Chapter 3

Growth and development: theoretical roots

3.1 High development theory and the Big Push

THE term high development emerged in the 1940s and 1950s and is often used to describe a set of ideas on the importance of increasing returns and pecuniary external economies for economic development.[1] The key argument is that individual rates of return differ from social rates of return, because individuals ignore positive externalities. A *laissez-faire* market would therefore deliver too little investment, and public intervention would be required, up to the point of industrial programming. The stage for the high development theory was set by a number of seminal contributions, such as Paul Rosenstein-Rodan (1943), Ragnar Nurkse (1952), and Albert O. Hirschman (1958). Paul Rosenstein-Rodan (1943, pp 205-206) gives a nice illustration of the dilemma faced by poor and closed economies:

> "If a hundred workers who were previously in disguised unemployment [...] are put into a shoe factory, their wages will constitute additional income. If the newly employed workers spend all of their additional income on the shoes they produce, the shoe factory will succeed. In fact, however, they will not spend all of their additional income on shoes. There is no easy solution of creating an additional market in this way. The risk of not finding a market reduces the incentive to invest, and the

[1] The term was used for instance by Paul Krugman and Joseph Stiglitz during the 1992 Annual Conference on Development Economics of the World Bank.

> shoe factory investment project will probably be abandoned. Let us vary the example. Instead of putting a hundred previously unemployed workers in one shoe factory, let us put ten thousand workers in hundred factories and farms which between them will produce the bulk of wage-goods on which the newly employed workers spend their wages. What was not true in the case of one single shoe factory will become true for the complimentary system of hundred factories and farms. The new producers will be each other's customers".

In this story, the pecuniary externalities in the form of additional demand created by employment in the modern sector constitute the wedge between social and private returns to investment. If private returns are too low, firms may be reluctant to invest in the first place. In this case, only a coordinated effort to locate many firms at the same time would create the network necessary to render each individual firm potentially profitable. Since private business seems unable to coordinate such an effort, in particular in the constrained capital market of the 1950s, this could be the government's job. The economy may need a big push from the outside to get started, just like an airplane needs a critical ground speed before it can be airborne.

The basic appeal of the big-push story is that it offers an explanation for the twin-peak phenomenon described in section (2.5)—ie, the fact that some countries not only fail to catch up, but fall even further behind. If poor countries could overcome the vicious circle of poverty, low demand, and low private returns to investment, they might start catching up, as some managed with admirable success. However, market mechanisms may not suffice breaking this circle. The model presented in chapter (6) uses a similar approach to creating a vicious circle between institutions and capital accumulation.

The high development theory included many ideas, which reappeared later with the new growth and new trade theory. Moreover, the big-push idea has been rejuvenated and refined in more recent contributions, such as that of Murphy, Shleifer, and Vishny (1989), which added the formal rigour that was missing in the original papers. Paul Krugman (1992) offers a streamlined version of the model by Murphy et al (1989), which illustrates the basic spirit.

Consider a closed economy with two sectors, a modern and traditional sector. Labour is the only input factor. The traditional sector has constant returns, while the modern sector requires high fixed investments upfront, and hence, has increasing returns. Labour in the modern sector receives a premium expressed as a ratio $w > 1$ between wages in the modern and traditional sector.[2] The economy produces N goods, where units are chosen as to set productivity equal to one in each good. Unit labour requirements are decreasing in the modern sector because of increasing returns. Q_i is the production of good i in the modern sector. If all goods are produced in the modern sector, labour requirements for each are

$$L_i = F + c\, Q_i, \tag{3.1}$$

with c as marginal labour requirement, and F as the fixed upfront investment. For simplicity, F and c are equal for all goods. Demand for the goods is assumed to be symmetric—ie, each goods receives a share $1/N$ of expenditure. Each good produced in the traditional sector is supplied perfectly elastic at the marginal cost of production, which is unity in terms of traditional sector labour. In the modern sector, each good is produced by a single firm. The price the firm may charge is constrained by potential competition from the traditional sector, so that it may not charge more than unity itself, although it enjoys a monopoly in producing this good in the modern sector.

Figure (3.1) shows the production potential of the economy in question and illustrates the obstacle of overcoming the initial threshold. If all labour input would be used in the modern sector, an amount of Q_2 for each good could be produced, while only Q_1 of each good could be produced in the traditional sector. If

$$\frac{(L/N) - F}{c} > \frac{L}{N}, \tag{3.2}$$

then $Q_2 > Q_1$—ie, fixed investments may not be too high and the marginal costs advantage must be sufficiently large. However, even if the economy could potentially produce much more with modern techniques, the first-mover who establishes a modern firm in an otherwise traditional economy may face a substantial disadvantage because the produced good may

[2] This assumption is found frequently in high development theory, such as Rosenstein-Rodan (1943, p 204). The traditional sector, such as a family farm, is regarded as the employer of last resort, where the marginal product is very low (Lewis, 1954). Moreover, workers in the modern sector may need a richer diet to master the additional workload, and therefore need higher wages. During the empirical assessment of multiple equilibria, Messrs Graham and Temple (2001, pp. 4-8) include a similar wage differential, which *inter alia* corresponds to empirical observations and helps to stabilise the low-income equilibrium (see section 2.5).

