2 The Schumpeterian Approach

The earliest attempt at providing a Schumpeterian approach to endogenous growth theory was that of Segerstrom, Anant, and Dinopoulos (1990), who modeled sustained growth as arising from a succession of product improvements in a fixed number of sectors, but with no uncertainty in the innovation process.

In this chapter we will sketch a simple model (see Aghion and Howitt 1988, 1992a) where growth is generated by a random sequence of quality improving (or "vertical") innovations that themselves result from (uncertain) research activities.

This model of growth with vertical innovations has the natural property that new inventions make old technologies or products *obsolete*. This obsolescence (or "creative destruction") feature in turn has both *positive* and *normative* consequences. On the *positive* side, it implies a negative relationship between current and future research, which results in the existence of a unique steady-state (or balanced growth) equilibrium and also in the possibility of cyclical growth patterns. On the *normative* side, although current innovations have positive externalities for future research and development, they also exert a negative externality on incumbent producers. This *business-stealing effect* in turn introduces the possibility that growth be *excessive* under laissez-faire, a possibility that did not arise in the endogenous growth models surveyed in the previous chapter.

2.1 A Basic Setup

The basic model abstracts from capital accumulation completely.[1] The economy is populated by a continuous mass L of individuals with linear intertemporal preferences: $u(y) = \int_0^\infty y_\tau e^{-r\tau} d\tau$, where r is the rate of time preference, also equal to the interest rate. Each individual is endowed with one unit flow of labor, so L is also equal to the aggregate flow of labor supply. Output of the consumption good depends on the input of an intermediate good, x, according to[2]

$$y = Ax^\alpha, \tag{2.1}$$

1. The implications of introducing human and physical capital accumulation are explored in the next chapter.

2. Throughout much of this book we restrict attention to the simple Cobb-Douglas form (2.1), because it is simple and involves handy formulas that the reader will quickly get used to. However, almost all of the analysis could be conducted under the more general hypothesis

$$y = AF(x),$$

where the production function F has a positive and diminishing marginal product, and where monopolist's marginal revenue schedule: $A\left(F'(x) + xF''(x)\right)$ is decreasing in x.

where $0 < \alpha < 1$. Innovations consist of the invention of a new variety of intermediate good that replaces the old one, and whose use raises the technology parameter, A, by the constant factor, $\gamma > 1$.

Society's fixed stock of labor has two competing uses. It can produce intermediate goods, one for one, and it can be used in research. That is:

$$L = x + n, \qquad \text{(L)}$$

where x is the amount of labor used in manufacturing and n is the amount of labor used in research.

When the amount n is used in research, innovations arrive randomly with a Poisson arrival rate λn, where $\lambda > 0$ is a parameter indicating the productivity of the research technology. The firm that succeeds in innovating can monopolize the intermediate sector until replaced by the next innovator. As in the basic Romer model, there are positive spillovers from the activities that generate growth in A, in two senses. The monopoly rents that the innovator can capture are generally less than the consumer surplus created by the intermediate good, and, more important, the invention makes it possible for other researchers to begin working on the next innovation. However, there is a negative spillover in the form of a "business-stealing effect," whereby the successful monopolist destroys the surplus attributable to the previous generation of intermediate good by making it obsolete.

The research sector is portrayed as in the patent-race literature that has been surveyed by Tirole (1988) and Reinganum (1989). The amount of labor devoted to research is determined by the arbitrage condition

$$w_t = \lambda V_{t+1}, \qquad \text{(A)}$$

where t is not time but the number of innovations that have occurred so far, w_t the wage, and V_{t+1} the discounted expected payoff to the $(t+1)^{th}$ innovation. The left-hand side is the value of an hour in manufacturing, whereas the right-hand side is the expected value of an hour in research—the flow probability λ of an innovation times the value V_{t+1}.

This arbitrage equation governs the dynamics of the economy over its successive innovations. Together with the labor market equation (L), it constitutes the backbone of the basic Schumpeterian model.

The value V_{t+1} is determined by the following asset equation:

$$rV_{t+1} = \pi_{t+1} - \lambda n_{t+1} V_{t+1},$$

which says that the expected income generated by a license on the $(t+1)^{th}$ innovation during a unit time interval, namely rV_{t+1}, is equal to the profit flow π_{t+1} attainable by the $(t+1)$ intermediate good monopolist minus the expected "capital loss" that will occur when the $(t+1)^{th}$ innovator is replaced by a new innovator and therefore loses V_{t+1}. The flow probability of this loss is the arrival rate λn_{t+1}. Put in slightly different terms, the value V_{t+1} of the $(t+1)^{th}$ innovation is the net present value of an asset that yields π_{t+1} until it disappears, which it does at the expected rate λn_{t+1}.[3]

Box 2.1
Poisson Processes

Throughout these chapters we will often assume that some random event X is governed by a "Poisson process," with a certain "arrival rate" μ. What this means mathematically is that the time T you will have to wait for X to occur is a random variable whose distribution is exponential with parameter μ:

$$F(T) \equiv \text{Prob}\{\text{Event occurs before } T\} = 1 - e^{-\mu T}.$$

So the probability density of T is

$$f(T) = F'(T) = \mu e^{-\mu T}.$$

That is, the probability that the event will occur sometime within the short interval between T and $T + dt$ is approximately $\mu e^{-\mu T} dt$. In particular, the probability that it will occur within dt from *now* (when $T = 0$) is approximately μdt. In this sense μ is the probability per unit of time that the event will occur now, or the "flow probability" of the event.

