Chapter 4

Skill-biased technical change. Balanced growth theorems

This chapter is both an alternative and a supplement to the pages 60-64 in Acemoglu, where the concepts of neutral technical change and balanced growth, including Uzawa's theorem, are discussed.

Since "neutral" technical change should be seen in relation to "biased" technical change, Section 1 below introduces the concept of "biased" technical change. Like regarding neutral technical change, also regarding biased technical change there exist three different definitions, Hicks', Harrod's, and what the literature has dubbed "Solow's". Below we concentrate on Hick's definition − with an application to the role of technical change for the evolution of the skill premium. So the focus is on the production factors skilled and unskilled labor rather than capital and labor. While regarding capital and labor it is Harrod's classifications that are most used in macroeconomics, regarding skilled and unskilled labor it is Hicks'.

The remaining sections discuss the concept of balanced growth and present three fundamental propositions about balanced growth. In view of the generality of the propositions, they have a broad field of application. Our propositions 1 and 2 are slight extensions of part 1 and 2, respectively, of what Acemoglu calls Uzawa's Theorem I (Acemoglu, 2009, p. 60). Our Proposition 3 essentially corresponds to what Acemoglu calls Uzawa's Theorem II (Acemoglu, 2009, p. 63).

4.1 The rising skill premium

4.1.1 Skill-biased technical change in the sense of Hicks: An example

Let aggregate output be produced through a differentiable three-factor production function $\tilde{F}$:

$$Y = \tilde{F}(K, L_1, L_2, t),$$

where K is capital input, L_1 is input of unskilled labor (also called blue-collar labor below), and L_2 is input of skilled labor. Suppose technological change is such that the production function can be rewritten

$$\tilde{F}(K, L_1, L_2, t) = F(K, H(L_1, L_2, t)), \tag{4.1}$$

where the "nested" function $H(L_1, L_2, t)$ represents input of a "human capital" aggregate. Let F be CRS-neoclassical w.r.t. K and H and let H be CRS-neoclassical w.r.t. (L_1, L_2). Finally, let $\partial H/\partial t > 0$. So "technical change" amounts to "technical progress".

In equilibrium under perfect competition in the labor markets the relative wage, often called the "skill premium", will be

$$\frac{w_2}{w_1} = \frac{\partial Y/\partial L_2}{\partial Y/\partial L_1} = \frac{F_H \partial H/\partial L_2}{F_H \partial H/\partial L_1} = \frac{H_2(L_1, L_2, t)}{H_1(L_1, L_2, t)} = \frac{H_2(1, L_2/L_1, t)}{H_1(1, L_2/L_1, t)}, \tag{4.2}$$

where we have used Euler's theorem (saying that if H is homogeneous of degree one in its first two arguments, then the partial derivatives of H are homogeneous of degree zero w.r.t. these arguments).

Time is continuous (nevertheless the time argument of a variable, x, is in this section written as a subscript t). Hicks' definitions are now: If for all $L_2/L_1 > 0$,

$$\frac{d\left(\frac{H_2(1,L_2/L_1,t)}{H_1(1,L_2/L_1,t)}\right)}{dt}\Bigg|_{\frac{L_2}{L_1}\text{ constant}} \gtreqless 0, \text{ then technical change is}$$

$$\begin{cases} \text{skill-biased in the sense of Hicks,} \\ \text{skill-neutral in the sense of Hicks.} \\ \text{blue collar-biased in the sense of Hicks,} \end{cases} \tag{4.3}$$

respectively.

In the US the skill premium (measured by the wage ratio for college grads vis-a-vis high school grads) has had an upward trend since 1950 (see for instance Jones and Romer, 2010).[1] If in the same period the relative supply of skilled labor had been roughly constant, by (17.3) in combination with (17.2), a possible explanation could be that technological change has been skill-biased in the sense of Hicks. In reality, in the same period also the relative supply of skilled labor has been rising (in fact even faster than the skill premium). Since in spite of this the skill premium has *risen,* it suggests that the extend of "skill-biasedness" has been even stronger.

We may alternatively put it this way. As the H function is CRS-neoclassical w.r.t. L_1 and L_2, we have $H_{22} < 0$ *and* $H_{12} > 0$, cf. Chapter 2. Hence, by (17.2), a rising L_2/L_1 without technical change would imply a *declining* skill premium. That the opposite has happened must, within our simple model, be due to (a) there *has* been technical change, and (b) technical change has *favoured skilled labor* (which means that technical change has been skill-biased in the sense of Hicks).

An additional aspect of the story is that skill-biasedness helps *explain* the observed increase in the relative *supply* of skilled labor. If for a constant relative supply of skilled labor, the skill premium is increasing, this increase strengthens the incentive to go to college. Thereby the relative supply of skilled labor (reflecting the fraction of skilled labor in the labor force) tends to increase.

4.1.2 Capital-skill complementarity

An additional potential source of a rising skill premium is *capital-skill complementarity.* Let the aggregate production function be

$$Y = \tilde{F}(K, L_1, L_2, t) = F(K, A_{1t}L_1, A_{2t}L_2) = (K+A_{1t}L_1)^{\alpha}(A_{2t}L_2)^{1-\alpha},\ 0 < \alpha < 1,$$

where A_{1t} and A_{2t} are technical coefficients that may be rising over time. In this production function capital and unskilled labor are perfectly substitutable (the partial elasticity of factor substitution between them is $+\infty$). On the other hand there is *direct complementarity* between capital and skilled labor, i.e., $\partial^2 Y/(\partial L_2 \partial K) > 0$.

Under perfect competition the skill premium is

$$\begin{aligned} \frac{w_2}{w_1} &= \frac{\partial Y/\partial L_2}{\partial Y/\partial L_2} = \frac{(K + A_{1t}L_1)^{\alpha}(1-\alpha)(A_{2t}L_2)^{-\alpha}A_{2t}}{\alpha(K + A_{1t}L_1)^{\alpha-1}A_{1t}(A_{2t}L_2)^{1-\alpha}} \\ &= \frac{1-\alpha}{\alpha}\left(\frac{K + A_{1t}L_1}{A_{2t}L_2}\right)\frac{A_{2t}}{A_{1t}}. \end{aligned} \tag{4.4}$$

[1] On the other hand, over the years 1915 - 1950 the skill premium had a downward trend (Jones and Romer, 2010).

Here, if technical change is absent (A_{1t} and A_{2t} constant), a rising capital stock will, *for fixed* L_1 *and* L_2, raise the skill premium.

A more realistic scenario is, however, a situation with an approximately constant real interest rate, cf. Kaldor's stylized facts. We have, again by perfect competition,

$$\frac{\partial Y}{\partial K} = \alpha(K + A_{1t}L_1)^{\alpha-1}(A_{2t}L_2)^{1-\alpha} = \alpha\left(\frac{K + A_{1t}L_1}{A_{2t}L_2}\right)^{\alpha-1} = r_t + \delta, \quad (4.5)$$

where r_t is the real interest rate at time t and δ is the (constant) capital depreciation rate. For $r_t = r$, a constant, (17.5) gives

$$\frac{K + A_{1t}L_1}{A_{2t}L_2} = \left(\frac{r+\delta}{\alpha}\right)^{-\frac{1}{1-\alpha}} \equiv c, \quad (4.6)$$

a constant. In this case, (17.4) shows that capital-skill complementarity is *not sufficient* for a rising skill premium. A rising skill premium requires that technical change brings about a rising A_{2t}/A_{1t}. So again an observed rising skill premium, along with a more or less constant real interest rate, suggests that technical change is skill-biased.