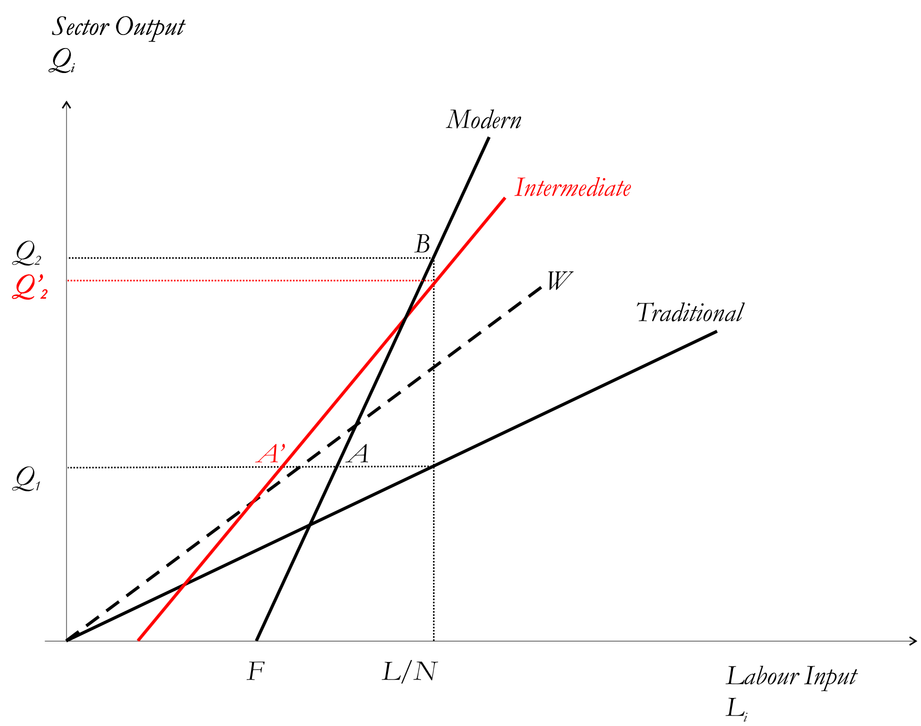

Figure 3.1: The Big Push; taken from Krugman (1992, p 18); Intermediate added.

meet too little demand to be profitable compared to the traditional sector. If only one firm switches from traditional to modern, the amount of goods it produces is by and large the same, Q_1. It pays a wage premium which creates additional demand, but if N is a big number, as assumed, this increase is negligible. Hence, the first-moving firm can only produce and sell Q_1. However, it can do so with less labour than before; let L_1^p denote labour requirements of the first firm entering the modern sector.

$$L_1^p = F + c \left(\frac{L}{N}\right) \tag{3.3}$$

From (3.2) follows that $L_1^p < L/N$. If the wage ratio is such that

$$w\ L_1^p > \frac{L}{N}, \tag{3.4}$$

then the costs of production in the modern sector are higher than in the traditional sector; therefore, no firm would enter the modern sector alone. The savings in employment through the superior technology of the modern sector are more than offset by the higher wages that must be paid. This is drawn in figure (3.1), where A indicates production of a single forerunner firm in the modern sector. It can only sell Q_1, employs less labour, but has higher costs than the traditional sector, because the dashed line, indicating the wage ratio, passes above point A. If all firms would produce in the modern sector, illustrated by point B, production for each good would increase to Q_2, and it would become profitable as the dashed line passes below.

Market mechanisms were insufficient to shift the economy out of traditional production. External economies, here in the form of pecuniary demand spillovers, and economies of scale in the modern sector because of the fixed investment necessary, as well as a traditional sector which pays lower wages, give rise to multiple equilibria: Some countries proceed to economic take-off, while other do not.[3]

If multiple equilibria are considered responsible for the lack of development, then the early authors of high development theory drew an apparently logical conclusion: industrial programming. Ragnar Nurkse (1953, p 13) calls for "[a] frontal attack—a wave of capital investments in a number of different industries". A strategy of balanced growth should encompass a broad range of industries which would create a market for each others products. As a response, but in the same spirit, Albert O. Hirschman (1958) emphasised

[3] Krugman is the first to note that some of the assumptions are hardly convincing. The purpose of this exercise has been to show how the main ideas of the high development theory could be translated into a simple model (Krugman, 1992, p 19).

the importance of strategic industries, which would trigger development by forward and backward linkages; a strategy of unbalanced growth. The experience since then has been mixed. Outward oriented strategies, such as export promotion in South-East Asia, are by and large considered a success, whereas inward-oriented strategies, such as import substitution have delivered disappointing results when widely applied in Latin America and in part of Asia (Krueger, 1980). By the same token, sections (2.3) and (6.4.2) call for some caution against over-enthusiastic investment programming.

The Big-Push story is not without caveats, in particular if it is based on demand spillovers. An intial problem is firm size. Because of the fixed investments but lower marginal labour requirements, a factory must have a sufficient size in order to undercut the costs of production in the traditional sector. However, the notion of two distinct economic sectors—a modern and a traditional one—is simplistic. Entrepreneurs face more choices than simply whether to produce, say, shoes either in the backyard of the family farm, or with a super-efficient, highly automated facility. It seems more plausible assuming that there are many different production designs, and that what is referred to as traditional and modern are just extreme versions of a continuum of different production technologies. Using a technology between traditional and modern would mean that less fixed investments, F, are needed but therefore higher marginal labour requirements, c, as compared to the modern sector. Such an intermediate technology is included in figure (3.1). As drawn, point A'—ie, the output of a single firm using intermediate technology, while all others stay in the traditional sector—is above the dashed line and therefore profitable. Moreover, when all firm proceed to intermediate, output is shifted to Q_2'. At this point, firms would find it profitable to switch to the modern sector and Q_2 would be realised. For the persistence of the development trap, condition (3.2) must therefore hold for all possible combinations of F and c, which are given by the available production technologies. Multiple equilibria persist only if no intermediate technology exists which violates (3.2). This is far more restrictive than the original setting.