For example, in the present chapter the event that an individual researcher discovers innovation number $t + 1$ is governed by a Poisson process with the arrival rate λ. The expression λV_{t+1} on the right-hand side of the arbitrage equation (A) represents the expected income of an individual researcher, because over a short interval of length dt the researcher will make an innovation worth V_{t+1} with probability λdt.

If X_1 and X_2 are two distinct events governed by independent Poisson processes with respective arrival rates μ_1 and μ_2, then the flow probability that at least one of the events will occur is just the sum of the two independent flow probabilities $\mu_1 + \mu_2$, because the probability that both events will occur at once is negligible. In this sense, independent Poisson processes are "additive." This is why, in the present chapter, when n_t independent researchers each innovate with a Poisson arrival rate λ, the Poisson arrival rate of innovations to the economy as a whole is the sum λn_t of the individual arrival rates.

If a sequence of independent events takes place, each governed by the same independent process with the constant arrival rate μ, then the expected number of arrivals per unit of time is obviously the arrival rate μ. For example, in the present chapter the expected number of innovations per year in a balanced growth equilibrium is the arrival rate λn.

Moreover, the number of events x that will take place over any interval of length Δ is distributed according to the "Poisson distribution" that you will find described in most statistics textbooks:

$$g(x) = \text{prob}\{x \text{ events occur}\} = \frac{(\mu\Delta)^x e^{\mu\Delta}}{x!},$$

whose expected value is the arrival rate times the length of the interval $\mu\Delta$. This distribution is used in section 2.3 to express the expected present value of future output in a balanced growth equilibrium.

3. n_{t+1} is the amount of labor devoted to R&D after the $(t+1)^{th}$ innovation.

Note that this equation presupposes that the incumbent innovator does not perform R&D, so that λn_{t+1} is indeed the probability of that innovator losing his or her monopoly rents. In fact, there is a simple reason why the incumbent innovator chooses to do no research; all the other researchers have immediate access to the incumbent technology A_{t+1} as a benchmark for their own research, and the value to the incumbent innovator of making the next innovation is $V_{t+2} - V_{t+1}$, which is strictly less than the value V_{t+2} to an outside researcher. This is an example of the "Arrow effect," or "replacement effect."

We thus have

$$V_{t+1} = \pi_{t+1}/(r + \lambda n_{t+1}). \tag{2.2}$$

The denominator of (2.2), which can be interpreted as the obsolescence-adjusted interest rate, shows the effects of creative destruction. The more research is expected to occur following the next innovation, the shorter the likely duration of the monopoly profits that will be enjoyed by the creator of the next innovation, and hence the smaller the payoff to innovating.

The model is now almost entirely specified, except for the profit flow π_t and also the flow demand for manufacturing labor x_t. Both are determined by the same profit-maximization problem solved by the intermediate producer that uses the t^{th} innovation. This producer could either be thought of as being the t^{th} innovator (who then sets up a new intermediate firm), or as an existing intermediate firm that purchases (at price V_t) the patent for the innovation from the t^{th} innovator. In either case, the t^{th} innovator is able to extract the whole expected net present value (NPV) of (monopoly) profits generated by that innovation during the lifetime of this innovation, namely V_t.

The t^{th} incumbent innovator will determine π_t and x_t by solving

$$\pi_t = \max_x [p_t(x)x - w_t x],$$

where w_t is the wage and $p_t(x)$ the price at which the t^{th} innovator (or intermediate firm) can sell the flow x of intermediate input to the final good sector. We assume the final good sector to be competitive, so that $p_t(x)$ must equal the marginal product of the intermediate input x in producing the final (or consumption) good. Thus, from equation (2.1), $p_t(x) = A_t \alpha x^{\alpha-1}$ is the inverse demand curve facing the t^{th} innovator.

The first-order condition to the above maximization program yields immediately the following expressions for x_t and π_t:

$$x_t = \arg\max_x \{A_t \alpha x^{\alpha} - w_t x\}$$

$$= \left(\frac{\alpha^2}{w_t/A_t}\right)^{1/(1-\alpha)} \equiv \left(\frac{w_t}{A_t}\right).$$

and

$$\pi_t = \{A_t \alpha x_t^{\alpha} - w_t x_t\}$$

$$= \left(\frac{1}{\alpha} - 1\right) w_t x_t = A_t \widetilde{\pi}\left(\frac{w_t}{A_t}\right)$$

Before proceeding, note that x_t and π_t are both decreasing functions of the productivity-adjusted wage rate $\omega_t = \frac{w_t}{A_t}$. That π_t decreases with respect to ω_t in turn introduces an additional reason, besides creative destruction, for the negative dependency of current research on the amount of expected future research: specifically, a higher demand for future research labor will push future wage ω_{t+1} up, thereby decreasing the flow of profits π_{t+1} to be appropriated by the next innovator. This, in turn, will tend to discourage current research, that is, to drive n_t down.

The model is now fully characterized by both:

- the *arbitrage equation* (A), which reflects the fact that labor can be freely allocated between manufacturing and research, and which can now be reexpressed (after substituting for V_{t+1} and π_{t+1}, noting $A_{t+1} = \gamma A_t$, and dividing both sides of equation (A) by A_t) as

$$\omega_t = \lambda \frac{\gamma \widetilde{\pi}(\omega_{t+1})}{r + \lambda n_{t+1}}. \tag{A}$$

- the *labor market clearing equation* (L), which reflects the frictionless nature of the labor market and determines the growth-adjusted wage rate ω_t as a function of the residual supply of manufacturing labor $L - n_t$:

$$L = n_t + \widetilde{x}(\omega_t), \tag{L}$$

where the demand for manufacturing labor $x_t = \widetilde{x}(\omega_t)$ is a decreasing function of the growth-adjusted wage rate ω_t.