We may rewrite (4.6) as

$$\frac{K}{A_{2t}L_2} = c - \frac{A_{1t}L_1}{A_{2t}L_2},$$

where the conjecture is that $A_{1t}L_1/(A_{2t}L_2) \rightarrow 0$ for $t \rightarrow \infty$. The analysis suggests the following story. Skill-biased technical progress generates rising productivity as well as a rising skill premium. The latter induces more and more people to go to college. The rising level of education in the labor force raises productivity further. This is a basis for further capital accumulation, continuing to replace unskilled labor, and so on.

In particular since the early 1980s the skill premium has been sharply increasing in the US (see Acemoglu, p. 498). This is also the period where ICT technologies took off.

4.2 Balanced growth and constancy of key ratios

The focus now shifts to homogeneous labor vis-a-vis capital.

We shall state general definitions of the concepts of "steady state" and "balanced growth", concepts that are related but not identical. With respect to "balanced growth" this implies a minor deviation from the way Acemoglu briefly defines it informally on his page 57. The main purpose of the present chapter is to lay bare the connections between these two concepts as well as their relation to the hypothesis of Harrod-neutral technical progress and Kaldor's stylized facts.

4.2.1 The concepts of steady state and balanced growth

A basic equation in many one-sector growth models for a closed economy in continuous time is

$$\dot{K} = I - \delta K = Y - C - \delta K \equiv S - \delta K, \tag{4.7}$$

where K is aggregate capital, I aggregate gross investment, Y aggregate output, C aggregate consumption, S aggregate gross saving ($\equiv Y - C$), and $\delta \geq 0$ is a constant physical capital depreciation rate.

Usually, in the theoretical literature on dynamic models, a *steady state* is defined in the following way:

Definition 3 *A steady state of a dynamic model is a stationary solution to the fundamental differential equation(s) of the model.*

Or briefly: a steady state is a stationary point of a dynamic process.

Let us take the Solow growth model as an example. Here gross saving equals sY, where s is a constant, $0 < s < 1$. Aggregate output is given by a neoclassical production function, F, with CRS and Harrod-neutral technical progress: $Y = F(K, AL) = ALF(\tilde{k}, 1) \equiv ALf(\tilde{k})$, where L is the labor force, A is the level of technology, and $\tilde{k} \equiv K/(AL)$ is the (effective) capital intensity. Moreover, $f' > 0$ and $f'' < 0$. Solow assumes $L(t) = L(0)e^{nt}$ and $A(t) = A(0)e^{gt}$, where $n \geq 0$ and $g \geq 0$ are the constant growth rates of the labor force and technology, respectively. By log-differentiating $\tilde{k}$ w.r.t. t,[2] we end up with the *fundamental differential equation* ("law of motion") of the Solow model:

$$\dot{\tilde{k}} = sf(\tilde{k}) - (\delta + g + n)\tilde{k}. \tag{4.8}$$

Thus, in the Solow model, a (non-trivial) steady state is a $\tilde{k}^* > 0$ such that, if $\tilde{k} = \tilde{k}^*$, then $\dot{\tilde{k}} = 0$. In passing we note that, by (4.8), such a $\tilde{k}^*$ must satisfy the equation $f(\tilde{k}^*)/\tilde{k}^* = (\delta + g + n)/s$, and in view of $f'' < 0$, it is unique if it exists.

The most common definition in the literature of *balanced growth* for an aggregate economy is the following:

[2] Or by directly using the fraction rule, see Appendix A to Chapter 3.

Definition 4 *A balanced growth path is a path* $(Y, K, C)_{t=0}^{\infty}$ *along which the quantities* $Y, K,$ *and* C *are positive and grow at constant rates (not necessarily positive and not necessarily the same).*

Acemoglu, however, defines (Acemoglu, 2009, p. 57) balanced growth in the following way: "balanced growth refers to an allocation where output grows at a constant rate and capital-output ratio, the interest rate, and factor shares remain constant". My problem with this definition is that it mixes growth of aggregate quantities with income distribution aspects (interest rate and factor income shares). And it is not made clear what is meant by the output-capital ratio if the relative price of capital goods is changing over time. So I stick to the definition above which is quite standard and is known to function well in many different contexts.

Note that in the Solow model (as well as in many other models) we have that if the economy is in a steady state, $\tilde{k} = \tilde{k}^*$, then the economy features balanced growth. Indeed, a steady state of the Solow model implies by definition that $\tilde{k} \equiv K/(AL)$ is constant. Hence K must grow at the same *constant* rate as AL, namely $g + n$. In addition, $Y = f(\tilde{k}^*)AL$ in a steady state, showing that also Y must grow at the constant rate $g + n$. And so must then $C = (1 - s)Y$. So in a steady state of the Solow model the path followed by $(Y, K, C)_{t=0}^{\infty}$ is a balanced growth path.

As we shall see in the next section, in the Solow model (and many other models) the reverse also holds: if the economy features balanced growth, then it is in a steady state. But this equivalence between steady state and balanced growth does not hold in all models.

4.2.2 A general result about balanced growth

An interesting fact is that, given the dynamic resource constraint (4.7), we have *always* that if there is balanced growth with positive gross saving, then the ratios Y/K and C/Y are constant (by "*always*" is meant: independently of how saving is determined and of how the labor force and technology evolve). And also the other way round: as long as gross saving is positive, constancy of the Y/K and C/Y ratios is enough to ensure balanced growth. So balanced growth and constancy of key ratios are essentially equivalent.

This is a very practical general observation. And since Acemoglu does not state any balanced growth theorem at this general level, we shall do it here, together with a proof. Letting g_x denote the growth rate of the (positively valued) variable x, i.e., $g_x \equiv \dot{x}/x$, we claim:

Proposition 1 *(the balanced growth equivalence theorem). Let* $(Y, K, C)_{t=0}^{\infty}$

be a path along which Y, K, C, and $S \equiv Y - C$ are positive for all $t \geq 0$. Then, given the accumulation equation (4.7), the following holds:

(i) if there is balanced growth, 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 Y, K, and C grow at the same constant rate, i.e., not only is there balanced growth, but the growth rates of Y, K, and C are the same.

Proof Consider a path $(Y, K, C)_{t=0}^{\infty}$ along which Y, K, C, and $S \equiv Y - C$ are positive for all $t \geq 0$. (i) Assume there is balanced growth. Then, by definition, g_Y, g_K, and g_C are constant. Hence, by (4.7), we have that $S/K = g_K + \delta$ is constant, implying

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

Further, since $Y = C + S$,

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

Now, let us provisionally assume that $g_K \neq g_C$. Then (**) gives

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

which is a constant since g_Y, g_K, and g_C are constant. Constancy of C/Y requires that $g_C = g_Y$, hence, by (***), $C/Y = 1$, i.e., $C = Y$. In view of $Y = C + S$, however, this outcome contradicts the given condition that $S > 0$. Hence, our provisional assumption and its implication, (***), are falsified. Instead we have $g_K = g_C$. 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 constant. Then $g_Y = g_K = g_C$, so that C/K is a constant. We now show that this implies that g_K is constant. Indeed, from (4.7), $S/Y = 1 - C/Y$, so that also S/Y is constant. It follows that $g_S = g_Y = g_K$, so that S/K is constant. By (4.7),

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

so that 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 g_C and g_Y had a constant growth rate. This growth rate, however, could easily differ from that of K. Suppose $Y = AK^{\alpha}L^{1-\alpha}$, $g_A = \gamma$ and $g_L = n$ (γ and n constants). Then we would have $g_C = g_Y = \gamma - \alpha\delta + (1-\alpha)n$, which could easily be strictly positive and thereby different from $g_K = -\delta \leq 0$ so that (i) no longer holds. □

The nice feature is that this proposition holds for *any* model for which the simple dynamic resource constraint (4.7) is valid. No assumptions about for example CRS and other technology aspects or about market form are involved. Note also that Proposition 1 suggests a link from balanced growth to steady state. And such a link *is* present in for instance the Solow model. Indeed, by (i) of Proposition 1, balanced growth implies constancy of Y/K, which in the Solow model implies that $f(\tilde{k})/\tilde{k}$ is constant. In turn, the latter is only possible if $\tilde{k}$ is constant, that is, if the economy is in steady state.