A second point is closed economies. If the domestic market is too small to allow modern firms to gain a profitable size, why not export what cannot be sold at home? In fact, if market size was really the constraining factor, how could small economies with minuscule domestic demand ever become economically successful? Some small open economies, such as Hong Kong, Singapore, or Estonia, are remarkable success stories. Once trade linkages are introduced, market size ceases to be the main constraining factor, and firms can embrace modern technologies irrespective of domestic demand (cf Stiglitz, 1992).

Even if the Big-Push argument is not without its flaws, it is important not to throw out the baby with the bathwater. Features, such as increasing returns and externalities, are widely applied in new growth and new trade models. Multiple equilibria are a powerful instrument for explaining divergent economic development. Chapter (6) introduces a model which uses the idea of a big push, caused by temporarily increasing returns, in an open economy setting.

3.2 Growth and theory

The convergence controversy, as laid out in section (2.2), has led to substantial revisions and amendments of neoclassical growth theory. The seminal papers by Robert Solow (1956) and Trevor Swan (1956) show that in a closed economy with perfect competition, capital accumulation only leads to transitory per capita growth. Once the steady state is assumed, fresh capital only replaces depreciations or offsets increases in the workforce. Hence, the capital endowment per worker remains constant and so does per capita output. While there is no endogenous growth, improvements in technology may further increase productivity and therefore incomes per worker. However, technology is treated as a pure public good which develops exogenously (Solow, 1956; Swan, 1956; Barro and Sala-i-Martin, 1995, pp 17-26). Although these models have never been intended to be used as development models, they are widely applied as such. A quick back-of-the-envelope calculation reveals some of the problems.

Output, Y, is produced by combining capital, K, with labour, L, using a level of technology, A:

$$Y = F(A, K, L) \tag{3.5}$$

A Cobb-Douglas specification will be used as an explicit form of the production function because it satisfies neoclassical requirements, such as the Inada-condition, and homogeneity of degree one:

$$Y = AK^{\alpha}L^{1-\alpha} \tag{3.6}$$

The exponents α and $1-\alpha$ are the elasticities of factor inputs for capital and labour. Under the neoclassical assumption of perfect competition, α

equals the share of total income paid to capital, while $1-\alpha$ represents the compensation to labour. The entire national product is allocated to holders of capital and labour.

Output per worker or capita[4] is a better indicator of welfare, hence, equation (3.6) is divided by L,

$$y = Ak^{\alpha}, \tag{3.7}$$

where $y = \frac{Y}{L}$ denotes output per worker, and $k = \frac{K}{L}$ denotes capital endowment per worker, or capital density. Let $\hat{y}$ be the exponential rate of output growth, with $\hat{y} = \frac{\partial \ln y}{\partial t}$, and $\hat{k}$ and $\hat{A}$ the growth rates of capital and technology respectively:

$$\hat{y} = \alpha\hat{k} + \hat{A} \tag{3.8}$$

Moreover, in a closed economy, the growth of the capital stock per worker, $\dot{k}$ is determined by the domestic savings rate, s, and the growth of the labour force, n:[5]

$$\dot{k} = sy - nk \tag{3.9}$$

$$\hat{k} = s\frac{y}{k} - n \tag{3.10}$$

$$\text{from (3.7) follows: } k = A^{-\frac{1}{\alpha}} y^{\frac{1}{\alpha}} \tag{3.11}$$

$$\text{substituting in (3.10): } \hat{k} = sA^{\frac{1}{\alpha}} y^{-\frac{1}{\alpha}} - n \tag{3.12}$$

$$\text{substituting in (3.8): } \hat{y} = \alpha(sA^{\frac{1}{\alpha}} y^{\frac{\alpha-1}{\alpha}} - n) + \hat{A} \tag{3.13}$$

The key parameter in the analysis is the exponent α. Because α represents the share of total income passed to capital, its value can be taken from national accounts. By and large, a benchmark of 0.4 for α seems reasonable for a large sample of countries (Romer, 1994, p 6).

In 1960, Switzerland was the richest country (today it is only forth) with a GDP per capita of $15,214. Since then it grew by 1.33 percent annually; approximately the same rate as the Congo, Jordan, Argentina, or the Philippines. However, the average Filipino had to live on $2,090—ie, in 1960 Swiss income exceeded the Philippine by factor 7. Because

[4]For simplicity, workforce and population will be used as synonyms during this text. Although the workforce is much smaller than total population, and participation rates vary across countries, the general quality of arguments does not seem to be affected.

[5]Depreciations additionally reduce the capital stock; however, they are hard to measure, and are therefore skipped here.

$y_{Philippines} = \frac{1}{7} y_{Switzerland}$, and y is raised to the power $\frac{\alpha-1}{\alpha}$ (equals 1.5 for $\alpha = 0.4$), equation (3.13) suggests that Swiss savings rates should have been around 18 times higher than Philippine savings rates $[(\frac{1}{7})^{1.5} \approx 18.52]$ to generate the same growth rates. In a closed economy, domestic savings can be proxied by the investment rate. Indeed, Swiss did save a lot more than Filipinos, but with a (gross)[6] investment rate twice as high as in the Philippines, the magnitudes of differences are far away from the factor 18 as predicted by equation (3.13) (figures taken from the Penn World Tables by Heston, Summers, and Aten, 2002).[7]

In March 2003, major newspapers reported that Don Johnson, the aging star of the Miami-Vice TV-series, was caught with documents worth about $8 billion while crossing the German-Swiss border.[8] What seems strange, apart from the weird picture as such, is that anybody still wants to invest in Switzerland. Again, following equation (3.7), the marginal product of an additional unit of capital should be 18 times higher in the Philippines than in Switzerland. Why was Mr Johnson not caught at the Philippine border?