2.2 Steady-State Growth

2.2.1 Comparative Statics on the Steady-State Level of Research

A steady-state (or balanced growth) equilibrium is simply defined as a stationary solution to system (A) and (L), with $\omega_t \equiv \omega$ and $n_t \equiv n$. In other words, both the allocation of labor between research and manufacturing and the productivity-adjusted wage rate remain constant over time, so that wages, profit, and final output are all scaled up by the same $\gamma > 1$ each time a new innovation occurs.

In a steady state the *arbitrage* and *labor market clearing* equations simply become

$$\omega = \lambda \frac{\gamma \tilde{\pi}(\omega)}{r + \lambda n} \qquad (\widehat{A})$$

$$n + \tilde{x}(\omega) = L. \qquad (\widehat{L})$$

Because the two curves corresponding to $(\widehat{A})$ and $(\widehat{L})$ in the (n, ω) space are respectively downward and upward sloping, the steady-state equilibrium $(\widehat{n}, \widehat{\omega})$ is unique. Using figure 2.1, we easily see that the equilibrium level of research $\widehat{n}$ will be raised by a lower interest rate r, a higher size of the labor market L, a higher productivity of R&D λ, and a higher size of innovation γ. These comparative-statics results are all intuitive: (a) A decrease in the rate of interest increases the marginal benefit to research, by raising the present value of monopoly profits. (b) An increase in the size of each innovation also increases the marginal benefit to research, by raising the size of the next interval's monopoly profits relative to this interval's productivity. (c) An increase in the endowment of skilled labor both increases the marginal benefit and reduces the marginal cost of research, by reducing the wage of skilled labor. (d) An increase in the arrival parameter decreases both the marginal cost and the marginal benefit of research, because on one hand it results in more "effective" units of research for any given level of employment, but on the other hand it also increases the rate of creative destruction during the next interval. The former effect turns out to dominate.

Also, using the fact that in a steady state the productivity-adjusted profit flow $\tilde{\pi}$ is equal (from above) to

$$\tilde{\pi} = \frac{1-\alpha}{\alpha}\omega x = \frac{1-\alpha}{\alpha}\omega(L-n),$$

we see that $(\widehat{A})$ can be rewritten as

$$1 = \lambda \frac{\gamma \frac{1-\alpha}{\alpha}(L-n)}{r + \lambda n}, \qquad (2.3)$$

according to which the steady-state level of research $\widehat{n}$ is a *decreasing* function of α; that is, a decreasing function of the elasticity of the demand curve faced by the intermediate monopolist.

In other words, *product market competition is unambiguously bad for growth*: the more competition, the lower the size of monopoly rents that will be appropriated by successful innovators, and therefore the smaller the incentives to innovate. This unambiguous—but also somewhat simplistic—prediction of the basic Schumpeterian model will be discussed in detail in Chapter 7.

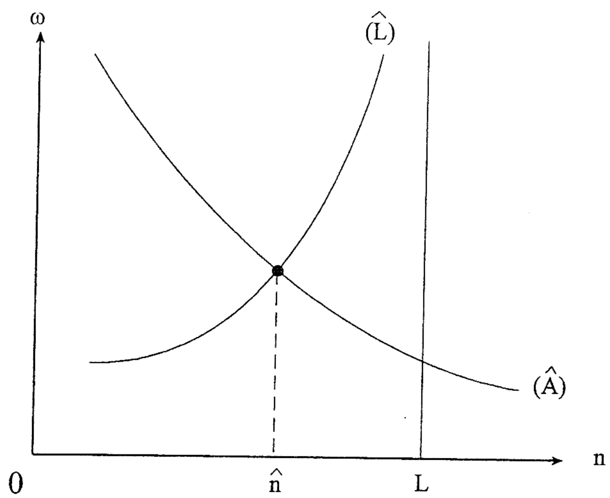

Figure 2.1

2.2.2 Comparative Statics on the Steady-State Rate of Growth

In a steady state the flow of consumption good (or final output) produced during the time-interval between the t^{th} and the $(t+1)^{th}$ innovation is

$$y_t = A_t \widehat{x}^{\alpha} = A_t (L - \widehat{n})^{\alpha}$$

which implies that

$$y_{t+1} = \gamma y_t \qquad (2.4)$$

Now, the reader should remember that the variable "t" does not refer to real time, but rather to the sequence of innovations $t = 1, 2, 3$, and so on. What happens to the evolution of final output in real time, that is, as a function of τ?

From equation (2.4) we know that the log of final output $\ln y(\tau)$ increases by an amount equal to $\ln \gamma$ each time a new innovation occurs. However, the real time interval between two successive innovations is random. Therefore, as shown in figure 2.2, the time path of the log of final output $\ln y(\tau)$ will itself be a random step function, with the size of each step being equal to $\ln \gamma > 0$ and with the time interval between each step being exponentially distributed with parameter $\lambda \widehat{n}$. Taking a unit-time interval between τ and $\tau + 1$, we have: $\ln y(\tau+1) = \ln y(\tau) + (\ln \gamma)\varepsilon(\tau)$, where $\varepsilon(\tau)$ is the number of innovations between τ and $\tau + 1$. Given that $\varepsilon(\tau)$ is distributed Poisson with parameter $\lambda \widehat{n}$, we have: $E(\ln y(\tau+1) - \ln y(\tau)) = \lambda \widehat{n} \ln \gamma$, where the LHS is nothing but the average growth rate.

