \gamma\alpha \ln k = \ln k^{\gamma\alpha},$$

so that $\lim_{\beta \to 0} q = k^{\gamma\alpha}$, which proves (*). As to (**), note that

$$\lim_{\beta \to -\infty} k^\beta = \lim_{\beta \to -\infty} \frac{1}{k^{-\beta}} \to \begin{cases} 0 \text{ for } k > 1, \\ 1 \text{ for } k = 1, \\ \infty \text{ for } k < 1. \end{cases}$$

Hence, by (4.47),

$$\lim_{\beta \to -\infty} \ln q = \begin{cases} 0 \text{ for } k \geq 1, \\ \lim_{\beta \to -\infty} \frac{m'(\beta)}{1} = \gamma \ln k = \ln k^\gamma \text{ for } k < 1, \end{cases}$$

where the result for $k < 1$ is based on L'Hôpital's rule for "∞/-∞". Consequently,

$$\lim_{\beta \to -\infty} q = \begin{cases} 1 \text{ for } k \geq 1, \\ k^\gamma \text{ for } k < 1, \end{cases}$$

which proves (**). □

Properties of the isoquants of the CES function The absolute value of the slope of an isoquant for (4.46) in the (L, K) plane is

$$MRS_{KL} = \frac{\partial Y/\partial L}{\partial Y/\partial K} = \frac{1-\alpha}{\alpha} k^{1-\beta} \to \begin{cases} 0 \text{ for } k \to 0, \\ \infty \text{ for } k \to \infty. \end{cases} \tag{*}$$

This holds whether $\beta < 0$ or $0 < \beta < 1$.

Concerning the asymptotes and terminal points, if any, of the isoquant $Y = \bar{Y}$ we have from (4.46) $\bar{Y}^{\beta/\gamma} = A\left[\alpha K^\beta + (1-\alpha)L^\beta\right]$. Hence,

$$K = \left(\frac{\bar{Y}^{\frac{\beta}{\gamma}}}{A\alpha} - \frac{1-\alpha}{\alpha} L^\beta \right)^{\frac{1}{\beta}},$$

$$L = \left(\frac{\bar{Y}^{\frac{\beta}{\gamma}}}{A(1-\alpha)} - \frac{\alpha}{1-\alpha} K^\beta \right)^{\frac{1}{\beta}}.$$

From these two equations follows, when $\beta < 0$ (i.e., $0 < \sigma < 1$), that

$$\begin{aligned} K &\rightarrow (A\alpha)^{-\frac{1}{\beta}}\bar{Y}^{\frac{1}{\gamma}} \text{ for } L \rightarrow \infty, \\ L &\rightarrow [A(1-\alpha)]^{-\frac{1}{\beta}}\bar{Y}^{\frac{1}{\gamma}} \text{ for } K \rightarrow \infty. \end{aligned}$$

When instead $\beta > 0$ (i.e., $\sigma > 1$), the same limiting formulas obtain for $L \rightarrow 0$ and $K \rightarrow 0$, respectively.

Properties of the CES function on intensive form Given $\gamma = 1$, i.e., CRS, we have $y \equiv Y/L = A(\alpha k^{\beta} + 1 - \alpha)^{1/\beta}$ from (4.46). Then

$$\frac{dy}{dk} = A\frac{1}{\beta}(\alpha k^{\beta} + 1 - \alpha)^{\frac{1}{\beta}-1}\alpha\beta k^{\beta-1} = A\alpha\left[\alpha + (1-\alpha)k^{-\beta}\right]^{\frac{1-\beta}{\beta}}.$$

Hence, when $\beta < 0$ (i.e., $0 < \sigma < 1$),

$$\begin{aligned} y &= \frac{A}{(ak^{\beta} + 1 - \alpha)^{-1/\beta}} \rightarrow \begin{cases} 0 \text{ for } k \rightarrow 0, \\ A(1-\alpha)^{1/\beta} \text{ for } k \rightarrow \infty. \end{cases} \\ \frac{dy}{dk} &= \frac{A\alpha}{[\alpha + (1-\alpha)k^{-\beta}]^{(\beta-1)/\beta}} \rightarrow \begin{cases} A\alpha^{1/\beta} \text{ for } k \rightarrow 0, \\ 0 \text{ for } k \rightarrow \infty. \end{cases} \end{aligned}$$

If instead $\beta > 0$ (i.e., $\sigma > 1$),

$$\begin{aligned} y &\rightarrow \begin{cases} A(1-\alpha)^{1/\beta} \text{ for } k \rightarrow 0, \\ \infty \text{ for } k \rightarrow \infty. \end{cases} \\ \frac{dy}{dk} &\rightarrow \begin{cases} \infty \text{ for } k \rightarrow 0, \\ A\alpha^{1/\beta} \text{ for } k \rightarrow \infty. \end{cases} \end{aligned}$$

The output-capital ratio is $y/k = A\left[\alpha + (1-\alpha)k^{-\beta}\right]^{\frac{1}{\beta}}$ and has the same limiting values as dy/dk, when $\beta > 0$.