In the post Bretton-Woods era it is reasonable to drop the assumption of closed economies. For instance, the Central and Eastern European countries (CEEC) which have joined the European Union in May 2004 have abolished capital controls in compliance with the acquis communautaire. Prior to this date most only had modest controls in place, in order to discourage short-term capital flows. All CEECs have experienced substantial capital inflows at magnitudes of around 3 to 8 percent of GDP. However, Lipschitz, Lane, and Mourmouras (2002) have calculated that a strict application of the Cobb-Douglas function would predict a median inflow of around 500 percent of GDP in the CEECs. That has not been observed and exceeds any reasonable projection.

Of course, these startling results hinge on two critical assumptions: (i) Perfect competition ensures that total income is allocated to capital and labour according to their marginal products; and (ii), technology is a pure public good, and the level of technology, A, is the same for all countries. In particular the latter assumption seems somewhat outdated. Not many people

[6] Actually, net investments are what equation (3.13) calls for. However, data on depreciations is notoriously unreliable, hence, gross investment rates suffice as a quick guess.

[7] Similar exercises can be found in Lucas (1990, pp 93-94) and Romer (1994, pp 5-6).

[8] At the moment, Mr Johnson sues the German Ministry of Finance for publishing his involvement in this deal. Actually, Mr Johnson appearance seems unrelated to the documents, and consequently no allegations have been raised against him.

would argue that Sub-Saharan countries apply the same production technology as, say, the United States. Moreover, standard theory does not say anything about the determinants of A, other than time, and yet technology is the only source for growth once economies reach their steady-state. Moses Abramovitz (1993, p 218) refers to A as a measure of ignorance, because it reflects all that standard theory cannot explain. Hence, a better understanding of what drives A might not only improve the prediction of future growth, it might also help solving this puzzle.

In his 1994 article, Paul Romer lists 5 basic facts, which should be taken into account in the formulation of growth models and production functions.

1. There are many firms in a market economy.
2. Technological discoveries are non-rival.
3. Physical activities may be replicated.
4. Technological advancement comes from things people do.
5. Many individuals and firms have market power and earn monopoly rents on discoveries.

Neoclassical models, such as those by Messrs Solow and Swan, respect number 1-3 but not 4 nor 5 (Romer, 1994, p 13). This problem is addressed by endogenous-growth models, which include fact 4 and sometimes 4+5. Two general strands can be distinguished: (i) Spillover models, which follow Kenneth Arrow (1962); and (ii) Neo-Schumpeterian models, in the tradition of, yes, Joseph Schumpeter. While spillover models succeed in reconciling the data with a growth model and perfect competition, Neo-Schumpeter models allow for monopolies in the form of exclusive access to technology, which in turn induce deliberate investments in research and development.

A unifying feature of spillover models is that private returns differ from social returns because private actions include positive externalities; hence the name. The seminal approach by Kenneth Arrow (1962) regards technological advancement as a byproduct of current production. People learn while they do something. Inventions are not the result of deliberate effort, but happen somewhat accidently in the course of production. The more that is produced, the higher the likelihood of progress. To emphasise the technological content of A, the 'Learning-by-Doing' model relates technological advance only to the application of physical capital in production, hence the more capital-intensive the production the stronger the progress. Once

Table 3.1: **Production functions and spillovers**

Rober Solow (1956)	$Y = A(t) \cdot F(K, L)$
Kenneth Arrow (1962)	$Y_i = A(K) \cdot F(K_i, L_i)$
Paul Romer (1986)	$Y = A(R) \cdot F(R_i, K_i, L)$
Robert Lucas (1988)	$Y_i = A(H) \cdot F(K_i, H_i)$
Paul Romer (1994)	$Y = A(K, L) \cdot F(K_i, L_i)$

These stylised versions are taken from Romer (1994), pp 12-13.

R is investment in R&D; H is human capital; subscripts denote investments of individual firms.

inventions are made, they are accessible to everybody in the economy—ie, nobody is excluded from using the new technology.

The renaissance of spillover models in the 1980s started with Paul Romer (1986). Although Romer's model resembles an earlier version by Hirofumi Uzawa (1964), the renaissance is largely attributed to Mr Romer's contribution. The main difference to traditional, neoclassical growth models is the explicit inclusion of fact 4—ie, technological advance is the result of deliberate investments in research and development. Though individual firms engage in R&D, the resulting technological progress spreads to all firms. Obviously, this is not very encouraging for individuals. Moreover, it violates fact 3 because the production function is not homogenous of degree one in K and L. In a later article, Mr Romer presents a stylised model more in line with the 'Learning-by-Doing' view. While each unit of capital increases the level of technology via knowledge spillovers, increases in labour-supply depress advances, because labour-saving progress become less attractive. For simplicity, let the increase and reduction of A be of same magnitudes—ie, $A(K, L) = K^{\gamma} Y^{-\gamma}$. If one adds γ to α in equations (3.13), y will be raised to the power $\frac{\alpha+\gamma-1}{\alpha+\gamma}$. Taking the example of Switzerland and the Philippines as above: If γ were 0.3 Swiss investments rates would have to be only a little more than twice as large as Philippine's, because $(\frac{1}{7})^{\frac{\alpha+\gamma-1}{\alpha+\gamma}} \approx 2.3$. As said above, this is pretty much the relation of observed investment rates. Hence, this little exercise would suffice to reconcile the data with the theory (Romer, 1994, p 8). How to substantiate a γ of 0.3 is another question.

Table (3.1) offers an overview of the production functions applied in the different spill-over growth-models together with the seminal contribution by

Robert Solow. To underline the exogenous character of technological advance in the Solow model, A depends purely on time, t—ie, with more time elapsed more progress can be expected. Subsequent models try to model A endogenously. The aggregate stock of capital, deliberate investments in research and development or human capital, and the stocks of physical capital (+) together with labour supply (-) are thought to be major determinants of the technology level.