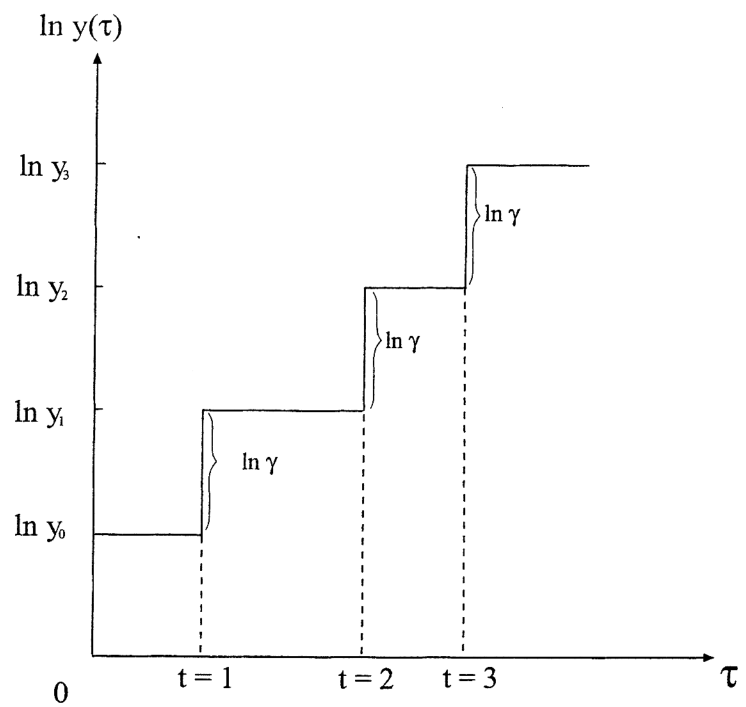

Figure 2.2

We thus end up with a very simple expression for the average growth rate in a steady state:

$$g = \lambda \widehat{n} \ln \gamma. \tag{G}$$

Combining this equation with the previous comparative-statics analysis on the steady-state level of research $\widehat{n}$, we are now able to sign the impact of parameter changes on the average growth rate. Increases in the size of the labor market L or a reduction of the interest rate r and in the degree of market competition α will increase $\widehat{n}$ and thereby also g. Increases in the size of innovation γ and/or in the productivity of R&D λ will also foster growth, directly (by increasing the factor λ ln γ) and also indirectly through increasing $\widehat{n}$.

Although the relationship between trade and growth is not our primary focus in this chapter,[4] these comparative-statics results suggest the following ambiguous effect of trade liberalization on long-run growth: on one hand, by increasing the size of the overall labor market pool, trade liberalization appears to be growth-enhancing; on the other hand, to the extent that it may also increase product market competition (or the possibility of imitating current innovations), trade liberalization may reduce the reward to new innovations and thereby discourage research and growth.

4. See Chapter 11.

2.3 Welfare Analysis

This section compares the laissez-faire average growth rate derived earlier with the average growth rate that would be chosen by a social planner whose objective was to maximize the expected present value of consumption $y(\tau)$. Because every innovation raises $y(\tau)$ by the same factor γ, the optimal policy consists of a fixed level of research. Expected welfare is

$$U = \int_0^\infty e^{-r\tau} y(\tau) d\tau = \int_0^\infty e^{-r\tau} \left(\sum_{t=0}^\infty \Pi(t,\tau) A_t x^\alpha \right) d\tau, \tag{2.5}$$

where $\Pi(t,\tau)$ equals the probability that there will be exactly t innovations up to time τ. Given that the innovation process is Poisson with parameter λn, we have

$$\Pi(t,\tau) = \frac{(\lambda n \tau)^t}{t!} e^{-\lambda n \tau}$$

The social planner will then choose (x, n) to maximize U subject to the labor resource constraint $L = x + n$. Using the fact that $A_t = A_0 \gamma^t$, we can reexpress the expected welfare U as

$$U(n) = \frac{A_0 (L-n)^\alpha}{r - \lambda n (\gamma - 1)}.$$

Then, the socially optimal level of research n^* will satisfy the first-order condition $U'(n^*) = 0$, which can be equivalently expressed as

$$1 = \frac{\lambda(\gamma - 1)\left(\frac{1}{\alpha}\right)(L - n^*)}{r - \lambda n^* (\gamma - 1)}. \tag{2.6}$$

This level of research would produce an average growth rate equal to

$$g^* = \lambda n^* \ln \gamma.$$

Whether the laissez-faire economy's average growth rate $g = \lambda \widehat{n} \ln \gamma$ is more or less than the optimal rate g^* will depend upon whether the steady-state equilibrium level of research $\widehat{n}$ determined earlier is greater or smaller than the socially optimal level n^*.

Now, in order to simplify the comparison between n^* and $\widehat{n}$, we can use equation (2.3), that is

$$1 = \frac{\lambda\gamma\left(\frac{1-\alpha}{\alpha}\right)(L-\widehat{n})}{r+\lambda\widehat{n}} \tag{2.7}$$

which determines $\widehat{n}$ just as (2.6) determines n^*. There are three differences between (2.6) and (2.7). The first is that the social discount rate $r - \lambda n(\gamma - 1)$ appears in (2.6) instead of the "private discount rate" $r + \lambda n$. The social rate is less than the rate of interest, whereas the private rate is greater. This difference corresponds to the *intertemporal spillover effect* discussed earlier. The social planner takes into account that the benefit to the next innovation will continue forever, whereas the private research firm attaches no weight to the benefits that accrue beyond the succeeding innovation.[5] This effect tends to generate insufficient research under laissez-faire.