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 distinction between the terms steady state and balanced growth.

4.3 The crucial role of Harrod-neutrality

Proposition 1 suggests that if one accepts Kaldor's stylized facts (see Chapter 1) as a characterization of the past century's growth experience, and if one wants a model consistent with them, one should construct the model such that it can generate balanced growth. For a model to be capable of generating balanced growth, however, technological progress must be of the Harrod-neutral type (i.e., be labor-augmenting), at least in a neighborhood of the balanced growth path. For a fairly general context (but of course not as general as that of Proposition 1), this was shown already by Uzawa (1961). We now present a modernized version of Uzawa's contribution.

Let the aggregate production function be

$$Y(t) = \tilde{F}(K(t), BL(t), t), \qquad B > 0, \tag{4.9}$$

where B is a constant that depends on measurement units. The only technology assumption needed is that $\tilde{F}$ has CRS w.r.t. the first two arguments ($\tilde{F}$ need not be neoclassical for example). As a representation of technical progress, we assume $\partial\tilde{F}/\partial t > 0$ for all $t \geq 0$ (i.e., as time proceeds, unchanged inputs result in more and more output). We also assume that the labor force evolves according to

$$L(t) = L(0)e^{nt}, \tag{4.10}$$

where n is a constant. Further, non-consumed output is invested and so (4.7) is the dynamic resource constraint of the economy.

Proposition 2 *(Uzawa's balanced growth theorem) Let* $P = (Y(t), K(t), C(t))_{t=0}^{\infty}$, *where* $0 < C(t) < Y(t)$ *for all* $t \geq 0$, *be a path satisfying the capital accumulation equation (4.7), given the CRS-production function (4.9) and the labor force path in (4.10). Then:*

(i) a necessary condition for this path to be a balanced growth path is that along the path it holds that

$$Y(t) = \tilde{F}(K(t), BL(t), t) = \tilde{F}(K(t), A(t)L(t), 0), \tag{4.11}$$

where $A(t) = Be^{gt}$ *with* $g \equiv g_Y - n$;

(ii) for any $g > 0$ *such that there is a* $q > \delta + g + n$ *with the property that the production function* $\tilde{F}$ *in (4.9) 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 the path* P *to be a balanced growth path with output-capital ratio* q*, is that the technology can be written as in (4.11) with* $A(t) = Be^{gt}$.

Proof (i)[3] Suppose the path $(Y(t), K(t), C(t))_{t=0}^{\infty}$ is a balanced growth path. By definition, g_K and g_Y are then constant, so that $K(t) = K(0)e^{g_K t}$ and $Y(t) = Y(0)e^{g_Y t}$. We then have

$$Y(t)e^{-g_Y t} = Y(0) = \tilde{F}(K(0), BL(0), 0) = \tilde{F}(K(t)e^{-g_K t}, BL(t)e^{-nt}, 0), \quad (*)$$

where we have used (4.9) with $t = 0$. In view of the precondition 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, (*) then implies

$$Y(t) = \tilde{F}(K(t)e^{g_Y t}e^{-g_K t}, BL(t)e^{g_Y t}e^{-nt}, 0) = \tilde{F}(K(t), Be^{(g_Y - n)t}L(t), 0).$$

We see that (4.11) holds for $A(t) = Be^{gt}$ with $g \equiv g_Y - n$.

[3] This part draws upon Schlicht (2006), who generalized a proof in Wan (1971, p. 59) for the special case of a constant saving rate.

(ii) Suppose (4.11) holds with $A(t) = Be^{gt}$. Let $g > 0$ be given such that there is a $q > g + n + \delta$ with the property that

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

for some constant $\tilde{k} > 0$. Our strategy is to prove the claim in (ii) 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 $s \equiv (\delta + g + n)/q$, i.e., $Y(t) - C(t) \equiv S(t) = sY(t)$ for all $t \geq 0$. Inserting this, together with $Y(t) = f(\tilde{k}(t))A(t)L(t)$, where $f(\tilde{k}(t)) \equiv \tilde{F}(\tilde{k}(t), 1, 0)$ and $\tilde{k}(t) \equiv K(t)/(A(t)L(t))$, into (4.7), we get the Solow equation (4.8). Hence $\tilde{k}(t)$ is constant if and only if $\tilde{k}(t)$ satisfies the equation $f(\tilde{k}(t))/\tilde{k}(t) = (\delta + g + n)/s \equiv q$. 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 equation. Letting $K(0)$ satisfy $K(0) = \tilde{k}BL(0)$, where $B = A(0)$, we thus have $\tilde{k}(0) = K(0)/(A(0)L(0)) = \tilde{k}$. So that the initial value of $\tilde{k}(t)$ equals the steady state value. It now follows that $\tilde{k}(t) = \tilde{k}$ for all $t \geq 0$, 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, $C(t) = (1 - s)Y(t)$, so that $C(t)/Y(t)$ is constant along the path P. By (ii) of Proposition 1 now follows that the path P is a balanced growth path, as was to be proved. □

The form (4.11) indicates that along a balanced growth path, technical progress must be purely "labor augmenting", that is, Harrod-neutral. It is in this case convenient to define a new CRS function, F, by $F(K(t), A(t)L(t)) \equiv \tilde{F}(K(t), A(t)L(t), 0)$. Then (i) of the proposition implies that at least along the balanced growth path, we can rewrite the production function this way:

$$Y(t) = \tilde{F}(K(t), A(0)L(t), t) = F(K(t), A(t)L(t)), \tag{4.12}$$

where $A(0) = B$, $A(t) = A(0)e^{gt}$ with $g \equiv g_Y - n$.

It is important to recognize that the occurrence of Harrod-neutrality says nothing about what the *source* of technological progress is. Harrod-neutrality should not be interpreted as indicating that the technological progress emanates specifically from the labor input. Harrod-neutrality only means that technical innovations predominantly are such that not only do labor and capital in combination become more productive, but this happens to *manifest itself* at the aggregate level in the form (4.12).[4]

What is the intuition behind the Uzawa result that for balanced growth to be possible, technical progress must have the purely labor-augmenting form?

[4]For a CRS Cobb-Douglas production function with technological progress, Harrod-neutrality is present whenever the output elasticity w.r.t capital (often denoted α) is constant over time.

First, notice that there is an asymmetry between capital and labor. Capital is an accumulated amount of non-consumed output. In contrast, in simple macro models labor is a non-produced production factor which (at least in the context of (4.10)) grows in an exogenous way. Second, because of CRS, the original formulation, (4.9), of the production function implies that

$$1 = \tilde{F}(\frac{K(t)}{Y(t)}, \frac{L(t)}{Y(t)}, t). \tag{4.13}$$

Now, since capital is accumulated non-consumed output, it tends to inherit the trend in output such that $K(t)/Y(t)$ must be constant along a balanced growth path (this is what Proposition 1 is about). Labor does not inherit the trend in output; indeed, the ratio $L(t)/Y(t)$ is free to adjust as time proceeds. When there is technical progress ($\partial\tilde{F}/\partial t > 0$) along a balanced growth path, this progress must manifest itself in the form of a changing $L(t)/Y(t)$ in (13.5) as t proceeds, precisely because $K(t)/Y(t)$ *must* be constant along the path. In the "normal" case where $\partial\tilde{F}/\partial L > 0$, the needed change in $L(t)/Y(t)$ is a *fall* (i.e., a rise in $Y(t)/L(t)$). This is what (13.5) shows. Indeed, 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.[5] It follows that along the balanced growth path, $Y(t)/L(t)$ is an increasing implicit function of t. If we denote this function $A(t)$, we end up with (4.12) with specified properties (g and q).