Still, most spillover models operate with perfect competition—ie, they circumvent fact 5 on Mr Romer's list. As a matter of fact, privately achieved innovations award monopoly rights; at least for a limited period of time and within certain legal restrictions. Patents, copyrights, trademarks and alike are a powerful source of economic rents. According to the consultancy firm Interbrand the Coca-Cola brand alone is worth close to $70 billion, roughly equal the GDP of Slovakia, and certainly an amount to be taken into account. Moreover, the current row over Trade Related aspects of Intellectual Property Rights (TRIPS) during the WTO negotiations illustrates how much there is at stake (The Economist, 2002).

The appeal of spillover models to economic development is that they allow for positive (and negative) externalities of certain economic activities, which in turn may create thresholds and multiple equilibria—as used in the Big-Push models. Moreover, comparison between Switzerland and the Philippines shows that variations in the technology parameter are a powerful instrument to account for the differing economic performance. This parameter must not necessarily be restricted to technological progress in a narrow sense, but may also encompass variations in social infrastructure or institutions. Looking at the determinants of this parameter may reveal feedback loops which give rise to development or poverty traps. The next section looks at an endogenous growth model in more detail, and section (3.4) describes an extension to developing countries.

3.3 Endogenous Growth: The Romer model

If technological progress is the key engine of steady-state growth, then it is important to understand where technological progress comes from. A purely exogenous rate of innovations and ideas is not satisfactory for it deprives economic agents of the possibility to influence this rate. However, it is impossible to change technological progress without knowing its determinants. This is a pity, given its enormous importance. Moreover, a lot of people and firms invest substantial amounts of time and money in order to develop new products and technologies. This must pay, otherwise one would expect firms to shut their R&D departments. Hence, it seems reasonable to draw a relation between the effort people put into research and its result: technological progress. This is not to deny that some of history's greatest inventions have been happened upon accidentally. However, the greater part of technological progress may be traced back to deliberate R&D.

A specific theory of endogenous growth was developed by Paul Romer (1986, 1990); it shows how developed countries—or the developed world as a whole—may enjoy sustained growth by investing into R&D.[9] The next section addresses the diffusion of new technologies to developing countries.

Romer's model economy has three sectors:

1. A final-goods sector: A large number of perfectly competitive firms combine labour and capital to produce a homogenous output good, Y.

2. A intermediate-goods sector: Firms purchase a design of a specific capital good from the research sector, which awards them a monopoly in producing this good. The good is then sold to the final-goods sector.

3. A research sector: Innovators produce new ideas, in the form of new capital goods, to which they hold eternal patent protection. They sell the design to the intermediate-goods sector.

The aggregate production function follows a Cobb-Douglas function where capital, K, and labour, L_y, are employed to produce output, Y. The level of *Harrod-neutral* technology is given by A:

$$Y = K^{\alpha}(A\ L_Y)^{1-\alpha} \tag{3.14}$$

The input factor labour is split between the final-goods sector, where output Y is produced, hence L_Y; and the research sector, where technological progress A is invented, hence L_A: $L = L_Y + L_A$.

Technological advance comes in the form of new ideas, which are invented in the research sector. New ideas appear if people try to discover them:

$$\dot{A} = \bar{\delta} L_A \tag{3.15}$$

[9] The exposition follows Jones (1995, 2002, pp 96-122).

Equation (3.15) says that changes in the stock of knowledge, $\dot{A}$, equal the amount of labour put into research, L_A, multiplied by the rate at which new discoveries are made, $\bar{\delta}$. This rate is determined by

$$\bar{\delta} = \delta \, A^{\phi}, \tag{3.16}$$

where δ and ϕ are constant parameters. The parameter ϕ indicates the amount of spillovers from previous ideas: with $\phi > 0$, the rate of discoveries would increase in the stock of ideas, perhaps because scientific advance also increases the productivity of research. With $\phi < 0$, the rate would decline in the A, for instance because the most obvious ideas have already been discovered. If both effects are of same size: $\phi = 0$. Furthermore, the number of people in research might negatively affect their average productivity, because the risk of duplication increases. The could be catched by raising labour input to the power of some parameter λ $(0 < \lambda < 1)$. Consequently, ideas are produced according to:

$$\dot{A} = \delta L_A^{\lambda} A^{\phi} \tag{3.17}$$

In steady-state, output, capital, and technology grow at the same and constant rate. Technology grows at:

$$\frac{\dot{A}}{A} = \delta \frac{L_A^{\lambda}}{A^{1-\phi}} \tag{3.18}$$

Technology grows at a constant rate if the numerator and denominator on the right-hand side of (3.18) grow at the same rate. Moreover, the supply of labour employed in research grows at the rate of population growth: $\dot{L}_A/L_A = n$. Hence, technology growth is given by:

$$\hat{A} = \frac{\dot{A}}{A} = \frac{\lambda n}{1-\phi} \tag{3.19}$$

In the steady state, the economy grows at a rate which is given by the parameters that determine the productivity of research and population growth. Growth increases if congestion effects in research are low ($\lambda \uparrow$) or spillovers from the stock of knowledge are strong ($\phi \uparrow$). Moreover, population growth has a positive effect on economic growth, because the burden of producing new ideas can be shared among more researchers.

There are two externalities involved, which may be ignored by individual researchers. Small, individual researchers may neglect their contribution to the stock of ideas and the corresponding spillover; as well as the higher risk of duplication once they enter research, which is akin to congestion on a highway. In fact, individual researchers take $\bar{\delta}$ as given and see constant returns to research by ignoring the externalities. Private and social returns may fall apart.