The second difference is the factor $(1-\alpha)$, which appears on the right-hand side of (2.7) but not in (2.6). This is an *appropriability effect* that reflects the private monopolists' inability to appropriate the whole output flow (he or she can appropriate only a fraction $(1-\alpha)$ of that output). This effect also tends to generate too little research under laissez faire.

The third difference is that the factor $(\gamma - 1)$ in the numerator (2.6) replaces γ in the RHS of the arbitrage equation. This corresponds to a *"business-stealing" effect*. The private research firm does not internalize the loss to the previous monopolist caused by an innovation. In contrast, the social planner takes into account that an innovation destroys the social return from the previous innovation.[6] This effect will tend to generate *too much* research under laissez-faire.

Thus, both the intertemporal-spillover and appropriability effects tend to make the average growth rate less than optimal, whereas the business-stealing effect tends to make it greater. Because these effects conflict with each other, the laissez-faire average growth rate may be more or less than optimal.

5. Two additional spillovers could easily be included. First, researchers could benefit from the flow of others' research, so that an individual firm's arrival rate would be a constant-returns function of its own and others research. Second, there could be an exogenous Poisson arrival rate μ of imitations that costlessly circumvent the patent laws and clone the existing intermediate good. Both would have the effect of lowering the average growth rate relative to its optimal value.

6. Under more general production technologies than Cobb-Douglas, there is a further "monopoly distortion effect" that arises from the fact that private firms choose x to maximize profits, not final consumption (whereas the former is always equal to $(1-\alpha)$ times the latter in the Cobb-Douglas case).

The appropriability and intertemporal-spillover effects dominate when the size of innovations γ is large, in which case $\widehat{n} < n^*$. However, when there is much monopoly power (α close to zero) *and* innovations are not too large, the business-stealing effect dominates, in which case $\widehat{n} > n^*$. In that case, and unlike in the models analyzed in the previous chapter, *laissez-faire growth will be excessive*!

This new possibility is the main welfare implication of introducing obsolescence (or creative destruction) in the process of economic growth. The idea that laissez-faire growth can be "excessive" because of the negative externality that new innovators exert upon incumbent firms will come up again in the following chapters, in particular Chapter 4 on growth and unemployment.

2.4 Uneven Growth

Going back to the basic model of section 2.1, we have already hinted at a negative correlation between current and future research in equilibrium: a higher level of research n_{t+1} tomorrow will both imply more creative destruction $(r + \lambda n_{t+1} \uparrow)$ and less profit $(\pi_{t+1} = A_{t+1}\widetilde{\pi}(\omega_{t+1}) \downarrow)$ after the next innovation $(t+1)$ occurs. This in turn will unambiguously discourage current research, that is, n_t will decrease.

In other words, the two basic equations (A) and (L) boil down to a single negative relationship between n_t and n_{t+1}[7]

$$n_t = \psi(n_{t+1}), \quad \psi' < 0. \tag{2.8}$$

A perfect foresight equilibrium (PFE) is defined as a sequence $(n_t)_0^\infty$ satisfying (2.8) for all $t \geq 0$. In figure 2.3, the sequence $\{n_0, n_1, \ldots\}$ constructed from the clockwise spiral starting at n_0 constitutes a PFE. The steady-state $\widehat{n}$ analyzed in the previous sections is defined as the fixed point of the mapping ψ, or equivalently as the intersection between the ψ-curve and the 45^o-line. Other equilibria may also exist. A two-cycle is a pair (n^0, n^1) such that $n^0 = \psi(n^1)$ and $n^1 = \psi(n^0)$. It defines a PFE of period two. If both n^0 and n^1 are positive, the PFE is a "real" two-cycle. If either n^0 or n^1 is zero, it is a "no-growth trap." In a real two-cycle, the prospect of high research in odd

7. Dividing both sides of equation (A) by A_t, and using the fact that $n_{t+1} = L - \widetilde{x}(\omega_{t+1})$ (equation (L)), we obtain

$$\omega_t = \frac{\lambda\gamma\widetilde{\pi}(\omega_{t+1})}{r + \lambda(L - \widetilde{x}(\omega_{t+1}))} = \theta(\omega_{t+1}),$$

where $\theta' < 0$. Using again the fact that $n_t = L - \widetilde{x}(\omega_t)$ and therefore is increasing in ω_t for all t, we indeed get $n_t = \psi(n_{t+1})$ with $\psi' < 0$. □

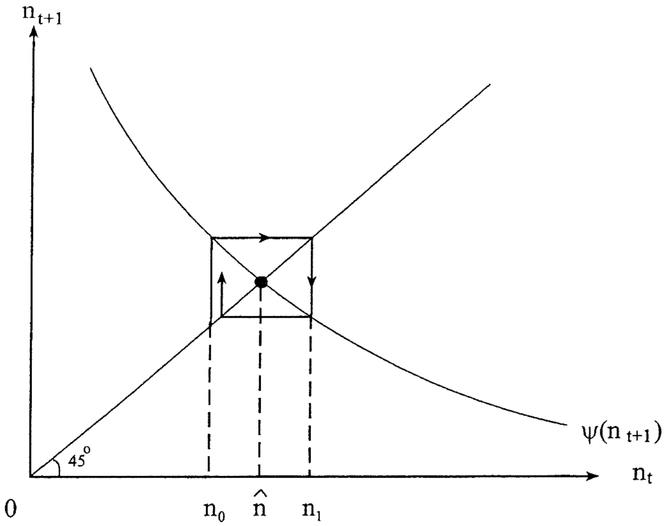

Figure 2.3

intervals discourages research in even intervals, and the prospect of low research in even intervals stimulates research in odd intervals. On such a PFE, growth will be cyclical. Although the analysis in this section does not claim to explain the relationship between growth and cycles (see rather Chapter 8), it nevertheless points at the possibility of deriving causal interactions between fluctuation and growth. As we will argue in Chapter 8, the two phenomena have traditionally been analyzed separately prior to endogenous growth theory.