The generality of Uzawa's theorem is noteworthy. The theorem assumes CRS, but does not presuppose that the technology is neoclassical, not to speak of satisfying the Inada conditions.[6] And the theorem holds for exogenous as well as endogenous technological progress. It is also worth mentioning that the proof of the sufficiency part of the theorem is *constructive*. It provides a method to construct a hypothetical balanced growth path (BGP from now).[7]

A simple implication of the Uzawa theorem is the following. Interpreting the $A(t)$ in (4.11) as the "level of technology", we have:

COROLLARY Along a BGP with positive gross saving and the technology level, $A(t)$, growing at the rate g, output grows at the rate $g + n$ while labor productivity, $y \equiv Y/L$, and consumption per unit of labor, $c \equiv C/L$, grow at the rate g.

[5] This way of presenting the intuition behind the Uzawa result draws upon Jones and Scrimgeour (2008).

[6] Many accounts of the Uzawa theorem, including Jones and Scrimgeour (2008), presume a neoclassical production function, but the theorem is much more general.

[7] Part (ii) of Proposition 2 is left out in Acemoglu's book.

Proof That $g_Y = g + n$ follows from (i) of Proposition 2. As to the growth rate of labor productivity we have

$$y_t = \frac{Y(0)e^{g_Y t}}{L(0)e^{nt}} = y(0)e^{(g_Y - n)t} = y(0)e^{gt}.$$

Finally, by Proposition 1, along a BGP with $S > 0$, c must grow at the same rate as y. □

We shall now consider the implication of Harrod-neutrality for the income shares of capital and labor when the technology is neoclassical and markets are perfectly competitive.

4.4 Harrod-neutrality and the functional income distribution

There is one facet of Kaldor's stylized facts we have so far not related to Harrod-neutral technical progress, namely the long-run "approximate" constancy of both the income share of labor, wL/Y, and the rate of return to capital. At least with neoclassical technology, profit maximizing firms, and perfect competition in the output and factor markets, these properties are inherent in the combination of constant returns to scale, balanced growth, and the assumption that the relative price of capital goods (relative to consumption goods) is constant over time. The latter condition holds in models where the capital good is nothing but non-consumed output, cf. (4.7).[8]

To see this, we start out from a neoclassical CRS production function with Harrod-neutral technological progress,

$$Y(t) = F(K(t), A(t)L(t)). \tag{4.14}$$

With $w(t)$ denoting the real wage at time t, in equilibrium under perfect competition the labor income share will be

$$\frac{w(t)L(t)}{Y(t)} = \frac{\frac{\partial Y(t)}{\partial L(t)}L(t)}{Y(t)} = \frac{F_2(K(t), A(t)L(t))A(t)L(t)}{Y(t)}. \tag{4.15}$$

In this simple model, without natural resources, (gross) capital income equals non-labor income, $Y(t) - w(t)L(t)$. Hence, if $r(t)$ denotes the (net) rate of return to capital at time t, then

$$r(t) = \frac{Y(t) - w(t)L(t) - \delta K(t)}{K(t)}. \tag{4.16}$$

[8]The reader may think of the "corn economy" example in Acemoglu, p. 28.

Denoting the (gross) capital income share by $\alpha(t)$, we can write this $\alpha(t)$ (in equilibrium) in three ways:

$$\begin{aligned}
\alpha(t) &\equiv \frac{Y(t)-w(t)L(t)}{Y(t)} = \frac{(r(t)+\delta)K(t)}{Y(t)}, \\
\alpha(t) &= \frac{F(K(t),A(t)L(t)) - F_2(K(t),A(t)L(t))A(t)L(t)}{Y(t)} = \frac{F_1(K(t),A(t)L(t))K(t)}{Y(t)}, \\
\alpha(t) &= \frac{\frac{\partial Y(t)}{\partial K(t)}K(t)}{Y(t)}, \qquad (4.17)
\end{aligned}$$

where the first row comes from (4.16), the second from (4.14) and (4.15), the third from the second together with Euler's theorem.[9] Comparing the first and the last row, we see that in equilibrium

$$\frac{\partial Y(t)}{\partial K(t)} = r(t) + \delta.$$

In this condition we recognize one of the first-order conditions in the representative firm's profit maximization problem under perfect competition, since $r(t)+\delta$ can be seen as the firm's required gross rate of return.[10]

In the absence of uncertainty, the equilibrium real interest rate in the bond market must equal the rate of return on capital, $r(t)$. And $r(t)+\delta$ can then be seen as the firm's cost of disposal over capital per unit of capital per time unit, consisting of interest cost plus capital depreciation.

Proposition 3 *(factor income shares and rate of return under balanced growth) Let the path $(K(t),Y(t),C(t))_{t=0}^{\infty}$ be a BGP in a competitive economy with the production function (4.14) and with positive saving. Then, along the BGP, the $\alpha(t)$ in (4.17) is a constant, $\alpha \in (0,1)$. The labor income share will be $1-\alpha$ and the (net) rate of return on capital will be $r = \alpha q - \delta$, where q is the constant output-capital ratio along the BGP.*

Proof By CRS we have $Y(t) = F(K(t),A(t)L(t)) = A(t)L(t)F(\tilde{k}(t),1) \equiv A(t)L(t)f(\tilde{k}(t))$. In view of part (i) of Proposition 2, by balanced growth, $Y(t)/K(t)$ is some constant, q. Since $Y(t)/K(t) = f(\tilde{k}(t))/\tilde{k}(t)$ and $f'' < 0$, this implies $\tilde{k}(t)$ constant, say equal to $\tilde{k}^*$. But $\partial Y(t)/\partial K(t) = f'(\tilde{k}(t))$, which

[9] From Euler's theorem, $F_1K + F_2AL = F(K,AL)$, when F is homogeneous of degree one.

[10] With natural resources, say land, entering the set of production factors, the formula, (4.16), for the rate of return to capital should be modified by subtracting land rents from the numerator.

then equals the constant $f'(\tilde{k}^*)$ along the BGP. It then follows from (4.17) that $\alpha(t) = f'(\tilde{k}^*)/q \equiv \alpha$. Moreover, $0 < \alpha < 1$, where $0 < \alpha$ follows from $f' > 0$ and $\alpha < 1$ from the fact that $q = Y/K = f(\tilde{k}^*)/\tilde{k}^* > f'(\tilde{k}^*)$, in view of $f'' < 0$ and $f(0) \geq 0$. Then, by the first equality in (4.17), $w(t)L(t)/Y(t) = 1 - \alpha(t) = 1 - \alpha$. Finally, by (4.16), the (net) rate of return on capital is $r = (1 - w(t)L(t)/Y(t))Y(t)/K(t) - \delta = \alpha q - \delta$. $\square$

This proposition is of interest by displaying a link from balanced growth to constancy of factor income shares and the rate of return, that is, some of the "stylized facts" claimed by Kaldor. Note, however, that although the proposition implies constancy of the income shares and the rate of return, it does not *determine* them, except in terms of α and q. But both q and, generally, α are endogenous and depend on $\tilde{k}^*$,[11] which will generally be unknown as long as we have not specified a theory of saving. This takes us to theories of aggregate saving, for example the simple Ramsey model, cf. Chapter 8 in Acemoglu's book.

4.5 What if technological change is embodied?

In our presentation of technological progress above we have implicitly assumed that all technological change is *disembodied*. And the way the propositions 1, 2, and 3, are formulated assume this.