Continuity at the boundary of $\mathbb{R}^2_+$ When $0 < \beta < 1$, the right-hand side of (4.46) is defined and continuous also on the boundary of $\mathbb{R}^2_+$. Indeed, we get

$$Y = F(K,L) = A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{\gamma}{\beta}} \rightarrow \begin{cases} A\alpha^{\frac{\gamma}{\beta}}K^{\gamma} \text{ for } L \rightarrow 0, \\ A(1-\alpha)^{\frac{\gamma}{\beta}}L^{\gamma} \text{ for } K \rightarrow 0. \end{cases}$$

When $\beta < 0$, however, the right-hand side is not defined on the boundary. We circumvent this problem by redefining the CES function in the following way when $\beta < 0$:

$$Y = F(K,L) = \begin{cases} A\left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{\gamma}{\beta}} \text{ when } K > 0 \text{ and } L > 0, \\ 0 \text{ when either } K \text{ or } L \text{ equals } 0. \end{cases} \tag{4.49}$$

We now show that continuity holds in the extended domain. When $K > 0$ and $L > 0$, we have

$$Y^{\frac{\beta}{\gamma}} = A^{\frac{\beta}{\gamma}} \left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right] \equiv A^{\frac{\beta}{\gamma}} G(K, L). \tag{4.50}$$

Let $\beta < 0$ and $(K, L) \to (0, 0)$. Then, $G(K, L) \to \infty$, and so $Y^{\beta/\gamma} \to \infty$. Since $\beta/\gamma < 0$, this implies $Y \to 0 = F(0, 0)$, where the equality follows from the definition in (4.49). Next, consider a fixed $L > 0$ and rewrite (4.50) as

$$\begin{aligned} Y^{\frac{1}{\gamma}} &= A^{\frac{1}{\gamma}} \left[\alpha K^{\beta} + (1-\alpha)L^{\beta}\right]^{\frac{1}{\beta}} = A^{\frac{1}{\gamma}} L(\alpha k^{\beta} + 1 - \alpha)^{\frac{1}{\beta}} \\ &= \frac{A^{\frac{1}{\gamma}} L}{(ak^{\beta} + 1 - \alpha)^{-1/\beta}} \to 0 \text{ for } k \to 0, \end{aligned}$$

when $\beta < 0$. Since $1/\gamma > 0$, this implies $Y \to 0 = F(0, L)$, from (4.49). Finally, consider a fixed $K > 0$ and let $L/K \to 0$. Then, by an analogue argument we get $Y \to 0 = F(K, 0)$, (4.49). So continuity is maintained in the extended domain.

4.10 Exercises

4.1 *(the aggregate saving rate in steady state)*

a) In a well-behaved Diamond OLG model let n be the rate of population growth and k^* the steady state capital-labor ratio (until further notice, we ignore technological progress). Derive a formula for the long-run aggregate net saving rate, S^N/Y, in terms of n and k^*. *Hint:* use that for a closed economy $S^N = K_{t+1} - K_t$.

b) In the Solow growth model without technological change a similar relation holds, but with a different interpretation of the causality. Explain.

c) Compare your result in a) with the formula for S^N/Y in steady state one gets in *any* model with the same CRS-production function and no technological change. Comment.

d) Assume that $n = 0$. What does the formula from a) tell you about the level of net aggregate savings in this case? Give the intuition behind the result in terms of the aggregate saving by any generation in two consecutive periods. One might think that people's rate of impatience (in Diamond's model the rate of time preference ρ) affect S^N/Y in steady state. Does it in this case? Why or why not?

e) Suppose there is Harrod-neutral technological progress at the constant rate $g > 0$. Derive a formula for the aggregate net saving rate in the long run in a well-behaved Diamond model in this case.

f) Answer d) with "from a)" replaced by "from e)". Comment.

g) Consider the statement: "In Diamond's OLG model any generation saves as much when young as it dissaves when old." True or false? Why?

4.2 (*increasing returns to scale and balanced growth)*