Finally, one can show the existence of a no-growth trap equilibrium under suitable parameter conditions. A no-growth trap is the extreme case in which the prospect of high research in odd intervals shuts down research completely in even intervals.[8] Although the no-growth trap defines an infinite sequence $\{n_t\}_0^\infty$, the oscillation will cease after one innovation. From then on no growth will occur because the innovation process has stopped. Thus whether or not the economy grows at all can depend on the psychology of innovators. The expectation that the economy will grow at the steady-state rate $\lambda\hat{n}\ln\gamma > 0$ will be self-fulfilling, but so might the expectation that the economy will not grow at all!

8. It is straightforward to show (Aghion and Howitt 1992a) that a no-growth trap and a steady state with positive growth will both exist for small enough rates of interest, given all the other parameters of the model.

2.5 Discussion

The basic Schumpeterian model outlined in this chapter shares a number of limitations with previous endogenous growth models. A first limitation is its reliance on steady-state constructions. Both the Romer (1987, 1990a) model of horizontal innovations and the vertical (or quality ladders) model presented in this section make assumptions that ensure the existence of a steady state with balanced growth. These assumptions are quite severe and have nothing to recommend them except for tractability. We have already pointed out that in the Solow-Swan model with technological change one needs to assume Harrod-neutral (purely labor-augmenting) technical change at a constant exponential rate. In the Cass-Koopmans-Ramsey model one needs in addition a utility function in the one-parameter iso-elastic class. Analogous assumptions are needed even when technology is endogenized, and there is no good reason for thinking that they apply, even roughly.[9]

These strong assumptions rule out important phenomena, and answer important questions, by mere assertion. For example, they miss the stages of development in which resources are gradually reallocated from agriculture to manufacturing and then to services, all with different factor requirements and with different technological dynamics. The economy is always a scaled-up version of what it was years ago, and no matter how far it has developed already the prospects for future development are always a scaled-up version of what they were years ago. It seems just as likely that economies go through long phases of rises and decline or through periodic fluctuations, when introducing and/or diffusing new technological paradigms. (See Chapter 8, which discusses some first attempts at explaining aggregate output [and employment] fluctuations on the transition path between two technological paradigms.)

A second limitation lies in the description of *knowledge* as a parameter A, which fits into the aggregate production function much like any other factor of production. In reality, the growth of technological knowledge takes the form of new ideas or, in Schumpeter's terms, new combinations known to

9. More specifically, suppose that society's production possibilities at each date are given by the function

$$c = G(L, K, A; dk/dt, dA/dt), \qquad (2.9)$$

which indicates how much consumption can be produced by the resources L, K, and A, when some of those resources must also be devoted to making K and A grow at the specified rates. Then in order for a balanced growth path to exist, with capital, output, consumption, and the productivity parameter *A* all rising at the same constant exponential rate, the production function *G* must obviously be homogeneous of degree one in all arguments exept *L*, at all points on the balanced path. Furthermore, if this homogeneity property holds everywhere in a neighbourhood of the balanced path, then the path can be optimal only if the utility function is isoelastic. All models of balanced endogenous growth assume this homogeneity property and isoelastic utility.

produce useful output. Assuming that labor and capital can be aggregated, as in all the literature we have cited, strains credibility, but it is hard to know even what is meant by assuming that ideas can be aggregated. A more natural way to think about the growth of knowledge would be to use either the adaptive models of learning found in the macroeconomic literature on convergence (or nonconvergence) to rational-expectations equilibrium (for example, Frydman and Phelps 1983), or Bayesian models of learning by experimentation. Such models assume that people do not know the "true" parameter or function describing technology. Instead they must learn about them through some combination of experience, experimentation, intuitive guessing, and creative extrapolation. There is no guarantee in general that the corresponding learning processes converge quickly, or at all, to the truth, especially when the possibilities for controlled experimentation are severely limited, and when the behavior of the variables that people are trying to forecast is affected by the forecasts themselves. (See Chapters 6 and 8 for first attempts at introducing *learning* and *experimentation* as important components of the growth process.)

A third main limitation is the lack of attention to *institutions and transaction costs*. Evidence that the Solow residual accounts for most of long-run growth in per capita income is not compelling evidence of the primary importance of technological knowledge in the same growth. The Solow residual is just, as Abramowitz once put it, "a measure of our ignorance," and evidence that it "explains" a lot is just evidence that our ignorance is very large. Douglas North (1989) has argued forcibly that the growth of productivity that has occurred since the rise of the West is as much attributable to the development of institutions that have allowed us to reduce transaction costs, and thereby to exploit more fully the potential gains from exchange, as it is to our increased control over nature. Yet with few exceptions the endogenous-growth literature has focused on our increased ability to understand and master nature as the ultimate mainspring of growth.