As noted in Chapter 2, *disembodied technological change* occurs when new technical knowledge advances the combined productivity of capital and labor independently of whether the workers operate old or new machines. Consider again the aggregate dynamic resource constraint (4.7) and the production function (4.9):

$$\dot{K}(t) = I(t) - \delta K(t), \tag{4.18}$$

$$Y(t) = \tilde{F}(K(t), BL(t), t), \qquad \partial\tilde{F}/\partial t > 0. \tag{4.19}$$

Here $Y(t) - C(t)$ is aggregate gross investment, $I(t)$. For a given level of $I(t)$, the resulting amount of new capital goods per time unit $(\dot{K}(t) + \delta K(t))$, measured in efficiency units, is independent of *when* this investment occurs. It is thereby not affected by technological progress. Similarly, the interpretation of $\partial\tilde{F}/\partial t > 0$ in (4.19) is that the higher technology level obtained as time proceeds results in higher productivity of *all* capital and labor. Thus also

[11] As to α, there is of course a trivial exception, namely the case where the production function is Cobb-Douglas and α therefore is a given parameter.

firms that have only old capital equipment benefit from recent advances in technical knowledge. No new investment is needed to take advantage of the recent technological and organizational developments.[12]

In contrast, we say that technological change is *embodied*, if taking advantage of new technical knowledge requires construction of new investment goods. The newest technology is incorporated in the design of newly produced equipment; and this equipment will not participate in subsequent technological progress. Whatever the source of new technical knowledge, investment becomes an important bearer of the productivity increases which this new knowledge makes possible. Without new investment, the potential productivity increases remain potential instead of being realized.

As also noted in Chapter 2, we may represent embodied technological progress (also called investment-specific technological change) by writing capital accumulation in the following way,

$$\dot{K}(t) = q(t)I(t) - \delta K(t), \tag{4.20}$$

where $I(t)$ is gross investment at time t and $q(t)$ measures the "quality" (productivity) of newly produced investment goods. The increasing level of technology implies increasing $q(t)$ so that a given level of investment gives rise to a greater and greater additions to the capital stock, K, measured in efficiency units. As in our aggregate framework, q capital goods can be produced at the same minimum cost as one consumption good, we have $p \cdot q = 1$, where p is the equilibrium price of capital goods in terms of consumption goods. So embodied technological progress is likely to result in a steady decline in the relative price of capital equipment, a prediction confirmed by the data (see, e.g., Greenwood et al., 1997).

This raises the question how the propositions 1, 2, and 3 fare in the case of embodied technological progress. The answer is that a generalized version of Proposition 1 goes through. Essentially, we only need to replace (4.7) by (13.13) and interpret K in Proposition 1 as the *value* of the capital stock, i.e., we have to replace K by $\tilde{K} = pK$.

But the concept of Harrod-neutrality no longer fits the situation without further elaboration. Hence to obtain analogies to Proposition 2 and Proposition 3 is a more complicated matter. Suffice it to say that with embodied technological progress, the class of production functions that are consistent with balanced growth is smaller than with disembodied technological progress.

[12] In the standard versions of the Solow model and the Ramsey model it is assumed that all technological progress has this form - for no other reason than that this is by far the simplest case to analyze.

4.6 Concluding remarks

In the Solow model as well as in many other models with disembodied technological progress, a steady state and a balanced growth path imply each other. Indeed, they are in that model, as well as many others, two sides of the same process. There *exist* exceptions, however, that is, cases where steady state and a balanced growth are not equivalent (some open economy models and some models with *embodied* technical change). So the two concepts should be held apart.[13]

Note that the definition of balanced growth refers to *aggregate* variables. At the same time as there is balanced growth at the aggregate level, *structural change* may occur. That is, 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, Chapter 20).

In view of the key importance of Harrod-neutrality, a natural question is: has growth theory uncovered any *endogenous* tendency for technical progress to converge to Harrod-neutrality? Fortunately, in his Chapter 15 Acemoglu outlines a theory about a mechanism entailing such a tendency, the theory of "directed technical change". Jones (2005) suggests an alternative mechanism.

4.7 References

Acemoglu, D., 2009, *Introduction to Modern Economic Growth*, Princeton University Press: Oxford.

Barro, R., and X. Sala-i-Martin, 2004, *Economic Growth*, second edition, MIT Press: Cambridge (Mass.)

Duffy. J., C. Papageorgiou, and F. Perez-Sebastian, 2004, Capital-Skill Complementarity? Evidence from a Panel of Countries, *The Review of Economics and Statistics*, vol. 86(1), 327-344.

Gordon, R. J., 1990. *The Measurement of Durable goods Prices.* Chicago University Press: Chicago.

[13] Here we deviate from Acemoglu, p. 65, where he says that he will use the two terms "interchangingly". We also deviate from Barro and Sala-i-Martin (2004, pp. 33-34) who *define* a steady state as synonymous with a balanced growth path as the latter was defined above.

Greenwood, J., Z. Hercowitz, and P. Krusell, 1997. Long-Run Implications of Investment-Specific Technological Change. *American Economic Review 87* (3), 342-362.

Groth, C., K.-J. Koch, and Thomas Steger, 2010, When growth is less than exponential, *Economic Theory 44,* 213-242.

Groth, C., and R. Wendner, 2014. Embodied Learning by Investing and Speed of Convergence, *J. of Macroeconomics* (forthcoming).

Jones, C. I., 2005, The shape of production functions and the direction of technical change. *Quarterly Journal of Economics*, no. 2, 517-549.

Jones, C. I., and D. Scrimgeour, 2008, The steady-state growth theorem: Understanding Uzawa (1961), *Review of Economics and Statistics 90* (1), 180-182.

Jones, C. I., and P. M. Romer, 2010, The new Kaldor facts: Ideas, institutions, population, and human capital, *American Economic Journal: Macroeconomics*, vol. 2 (1), 224-245. Cursory.

Kongsamut, P., S. Rebelo, and D. Xie, 2001, Beyond balanced growth. *Review of Economic Studies 48,* 869-882.

Perez-Sebastian, F., 2008, "Testing capital-skill complementarity across sectors in a panel of Spanish regions", WP 2008.

Schlicht, E., 2006, A variant of Uzawa's theorem, *Economics Bulletin 6,* 1-5.

Stokey, N.L., 1996, Free trade, factor returns, and factor accumulation, *J. Econ. Growth*, vol. 1 (4), 421-447.

Uzawa, H., 1961, Neutral inventions and the stability of growth equilibrium, *Review of Economic Studies 28,* No. 2, 117-124.

Wan, H. Y. Jr., 1971, *Economic Growth*, Harcourt Brace: New York.

Chapter 5

Growth accounting and the concept of TFP: Some warnings

5.1 Introduction

This chapter discusses the concepts of Total Factor Productivity, TFP, and TFP growth, and ends up with three warnings regarding uncritical use of them.

First, however, we should provide a precise definition of the TFP *level* which is in fact a tricky concept. Unfortunately, Acemoglu (p. 78) does not make a clear distinction between TFP *level* and TFP *growth*. Moreover, Acemoglu's point of departure (p. 77) assumes *a priori* that the way the production function is time-dependent can be represented by a one-dimensional index, $A(t)$. The TFP concept and the applicability of growth accounting are, however, not limited to this case.