The final-goods sector produces output by combining labour with a range of capital goods, x_j, which are bought from the intermediate sector. The range is given by the level of technology, A:

$$Y = L_Y^{1-\alpha} \int_0^A x_j \mathrm{d}j \tag{3.20}$$

For a given level of technology, equation (3.20) has constant returns to scale in labour and each capital good. At constant inputs, output increases in A. Firms in the final-goods sector maximise profits according to

$$\max L_Y^{1-\alpha} \int_0^A x_j \mathrm{d}j - w\ L_Y - \int_0^A p_j\ x_j \mathrm{d}j, \tag{3.21}$$

where the first term represents output (with price normalised to unity), the second the costs of labour, and the third term represents the cost of the capital goods. Wages are indicated by w, and the price of a capital good j is given by p_j. First-order conditions are:

$$w = (1-\alpha)\frac{Y}{L_Y} \tag{3.22}$$

$$p_j = \alpha\ L_Y^{1-\alpha} x_j^{\alpha-1} \tag{3.23}$$

Firms hire labour until its marginal product equals the wage, and they buy capital goods until the marginal product equals the price.

The intermediate sector produces capital goods by transforming one unit of raw capital into one unit of the capital good, although they must pay for the design of the capital good. They maximise profits π_j according to

$$\max \pi_j = p_j(x_j)\, x_j - r\, x_j, \tag{3.24}$$

where $p_j(x_j)$ is the demand function for the capital good given by (3.23), and r is the interest rate. The first-order condition with respect to x, dropping the subscripts, is:

$$p = \frac{1}{1 + \frac{p'(x)x}{p}} \tag{3.25}$$

From (3.23) one can derive:

$$\frac{p'(x)x}{p} = \alpha - 1 \tag{3.26}$$

Hence,

$$p = \frac{1}{\alpha} r \tag{3.27}$$

Monopolists charge a markup over marginal costs. The solution (3.27) is for each monopolist, so that all capital goods costs the same and are employed in equal amounts by the final-goods sector—ie, $x_j = x$. Therefore, the capital stock in the economy consists of A capital goods employed in equal amounts:

$$K = Ax \tag{3.28}$$

Combining equations (3.23), (3.27), and (3.28) allows calculating the interest rate:

$$\begin{aligned} \frac{r}{\alpha} &= \alpha\ L_Y^{1-\alpha} x^{\alpha-1} \\ r &= \alpha\ L_Y^{1-\alpha} \left(\frac{K}{A}\right)^{\alpha-1} \end{aligned}$$

Hence,

$$r = \alpha^2 \frac{Y}{K} \tag{3.29}$$

Equation (3.29) highlights an important result of the Romer model: The interest rate in the economy is lower than the marginal product of capital, $\frac{\partial Y}{\partial K} = \alpha \frac{Y}{K}$.

The difference is used to compensate researchers, and prevails only because firms in the intermediate-goods sector enjoy a monopoly in the provision of specific capital goods, which they buy from researchers. The price firms charge from the final-goods sector exceeds marginal costs, which is necessary because average costs would otherwise be higher than marginal costs and the intermediate-goods sector would collapse. Why? Firms in the intermediate-goods sector pay a fixed sum to a researcher in exchange for the right to produce a specific capital good. Marginal costs are constant, because they transform one unit of raw capital into one unit of capital good, as given by (3.24). Therefore, the intermediate-goods sector has increasing returns because average costs are declining in the output of capital goods. The monopoly rights firms buy, ensure that imperfect competition allows the charging of higher prices than marginal costs.

Profits in the intermediate-goods sector are given by:

$$\pi = \alpha(1-\alpha)\frac{Y}{A} \tag{3.30}$$

Firms in the intermediate-goods sector pay a price for a new idea, P_A, which equals its net present value relative to any other investment:

$$rP_A = \pi + \dot{P}_A \tag{3.31}$$

Equation (3.31) states that the profits made with a new idea, π, plus any change in the price of that idea, $\dot{P}_A$, must equal the amount that could otherwise be earned with the money, rP_A. In steady-state, output, Y, and capital, K, grow at the same rate; therefore the interest rate must be constant (see 3.29). Hence, π/P_A must be constant also, and π and P_A grow at the rate of population growth, n.[10]

$$P_A = \frac{\pi}{r-n} \tag{3.32}$$

The wage researchers earn must equal the wage in the final-goods sector because there are no entry barriers in either sector which would limit arbitrage. The wage in the research sector is given by

$$w_R = \bar{\delta} P_A, \tag{3.33}$$

which must equal the wage given by (3.22):

$$\bar{\delta} P_A = (1-\alpha)\frac{Y}{L_Y} \tag{3.34}$$

Substituting (3.32), (3.30), and (3.15) into (3.34), plus a little algebra returns

---

[10] In steady state, per capita output, y, and the level of technology grow at the same rate. Therefore Y/A grows at the rate of population growth, n. Since π is proportional to Y/A, it must grow at the rate n, too.

$$s_R = \frac{L_A}{L} = \frac{1}{1 + \frac{r-n}{\alpha \hat{A}}}, \tag{3.35}$$

with s_R as the fraction of total labour employed in the research sector.

There are two reasons why current s_R may be different from an optimum fraction. Knowledge spillovers and research congestion put a wedge between the private and social returns of research. Moreover, firms in the intermediate-goods sector buy research according to the monopoly profit they receive but neglect the concomitant consumer-surplus. With regard to basic science, in particular, the distortions are deemed so large that most research is done through publicly funded programmes, because private firms have little incentives to do it themselves.