A related shortcoming of all endogenous growth models is their representation of firms and R&D activities. Most of the existing literature models R&D as being performed by the same individual (or by indeterminate collections of individuals). In practice, however, research and development are not performed by the same individuals. Typically R&D takes place within firms where employee-inventors are subject to assignment contracts with their employers, the employer providing the financing and physical capital, and the employee providing skills and ideas. Contractual provisions on how to share property rights on inventions, on how to structure the monetary and nonmonetary compensations to the inventor, are far more complex than the simple representation in terms of individual patents. Going farther into understanding the *financial and institutional aspects* of R&D in situations where individual researchers are cash or credit-constrained should undoubtedly enrich the Schumpeterian approach to growth. (See Chapters 7, 13, and 14 for preliminary attempts in that direction.)

Moving away from the representative agent assumption would also allow these models to incorporate the *political dimension* of "creative destruction." The old paradigm makes economic growth appear as an unmixed blessing that raises everyone's welfare, thereby ignoring obsolescence and other resource-allocation aspects of growth that explain why there is always a vested interest opposed to the introduction of new technologies and new institutions. Until this distributional tussle is incorporated into the heart of endogenous growth theory (and again, the Schumpeterian model of creative destruction appears to be the most natural framework to use) it is hard to see how proponents of such models can claim any deep understanding of why some societies acquire and adopt new technologies and institutions more rapidly than others, or how they will ever understand the phenomenon that Olson (1982) calls the rise and decline of nations. Preliminary attempts at introducing distributional and political considerations into an endogenous growth framework will be analyzed and discussed in Chapters 9 and 10.

Although the following chapters will explore several important dimensions in which the Schumpeterian paradigm can be fruitfully applied or developed, the most immediate extensions of the basic model will now be addressed in the next and last section of this chapter.

2.6 Some Immediate Extensions of the Basic Schumpeterian Model

2.6.1 Technology Transfers and Cross-Country Convergence

The earlier model suggests that two independent economies should always *diverge* in log of GDP terms; indeed, to the extent that the log of GDP in each country follows a *random walk with drift*, the fact that country C has innovated more than country D (and therefore $A(C) > A(D)$), does not imply that the latter country is more likely to make the next innovation(s). Having a tendency to diverge even with the same parameters (λ, γ), two independent economies will a fortiori exhibit divergent development paths if $(\lambda_C, \gamma_C) \neq (\lambda_D, \gamma_D)$.

Although the basic Schumpeterian model appears to be strongly biased toward nonconvergence, a straightforward extension of that model can nevertheless account for β-convergence, that is, for the evidence that conditional on a given steady-state path, those countries that are currently farther below that path tend to grow faster. The following extension emphasizes knowledge spillovers (or technology transfers) across countries instead of decreasing returns to capital accumulation as the main source of β-convergence.

Consider an open economy, and suppose that the rest of the world grows at an average rate g. Thus, at any date τ, the average worldwide knowledge

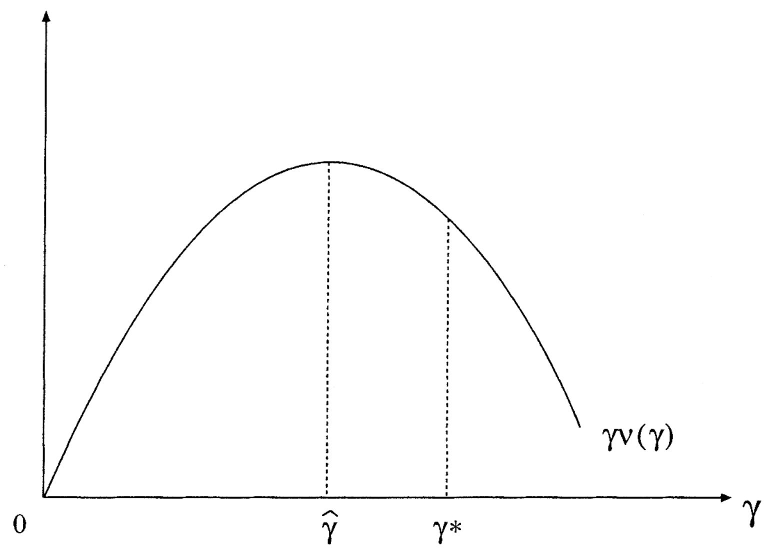

Figure 2.7

The economy's average growth rate (AGR) $\lambda\hat{n}v(\hat{\gamma})\ln\hat{\gamma}$ is affected by the fact that innovations are too small under laissez-faire, although the direction of the overall effect is ambiguous. The direct effect on $\ln\hat{\gamma}$ is to decrease AGR. The direct effect on the arrival rate working through $v(\hat{\gamma})$ is to increase AGR. The indirect effect on the arrival rate working through $\lambda\hat{n}$ is to decrease AGR.

In the nondrastic case, the above business-stealing effect whereby innovations are too small under laissez-faire is mitigated by an additional effect, namely that private innovators tend to increase the size of innovations in order to increase their profit margins. This margin is independent of the size γ in the drastic case, but it increases with γ in the nondrastic case. (Remember that the profit margin is $\alpha^{-1} - 1$ if the innovation is drastic and $\gamma^{1/\alpha} - 1$ if nondrastic.) However, this additional "profit-margin" effect does not necessarily overturn our earlier result to the effect that innovations are too small under laissez-faire. (See Aghion and Howitt 1992a).

2.7 Summary

This chapter has presented a simple model of growth through creative destruction. Output of the economy depends on how much intermediate input is employed and on the quality of that intermediate input. Succeeding vintages of intermediate goods embody quality improvements, which render their predecessors obsolete. These quality improvements also produce economic growth. They result from "research" activities by firms that generate a random sequence of product innovations. The uncertainty of the research process implies that growth will be stochastic.