For convenience, we treat time as continuous (although the timing of the variables is indicated merely by a subscript).[1]

5.2 TFP level and TFP growth

Let Y_t denote aggregate output (value added in fixed prices) at time t in a sector or the economy as a whole. Suppose Y_t is determined by the function

$$Y_t = F(K_t, H_t, t), \tag{5.1}$$

[1] I thank Niklas Brønager for useful discussions related to this chapter.

where K_t is an aggregate input of physical capital and H_t an index of quality-adjusted labor input.[2] The "quality-adjustment" of the input of labor (man-hours per year) aims at taking educational level and work experience into account. In fact, both output and the two inputs are aggregates of heterogeneous elements. The involved conceptual and measurement difficulties are huge and there are different opinions in the growth accounting literature about how to best deal with them. Here we ignore these problems. The third argument in (5.1) is time, t, indicating that the production function $F(\cdot\ ,\cdot\ ,t)$ is time-dependent. Thus "shifts in the production function", due to changes in efficiency and technology ("technical change" for short), can be taken into account. We treat time as continuous and assume that F is a neoclassical production function. When the partial derivative of F w.r.t. the third argument is positive, i.e., $\partial F/\partial t > 0$, technical change amounts to technical *progress*. We consider the economy from a purely supply-side perspective.[3]

We shall here concentrate on the fundamentals of TFP and TFP growth. These can in principle be described without taking the heterogeneity and changing quality of the labor input into account. Hence we shall from now on ignore this aspect and simplifying *assume* that labor is homogeneous and labor quality is constant. So (5.1) is reduced to the simpler case,

$$Y_t = F(K_t, L_t, t), \tag{5.2}$$

where L_t is the number of man-hours per year. As to measurement of K_t, some adaptation of the *perpetual inventory method*[4] is typically used, with some correction for under-estimated quality improvements of investment goods in national income accounting. The output measure is (or at least should be) corrected correspondingly, also for under-estimated quality improvements of consumption goods.

[2]Natural resources (land, oil wells, coal in the ground, etc.) constitute a third primary production factor. The role of this factor is in growth accounting often subsumed under K.

[3]Sometimes in growth accounting the left-hand side variable, Y, in (5.2) is the gross product rather than value added. Then non-durable intermediate inputs should be taken into account as a third production factor and enter as an additional argument of $\tilde{F}$ in (5.2). Since non-market production is difficult to measure, the government sector is usually excluded from Y in (5.2). Total Factor Productivity is by some authors called *Multifactor Productivity* and abbreviated MFP.

[4]Cf. Section 2.2 in Chapter 2.

5.2.1 TFP growth

The notion of Total Factor Productivity at time t, TFP_t, is intended to indicate a *level* of productivity. Nevertheless there is a tendency in the literature to evade a direct definition of this level and instead go straight away to a decomposition of output *growth*. Let us start the same way here but not forget to come back to the issue about what can be meant by the level of TFP.

The growth rate of a variable Z at time t will be denoted $g_{Z,t}$. We take the total derivative w.r.t. t in (5.2) to get

$$\dot{Y}_t = F_K(K_t, L_t, t)\dot{K}_t + F_L(K_t, L_t, t)\dot{L}_t + F_t(K_t, L_t, t) \cdot 1.$$

Dividing through by Y_t gives

$$\begin{aligned} g_{Y,t} &\equiv \frac{\dot{Y}_t}{Y_t} = \frac{1}{Y_t}\left[F_K(K_t, L_t, t)\dot{K}_t + F_L(K_t, L_t, t)\dot{L}_t + F_t(K_t, L_t, t) \cdot 1\right] \\ &= \frac{K_t F_K(K_t, L_t, t)}{Y_t} g_{K,t} + \frac{L_t F_L(K_t, L_t, t)}{Y_t} g_{L,t} + \frac{F_t(K_t, L_t, t)}{Y_t} \\ &\equiv \varepsilon_{K,t} g_{K,t} + \varepsilon_{L,t} g_{L,t} + \frac{F_t(K_t, L_t, t)}{Y_t}, \end{aligned} \tag{5.3}$$

where $\varepsilon_{K,t}$ and $\varepsilon_{L,t}$ are shorthands for $\varepsilon_K(K_t, L_t, t) \equiv \frac{K_t F_K(K_t,L_t,t)}{F(K_t,L_t,t)}$ and $\varepsilon_L(K_t, L_t, t) \equiv \frac{L_t F_L(K_t,L_t,t)}{F(K_t,L_t,t)}$, respectively, that is, the partial output elasticities w.r.t. the two production factors, evaluated at the factor combination (K_t, L_t) at time t. Finally, $F_t(K_t, L_t, t) \equiv \partial F/\partial t$, that is, the partial derivative w.r.t. the third argument of the function F, evaluated at the point (K_t, L_t, t).

The equation (5.3) is the *basic growth-accounting relation*, showing how the output growth rate can be decomposed into the "contribution" from growth in each of the inputs and a residual. The TFP *growth rate* is defined as the residual

$$g_{\mathrm{TFP},t} \equiv g_{Y,t} - (\varepsilon_{K,t} g_{K,t} + \varepsilon_{L,t} g_{L,t}) = \frac{F_t(K_t, L_t, t)}{Y_t}, \tag{5.4}$$

So the TFP growth rate is what is left when from the output growth rate is subtracted the "contribution" from growth in the factor inputs weighted by the output elasticities w.r.t. these inputs. This is sometimes interpreted as reflecting that part of the output growth rate which is *explained* by technical progress. One should be careful, however, not to identify a descriptive accounting relationship with deeper causality. Without a complete model, at most one can say that the TFP growth rate measures that fraction of output growth that is *not directly attributable* to growth in the capital and labor inputs. So:

> The TFP growth rate can be interpreted as reflecting the "*direct* contribution" to current output growth from current technical change (in a broad sense including learning by doing and organizational improvement).

Let us consider how the actual measurement of $g_{\text{TFP},t}$ can be carried out. The output elasticities w.r.t. capital and labor, $\varepsilon_{K,t}$ and $\varepsilon_{L,t}$, will, under perfect competition and absence of externalities and of increasing returns to scale, equal the income shares of capital and labor, respectively. Time series for these income shares and for Y, K, and L, hence also for $g_{Y,t}$, $g_{K,t}$, and $g_{L,t}$, can be obtained (directly or with some adaptation) from national income accounts. This allows straightforward measurement of the residual, $g_{\text{TFP},t}$.[5]

The decomposition in (5.4) was introduced already by Solow (1957). Since the TFP growth rate appears as a residual, it is sometimes called the *Solow residual*. As a residual it may reflect the contribution of many things, some wanted (current technical innovation in a broad sense including organizational improvement), others unwanted (such as varying capacity utilization, omitted inputs, measurement errors, and aggregation bias).

5.2.2 The TFP level

Now let us consider the *level* of TFP, that "something" for which we have calculated its growth rate without yet having defined what it really is. But knowing the growth rate of TFP for all t in a certain time interval, we in fact have a differential equation in the TFP level of the form $dx(t)/dt = g(t)x(t)$, namely:

$$d(\text{TFP}_t)/dt = g_{\text{TFP},t} \cdot \text{TFP}_t.$$

The solution of this simple linear differential equation is[6]

$$\text{TFP}_t = \text{TFP}_0 e^{\int_0^t g_{\text{TFP},\tau} d\tau}. \tag{5.5}$$

For a given initial value $\text{TFP}_0 > 0$ (which may be normalized to 1 if desired), the time path of TFP is determined by the right-hand side of (5.5). Consequently:

> The TFP level at time t can interpreted as reflecting the cumulative "*direct* contribution" to output since time 0 from cumulative technical change since time 0.

---

[5] Of course, data are in discrete time. So to make actual calculations we have to translate (5.4) into discrete time. The weights $\varepsilon_{K,t}$ and $\varepsilon_{L,t}$ can then be estimated by two-years moving averages of the factor income shares as shown in Acemoglu (2009, p. 79).

[6] See Appendix B of Chapter 3 in these lecture notes or Appendix B to Acemoglu.

Why do we say "*direct* contribution"? The reason is that the cumulative technical change since time 0 may also have an *indirect* effect on output, namely via affecting the output elasticities w.r.t. capital and labor, $\varepsilon_{K,t}$ and $\varepsilon_{L,t}$. Through this channel cumulative technical change affects the role of input growth for output growth. This possible indirect effect over time of technical change is not included in the TFP concept.