3.4 Growth and development

The previous section introduces a growth model, based on the work by Paul Romer, which illustrates the origins of technological progress. Once countries have reached their steady-state, their subsequent growth only occurs as the result of technological advances. However, the Romer model says little about why some countries are rich and others poor. The level of technology—ie, the range of capital goods—is not restricted to the developed world, but accessible by all countries in principle. Poor countries, though, may lack the skills to take full advantage of the most sophisticated technologies. These skills are expressed by the human capital of the workforce, education in particular. Hence, human capital is what limits the application of new technologies: Only those ideas may be applied for which sufficiently skilled workers exist in the economy:

$$Y = L^{1-\alpha} \int_0^h x_j^\alpha \mathrm{d}j \tag{3.36}$$

Comparing equation (3.36) with (3.20) reveals that the developing economy may only produce those capital goods for which workers with a corresponding level of skill can be found. An individual's skill level is denoted by h, which is the range of capital goods a worker is trained enough to produce. If h is smaller than A then the skill level is not high enough to produce all capital goods that are available (This production function may be found also in Easterly, King, Levine, and Rebelo, 1994).[11]

[11] The exposition follows Jones (2002, pp 124-130).

As in the previous section, one unit of capital good may be produced with one unit of raw capital: $\int_0^h x_j \mathrm{d}j = K$. Moreover, all capital goods are employed at equal amount, thus $x_j = x$ for all j. The production function may be simplified to:

$$Y = K^{\alpha}(h\ L)^{1-\alpha} \tag{3.37}$$

Skills are accumulated according to:

$$\dot{h} = \mu e^{\psi u} A^{\gamma} h^{1-\gamma} \tag{3.38}$$

Skill accumulation depends on two factors. First, the time spent on education and schooling, denoted with u; and second, the distance between the skill level in the economy and the technology frontier.[12]

The parameter ψ is the proportional increase in the skill level for an increase in u. As a rule of thumb, wages earned by an individual increase by around 10 percent for each additional year of schooling. This could be expressed by $\psi = 0.1$, if u were regarded as years of schooling (Jones, 2002, p 56).[13]

The term $A^{\gamma}h^{1-\gamma}$ indicates that workers with a skill level far behind the technology frontier learn faster than those close to the frontier. "This implies ... that it took much longer to learn to use computers thirty years ago, when they were new, than it does today." (Jones, 2002, p 127). In fact, this is similar to the idea of β-convergence, which suggests that countries grow faster the further away they are from their steady-state.

The technological frontier is challenged primarily by reserach conducted in the developed world—the contribution of developing countries is negligible. Hence, the growth rate of the level of technology, $\hat{A}$, is exogenously given for developing countries. Poor countries are kept busy learning how to use the existing stock of knowledge but contribute few, if any, new ideas.

In equation (3.37), the skill level, h, is in the production function in the form of *Harrod-neutral* technology. Thus, in steady state, output per worker y, and capital per worker, k, will grow at the same rate as h. Moreover, from equation (3.38) it is clear that A/h must be constant, because it is proportional to $\hat{h}$. Hence, A and h must grow at the same rate:

$$\hat{y} = \hat{k} = \hat{h} = \hat{A} \tag{3.39}$$

[12] The parameters are assumed to be $\mu > 0$ and $0 < \gamma \leq 1$.

[13] Robert Lucas (1988) introduces a similar function for human capital, $H = e^{\psi u}L$, where human capital, H, depends on schooling time.

Rewriting equation (3.38):

$$h = \left(\frac{\mu}{\hat{h}} e^{\psi u}\right)^{\frac{1}{\gamma}} A \tag{3.40}$$

Steady-state output per worker, y^* is given by:

$$y^* = \left(\frac{s}{n + \hat{A} + d}\right)^{\frac{\alpha}{1-\alpha}} h \tag{3.41}$$

Substituting (3.40) and knowing that $\hat{h} = \hat{A}$:

$$y^* = \left(\frac{s}{n + \hat{A} + d}\right)^{\frac{\alpha}{1-\alpha}} \left(\frac{\mu}{\hat{A}} e^{\psi u}\right)^{\frac{1}{\gamma}} A \tag{3.42}$$

Remember that $\hat{A}$ is determined by research efforts mainly in developed countries and is taken as given for poor countries. The first term on the right-hand side of (3.42) is familiar from the basic growth models of Robert Solow and Trevor Swan.[14] The second term details the effect of schooling and eduction on the development of y. It shows that more schooling increases the steady-state output. The technology frontier, A, enters as a third term. Even poor countries with little human capital profit from new ideas if only because they increase the speed of learning.

The intuition behind this argument is plausible. Poor countries usually employ less sophisticated technologies, and usually have lower enrollment and literacy rates—ie, less schooling and education. Moreover, empirical evidence finds that human capital, including health, is positively associated with growth (see table 2.2).

The question is whether if human capital is the best or only variable for explaining technology diffusion. And the answer is: probably not. Another obvious explanation why poor countries employ simpler technology is that they produce according to their comparative advantage—ie, lower wages. Since labour is cheap, firms in developing countries substitute fancy technology with manpower, which would be presumably more capital intensive. Apparently, the argument is somewhat circular, because low wages may reflect little schooling. However, causality is an important point. The model suggests that poor human capital limits the diffusion of new technology because workers with sufficient skills are required, which causes a lower steady-state income than otherwise. The reverse may hold, too: Because

[14] Again, s is the savings rate, and d the depreciation rate.

countries are poor, wages are low. Therefore, it pays to employ simple technology and more manpower, and the demand for capital goods from the technology frontier may be low. This holds in particular for innovations which increase labour-productivity (labour-augmenting or *Harrod-neutral* technological advance): such innovations are more welcome, and thus rewarded higher, if labour is expensive. According to this logic, poor countries should welcome capital-augmenting innovations (*Solow-neutral* innovation), because they would increase the effective supply of the scarce input factor.[15]

Consider for example mobile phones, a technology which was quickly adopted by many poor countries. According to the International Telecommunications Union, an organisation within the United Nations System, Africa is the fastest growing market for mobile phones. By 2003, mobile phone penetration in Africa has hit 6.2 percent, while only 3 percent used fixed telephones. More than 50m Africans now own a mobile subscription; by 2010 this figure is expected to reach between 100 and 200 million (International Telecommunications Union, 2004a). Africa is not an exception. Mobile phones are popular throughout the developing world, in particular where there are few fixed-telephone lines (International Telecommunications Union, 2004b).