Society has two uses for its fixed labor force: the manufacture of the latest generation of intermediate goods, and research aimed at discovering the next generation. Firms are motivated to hire labor for research by the prospect of monopoly rents. That is, if the research results in an innovation, the firm will get to monopolize the intermediate-goods industry until someone else comes along with a better product to replace it.

The model possesses a unique steady-state equilibrium in which society's division of labor between research and manufacturing remains unchanged over time, so that growth is stochastic but balanced. The average growth rate in steady-state equilibrium is an increasing function of the propensity to save, the productivity of the research technology that relates R&D employment to the expected arrival rate of innovations, and the degree of market power enjoyed by a successful innovator, all of which encourage more labor to be transferred from manufacturing to research.

The average growth rate may be either too low or too high to maximize welfare, because there are both positive and negative externalities in research, and it is not clear which will predominate. The positive externalities are the *intertemporal spillover* whereby the knowledge embedded in each innovation can be used by all future researchers, and the *appropriability effect* whereby the monopoly rents that motivate research firms constitute only part of the immediate social gain from an innovation, the rest being consumer surplus. The main negative externality is the *business-stealing effect* whereby a research firm does not internalize the loss to society from the obsolescence created by its innovation.

This chapter focuses mainly on steady-state balanced growth. However, there may also exist unbalanced equilibrium growth paths, in which the level of research switches with each innovation, between a high level and low level. When firms expect low research after the next innovation they are encouraged to do much research because the next successful innovator will retain its monopoly position for a long time. When they expect high research after the next innovation, however, they are discouraged by the prospect of rapid obsolescence. Not only does this make oscillatory equilibrium possible, it also creates the possibility of a *no-growth trap*; a situation in which people expect so much research to take place after the next innovation that no research at all is done, and the economy stagnates at the current level of output.

This chapter also shows how to incorporate "technology transfer" in the basic model. If two economies were completely unconnected, there would be no possibility of convergence; each would grow at a rate determined by its own research effort. But in reality, research in one country benefits from knowledge created in others; this provides a simple mechanism by which a laggard country would tend to catch up.

One can also allow contemporaneous technology spillovers in research, whereby the productivity of any one research firm depends on the economy-wide level of research. This implies that there can be more than one equilibrium growth rate. In a low-growth equilibrium, firms are discouraged from doing research because they cannot benefit from the work of many others; conversely, in a high-growth equilibrium they are encouraged to do a lot of research by the ease with which they can learn from each others' efforts.

Imperfections in capital markets can also be incorporated. Agency costs in discovering which research projects are more promising than others have the expected effect of reducing the steady-state level of research and rate of growth. This and the related idea of increasing returns to scale in the financial system both shed light on why the level of financial development is important in the growth process.

Most of this chapter assumes that innovations are "drastic," that each new generation of intermediate good is so much better than its predecessors that the latest innovator is not threatened by competition from previous innovators. It also shows, however, that all the comparative-statics results of the analysis go through in the nondrastic case. Finally, it allows research firms to choose not only the frequency but also the size of innovations. This generalization shows that under laissez-faire, innovations will be too small if they are drastic. In the nondrastic case, the tendency to make innovations too small is at least partly mitigated by the incentive for innovators to move away from their competitive fringe, which they can do by increasing the size of innovations.

Problems

Difficulty of Problems:

No star: normal

One star: difficult

Two stars: very difficult

1. Innovation by the leader (based on Barro and Sala-i-Martín 1995)

In the text we saw that the incumbent monopolist never does research. The evidence, however, tells us that a large proportion of improvements in the quality of existing products is done by the monopolist who is already manufacturing them. This can be due to the fact that the monopolist has a lower cost of research or a better knowledge of the product that implies a higher probability of success. Consider first an extension of the Schumpeterian model in which only the monopolist can undertake research.

a. Find the equation that determines the level of R&D employment.

b. Now suppose that both the incumbent monopolist and outsiders can do research, but the former has a higher probability of success, given by $\lambda_I z$, with $\lambda_I > \lambda$, and where z is the number of researchers employed by the firm. All other firms innovate with probability λz. Under which parameter values will the incumbent invest in R&D?

2. Welfare and patent auctions (based on Kremer 1996)

Despite the results of the quality ladder models, most empirical research has found that the rate of innovation is below the social optimum. Patents have the problem that the monopolist cannot appropriate the full social return, and hence too little is spent on research. They also lead to inefficient research targeted at inventing "around" the patent. The first-best solution are subsidies to R&D, but in practice they often result in more slack rather than more research, and they do not solve the problem of inventing around the patent. Government funded research might not be targeted at the most profitable innovations and is often burdened by bureaucracy. Kremer has proposed an auction mechanism that combines the advantages of direct funding of research with those of the patent system.

Assume that the economy behaves like the one-sector model of quality ladders presented in the text. Now consider the following system to purchase patents. The patent is auctioned and the (market) value of the patent is determined by the bidding. The government then offers to buy out patents at this private value times a fixed markup that covers the difference between the social and private values of the innovation. Inventors can decide whether or not to sell their patent. If sold, the patent is then placed in the public domain. However, there is a small proportion of the auctioned patents that are sold to the highest bidder, who then becomes a monopolist. This ensures truthful revelation, in as much as bidders are prevented from making offers at prices they would not be willing to pay.

Suppose that there is a standard patent system in the economy, and then the mechanism just described is introduced. Agents know that the mechanism will be used in all subsequent periods.

a. Find the social value of an innovation. What should the markup be?

b. Would this mechanism ensure static allocative efficiency? Would it ensure dynamic efficiency, that is, the socially optimal level of R&D?

c. Could this mechanism be implemented if the competitive economy generated excessive growth?