To clarify the matter we will compare the TFP calculation under Hicks-neutral technical change with that under other forms of technical change.

5.3 The case of Hicks-neutrality*

In the case of Hicks neutrality, by definition, technical change can be represented by the evolution of a one-dimensional variable, B_t, and the production function in (5.2) can be specified as

$$Y_t = F(K_t, L_t, t) = B_t F(K_t, L_t). \tag{5.6}$$

Here the TFP level is at any time, t, identical to the level of B_t if we normalize the initial values of both B and TFP to be the same, i.e., $\mathrm{TFP}_0 = B_0 > 0$. Indeed, calculating the TFP growth rate, (5.4), on the basis of (5.6) gives

$$g_{\mathrm{TFP},t} = \frac{F_t(K_t, L_t, t)}{Y_t} = \frac{\dot{B}_t F(K_t, L_t)}{B_t F(K_t, L_t)} = \frac{\dot{B}_t}{B_t} \equiv g_{B,t}, \tag{5.7}$$

where the second equality comes from the fact that K_t and L_t are kept fixed when the *partial* derivative of F w.r.t. t is calculated. The formula (5.5) now gives

$$\mathrm{TFP}_t = B_0 \cdot e^{\int_0^t g_{B,\tau} d\tau} = B_t.$$

The nice feature of Hicks neutrality is thus that we can write

$$\mathrm{TFP}_t = \frac{F(K_t, L_t, t)}{F(K_t, L_t, 0)} = \frac{B_t F(K_t, L_t)}{B_0 F(K_t, L_t)} = B_t, \tag{5.8}$$

using the normalization $B_0 = 1$. That is:

> *Under Hicks neutrality, current TFP appears as the ratio between the current output level and the hypothetical output level that would have resulted from the current inputs of capital and labor in case of no technical change since time 0.*

So in the case of Hicks neutrality the economic meaning of the TFP level is straightforward. The reason is that under Hicks neutrality the output elasticities w.r.t. capital and labor, $\varepsilon_{K,t}$ and $\varepsilon_{L,t}$, are *independent* of technical change.

5.4 The case of absence of Hicks-neutrality*

The above very intuitive interpretation of TFP is only valid under Hicks-neutral technical change. Neither under general technical change nor even under Harrod- or Solow-neutral technical change (unless the production function is Cobb-Douglas so that both Harrod and Solow neutrality imply Hicks-neutrality), will current TFP appear as the ratio between the current output level and the hypothetical output level that would have resulted from the current inputs of capital and labor in case of no technical change since time 0.

To see this, let us return to the general time-dependent production function in (5.2). Let X_t denote the ratio between the current output level at time t and the hypothetical output level, $F(K_t, L_t, 0)$, that would have obtained with the current inputs of capital and labor in case of no change in the technology since time 0, i.e.,

$$X_t \equiv \frac{F(K_t, L_t, t)}{F(K_t, L_t, 0)}. \tag{5.9}$$

So X_t can be seen as a factor of joint-productivity growth from time 0 to time t evaluated at the time-t input combination.

If this X_t should always indicate the level of TFP at time t, the growth rate of X_t should equal the growth rate of TFP. Generally, it does not, however. Indeed, defining $G(K_t, L_t) \equiv F(K_t, L_t, 0)$, by the rule for the time derivative of fractions[7], we have

$$\begin{aligned}
g_{X,t} &\equiv \frac{d\tilde{F}(K_t, L_t, t)/dt}{F(K_t, L_t, t)} - \frac{dG(K_t, L_t)/dt}{G(K_t, L_t)} \\
&= \frac{1}{Y_t}\left[F_K(K_t, L_t, t)\dot{K}_t + F_L(K_t, L_t, t)\dot{L}_t + F_t(K_t, L_t, t)\cdot 1\right] \\
&\quad -\frac{1}{G(K_t, L_t)}\left[G_K(K_t, L_t)\dot{K}_t + G_L(K_t, L_t)\dot{L}_t\right] \\
&= \varepsilon_K(K_t, L_t, t)g_{K,t} + \varepsilon_L(K_t, L_t, t)g_{L,t} + \frac{F_t(K_t, L_t, t)}{Y_t} \\
&\quad -(\varepsilon_K(K_t, L_t, 0)g_{K,t} + \varepsilon_L(K_t, L_t, 0)g_{L,t}) \qquad (5.10) \\
&= \left(\varepsilon_K(K_t, L_t, t) - \varepsilon_K(K_t, L_t, 0)\right)g_{K,t} + (\varepsilon_L(K_t, L_t, t) - \varepsilon_L(K_t, L_t, 0))g_{L,t} + g_{\mathrm{TFP},t} \\
&\neq g_{\mathrm{TFP},t} \qquad \text{generally,}
\end{aligned}$$

where $g_{\mathrm{TFP},t}$ is given in (5.4). Unless the partial output elasticities w.r.t. capital and labor, respectively, are unaffected by technical change, the conclusion is that TFP_t will differ from our X_t defined in (5.9). So:

[7] See Appendix A to Chapter 3 of these lecture notes.

In the absence of Hicks neutrality, current TFP does not generally appear as the ratio between the current output level and the hypothetical output level that would have resulted from the current inputs of capital and labor in case of no technical change since time 0.

A closer look at X_t vs. $\mathbf{TFP}_t$

As X_t in (5.9) is the time-t output arising from the time-t inputs relative to the fictional time-0 output from the same inputs, we consider X_t along with TFP as two alternative joint-productivity indices. From (5.10) we see that

$$g_{\mathrm{TFP},t} = g_{X,t} - (\varepsilon_K(K_t, L_t, t) - \varepsilon_K(K_t, L_t, 0))\, g_{K,t} - (\varepsilon_L(K_t, L_t, t) - \varepsilon_L(K_t, L_t, 0)) g_{L,t}.$$

So the growth rate of TFP equals the growth rate of the joint-productivity index X corrected for the cumulative impact of technical change since time 0 on the direct contribution to time-t output growth from time-t input growth. This impact comes about when the output elasticities w.r.t. capital and labor, respectively, are affected by technical change, that is, when $\varepsilon_K(K_t, L_t, t) \neq \varepsilon_K(K_t, L_t, 0)$ and/or $\varepsilon_L(K_t, L_t, t) \neq \varepsilon_L(K_t, L_t, 0)$.

Under Hicks-neutral technical change there will be no correction because the output elasticities are *independent* of technical change. In this case TFP coincides with the index X. In the absence of Hicks-neutrality the two indices differ. This is why we in Section 2.2 characterized the TFP level as the cumulative "*direct* contribution" to output since time 0 from cumulative technical change, thus excluding the possible indirect contribution coming about via the potential effect of technical change on the output elasticities w.r.t. capital and labor and thereby on the contribution to output from input growth.

Given that the joint-productivity index X is the more intuitive joint-productivity measure, why is TFP the more popular measure? There are at least two reasons for this. First, it can be shown that the TFP measure has more convenient balanced growth properties. Second, X is more difficult to measure. To see this we substitute (5.3) into (5.10) to get

$$g_{X,t} = g_{Y,t} - (\varepsilon_K(K_t, L_t, 0) g_{K,t} + \varepsilon_L(K_t, L_t, 0) g_{L,t}). \tag{5.11}$$

The relevant output elasticities, $\varepsilon_K(K_t, L_t, 0) \equiv \frac{K_t F_K(K_t, L_t, 0)}{F(K_t, L_t, 0)}$ and $\varepsilon_L(K_t, L_t, 0) \equiv \frac{L_t F_L(K_t, L_t, 0)}{F(K_t, L_t, 0)}$, are hypothetical constructs, referring to the technology as it was at time 0, but with the factor combination observed at time t, not at time 0. The nice thing about the Solow residual is that under the assumptions of perfect competition and absence of externalities, it allows measurement by using data on prices and quantities alone, that is, without knowledge of the production function. To evaluate g_X, however, we need estimates of the hypothetical output elasticities, $\varepsilon_K(K_t, L_t, 0)$ and $\varepsilon_L(K_t, L_t, 0)$. This requires knowledge about how the output elasticities depend on the factor combination and time, respectively, that is, knowledge about the production function.