The appeal of mobile phones (which were considered posh when introduced in the early 90s) to the developing world is that they allow the installation of a cable-based telephone network to be skipped. This did not matter in rich countries where telephone connection has been around for decades, even in the most remote areas. Not so in many poor countries, where mobile phones allowed the connection of previously unserved remote areas, at a much lower price. Mobile phones are an excellent example of capital-augmenting technologies, suggesting that these are readily employed by poor countries. Low human capital has not prevented firms from installing the necessary infrastructure, and selling phones, contracts, and other services to customers. Where workers lacked the required skills, they trained them.

The distinction in labour-augmenting and capital-augmenting technology, where the latter is more warmly embraced by developing countries than the former, casts doubt on the direction of causality suggested by this section's model. Instead, it seems that the diffusion of new ideas is driven by the demand for advanced technology, which in turn depends *inter alia* on the wage level. Consequently, it may be futile trying to increase steady-state

[15] Models of economic growth usually use only labour-augmenting innovations, because they are compatible with a long-term equilibrium, while capital-augmenting technological advance is not (Barro and Sala-i-Martin, 1995, pp 63-64).

output by deliberately increasing schooling and education (see also section 2.3).

Another good explanation why technology is not adopted by poor countries is a lack of infrastructure and institutional quality. Organisational innovations in particular are only valuable in an environment where complex business structures are possible. For instance, lean production and delivery-on-time, which contributed a lot to cost reductions in OECD countries, is only possible if reliable suppliers are at hand. Punctual deliveries are hampered by bad roads, an erratic bureaucracy, and difficulties in claiming compensation for bad or late deliveries. All of these factors are abundant in most developing countries.

3.5 Technology to institutions

Endogenous growth theory offers valuable insights into the determinants of technological advancement. This is important because technological progress is the only engine of growth once economies have reached their steady state. The application of the Romer model to developing countries suggests that human capital plays a crucial role in determining the speed of technology diffusion. However, this explanation has not proved to be entirely convincing. Labour-augmenting innovations increase the effective supply of an input factor which is already relatively abundant in poor countries; hence, it is little surprising that these technologies are adopted more slowly. Capital-augmenting progress, on the other hand, increases the effective supply of a relatively scarce input factor, and is much more warmly embraced in the developing world. The penetration of mobile phones in Africa serves as an example. Moreover, poor institutions may limit the degree to which advanced technologies are employed, because simple techniques may be less sensitive to the quality of institutions. For instance, software developers may shy away from markets where there is little respect for intellectual property rights.

It is worth remembering that the basic growth model by Messrs Solow and Swan used A to capture variations in growth which could not be explained by variations in inputs. The parameter A is more a residuum, which includes whatever affects the productivity of capital and labour. As some say, calling this parameter technology, with its growth rate given exogenously, was a great marketing trick, which paved the way for the overwhelming success of the model.[16] As mentioned above, variables other than technological progress may increase productivity; institutions being high on that list. Moreover, bad institutions may explain in the first place why countries have lower investment rates and less schooling, which then translate into lower steady-state output per worker (Jones, 2002, pp 136-149).

It is tempting to regard institutions as technology of societies, and therefore as something very similar to technological innovations. An efficient way to organise a society, with well designed rules, may be just like a new way to assemble cars. But it is not. The key difference between technology and institutions is that institutions are rules, which need an ongoing effort of enforcement in order to be effective. Technology is accumulated over time, when new ideas are added to the stock of knowledge. It is cumulative in the sense that once innovations are made, there is not much effort required to keep them. They are written in a book, and accessible to all who can buy and read it. Institutions, on the other hand, consists of rules and enforcement (see section 4.1): unenforced rules are worthless, because nobody complies. Enforcement is costly, and once societies stop investing into enforcement, institutions deteriorate.

In fact, the stock of knowledge deteriorates, too, if it is not cared for. Imagine that universities, schools, and libraries would be closed; would anybody still know to construct a nuclear power plant in say 50 years time? History provides numerous examples of societies that actually lost knowledge. Printing, for instance, was first developed in Greece around 1700 BC, according to archeological evidence found in Phaistos on Crete. But apparently, printing proved impractical and was forgotten until it was reinvented some 2,500 years later in China and 3,100 years later in Europe (Diamond, 1997, pp 287-289). However, the rate with which unused ideas depreciate is arguably much lower than the rate with which unenforced institutions collapse.

It is therefore reasonable to include the quality of institutions as an additional parameter, because technology and human capital capture its influence only incompletely. Jones (2002, p 147), for instance, proposes a production function, that includes human capital, h, as well as a parameter

[16] This comment is from John Nye at a conference on new institutional economics in Corsica 2004. Work by Moses Abramovitz (1993) points in a similar direction. It highlights the view that the meaning of the parameter A should not be narrowed too much.

I which denotes social infrastructure:

$$Y = IK^{\alpha}(hL)^{1-\alpha} \tag{3.43}$$

The concept of social infrastructure has a strong overlap with institutions. The main challenge ahead is to derive an idea of the determinants of I. Why is it that some countries have better institutions than others?