Now to the warnings concerning application of the TFP measure.

5.5 Three warnings

Balanced growth at the aggregate level, hence Harrod neutrality, seems to characterize the growth experience of the UK and US over at least a century (Kongsamut et al., 2001; Attfield and Temple, 2010). At the same time the aggregate elasticity of factor substitution is generally estimated to be significantly less than one (see, e.g., Antras, 2004). This amounts to rejection of the Cobb-Douglas specification of the aggregate production function and so, at the aggregate level, Harrod neutrality rules out Hicks neutrality.

Warning 1 Since Hicks-neutrality is empirically doubtful at the aggregate level, TFP_t can often *not* be identified with the simple intuitive joint-productivity measure X_t, defined in (5.9) above.

Warning 2 When Harrod neutrality obtains, relative TFP growth rates across sectors or countries can be quite deceptive.

Suppose there are n countries and that country i has the aggregate production function

$$Y_{it} = F^{(i)}(K_{it}, A_t L_{it}) \qquad i = 1, 2, ..., n,$$

where $F^{(i)}$ is a neoclassical production function with CRS and A_t is the level of labor-augmenting technology which, for simplicity, we assume shared by all the countries (these are open and "close" to each other). So technical progress is Harrod-neutral. Let the growth rate of A be a constant $g > 0$. Many models imply that $\tilde{k}_i \equiv K_{it}/(A_t L_{it})$ tends to a constant, $\tilde{k}_i^*$, in the long run, which we assume is also the case here. Then, for $t \to \infty$, $k_{it} \equiv K_{it}/L_{it} \equiv \tilde{k}_{it} A_t$ where $\tilde{k}_{it} \to \tilde{k}_i^*$ and $y_{it} \equiv Y_{it}/L_{it} \equiv \tilde{y}_{it} A_t$ where $\tilde{y}_{it} \to \tilde{y}_i^* = f^{(i)}(\tilde{k}_i^*)$; here $f^{(i)}$ is the production function on intensive form. So in the long run g_{k_i} and g_{y_i} tend to $g_A = g$.

Formula (5.4) then gives the TFP growth rate of country i in the long run as

$$\begin{aligned} g_{\mathrm{TFP}_i} &= g_{Y_i} - (\alpha_i^* g_{K_i} + (1-\alpha_i^*) g_{L_i}) = g_{Y_i} - g_{L_i} - \alpha_i^*(g_{K_i} - g_{L_i}) \\ &= g_{y_i} - \alpha_i^* g_{k_i} = (1-\alpha_i^*) g, \end{aligned} \tag{5.12}$$

where α_i^* is the output elasticity w.r.t. capital, $f^{(i)\prime}(\tilde{k}_i)\tilde{k}_i/f^{(i)}(\tilde{k}_i)$, evaluated at $\tilde{k}_i = \tilde{k}_i^*$. Under labor-augmenting technical progress, the TFP growth rate thus varies negatively with the output elasticity w.r.t. capital (the capital income share under perfect competition). Owing to differences in product and industry composition, the countries have different α_i^*'s. In view of (5.12), for two different countries, i and j, we get

$$\frac{TFP_i}{TFP_j} \rightarrow \begin{cases} \infty \text{ if } \alpha_i^* < \alpha_j^*, \\ 1 \text{ if } \alpha_i^* = \alpha_j^*, \\ 0 \text{ if } \alpha_i^* > \alpha_j^*, \end{cases} \tag{5.13}$$

for $t \rightarrow \infty$.[8] Thus, in spite of long-run growth in the essential variable, y, being the same across the countries, their TFP growth rates are very different. Countries with low α_i^*'s appear to be technologically very dynamic and countries with high α_i^*'s appear to be lagging behind. It is all due to the difference in α across countries; a higher α just means that a larger fraction of $g_{y_i} = g_{k_i} = g$ becomes "explained" by g_{k_i} in the growth accounting (5.12), leaving a smaller residual. And the level of α has nothing to do with technical progress.

We conclude that comparison of TFP levels across countries or time may misrepresent the intuitive meaning of productivity and technical progress when output elasticities w.r.t. capital differ and technical progress is Harrod-neutral (even if technical progress were at the same time Hicks-neutral as is the case with a Cobb-Douglas specification). It may be more reasonable to just compare levels of Y/L across countries and time.

Warning 3 Growth accounting is - as the name says - just about accounting and measurement. So do not confuse growth *accounting* with *causality* in growth analysis. To talk about causality we need a theoretical model supported by the data. On the basis of such a model we can say that this or that set of exogenous factors through the propagation mechanisms of the model cause this or that phenomenon, including economic growth. In contrast, considering the growth accounting identity (5.3) in itself, none of the terms have

[8] If F is Cobb-Douglas with output elasticity w.r.t. capital equal to α_i, the result in (5.12) can be derived more directly by first defining $B_t = A_t^{1-\alpha_i}$, then writing the production function in the Hicks-neutral form (5.6), and finally use (5.7).

priority over the others w.r.t. a causal role. And there are important omitted variables. There are simple illustrations in Exercises III.1 and III.2.

In a complete model with exogenous technical progress, part of $g_{K,t}$ will be *induced* by this technical progress. If technical progress is endogenous through learning by investing, as in Arrow (1962), there is mutual causation between $g_{K,t}$ and technical progress. Yet another kind of model might explain both technical progress and capital accumulation through R&D, cf. the survey by Barro (1999).

5.6 References

Antràs, P., 2004, Is the U.S. aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution, *Contributions to Macroeconomics*, vol. 4, no. 1, 1-34.

Attfield, C., and J.R.W. temple, 2010, Balanced growth and the great ratios: New evidence for the US and UK, *J. of Macroeconomics*, vol. 32, 937-956.

Barro, R.J., 1999, Notes on growth accounting, *J. of Economic Growth*, vol. 4 (2), 119-137.

Bernard, A. B., and C. I. Jones, 1996a, Technology and Convergence, *Economic Journal*, vol. 106, 1037-1044.

Bernard, A. B., and C. I. Jones, 1996b, Comparing Apples to Oranges: productivity convergence and measurement across industries and countries, *American Economic Review,* vol. 86, no. 5, 1216-1238.

Greenwood, J., and P. Krusell, 2006, Growth accounting with investment-specific technological progress: A discussion of two approaches, *J. of Monetary Economics.*

Hercowitz, Z., 1998, The 'embodiment' controversy: A review essay, *J. of Monetary Economics,* vol. 41, 217-224.

Hulten, C.R., 2001, Total factor productivity. A short biography. In: Hulten, C.R., E.R. Dean, and M. Harper (eds.), *New Developments in Productivity Analysis,* Chicago: University of Chicago Press, 2001, 1-47.

Kongsamut, P., S. Rebelo, and D. Xie, 2001, Beyond Balanced Growth, *Review of Economic Studies*, vol. 68, 869-882.

Sakellaris, P., and D.J. Wilson, 2004, Quantifying embodied technological progress, *Review of Economic Dynamics*, vol. 7, 1-26.

Solow, R.M., 1957, Technical change and the aggregate production function, *Review of Economics and Statistics,* vol. 39, 312-20.