

1	<i>Chapter 3</i>	1
2		2
3	HORIZONTAL INNOVATION IN THE THEORY OF GROWTH AND	3
4	DEVELOPMENT	4
5		5
6	GINO GANCIA	6
7	<i>CREI and Universitat Pompeu Fabra</i>	7
8		8
9	FABRIZIO ZILIBOTTI	9
10	<i>Institute for International Economic Studies, Stockholm University</i>	10
11		11
12	Contents	12
13		13
14	Abstract	2 14
15	Keywords	2 15
16	1. Introduction	3 16
17	2. Growth with expanding variety	6 17
18	2.1. The benchmark model	6 18
19	2.2. Two variations of the benchmark model: “lab-equipment” and “labor-for intermediates”	10 19
20	2.3. Limited patent protection	12 20
21	3. Trade, growth and imitation	14 21
22	3.1. Scale effects, economic integration and trade	14 22
23	3.2. Innovation, imitation and product cycles	17 23
24	4. Directed technical change	20 24
25	4.1. Factor-biased innovation and wage inequality	21 25
26	4.2. Appropriate technology and development	26 26
27	4.3. Trade, inequality and appropriate technology	30 27
28	5. Complementarity in innovation	34 28
29	6. Financial development	40 29
30	7. Endogenous fluctuations	47 30
31	7.1. Deterministic cycles	48 31
32	7.2. Learning and sunspots	52 32
33	8. Conclusions	56 33
34	Acknowledgements	56 34
35	Uncited references	56 35
36	References	56 36
37		37
38		38
39		39
40		40
41		41
42	<i>Handbook of Economic Growth, Volume 1A. Edited by Philippe Aghion and Steven N. Durlauf</i>	42
43	© 2005 Elsevier B.V. All rights reserved	
	DOI: 10.1016/S1574-0684(05)01003-8	43

1 **Abstract**

2 We analyze recent contributions to growth theory based on the model of expanding
3 variety of Romer [Romer, P. (1990). “Endogenous technological change”. *Journal of*
4 *Political Economy* 98, 71–102]. In the first part, we present different versions of the
5 benchmark linear model with imperfect competition. These include the “lab-equipment”
6 model, “labor-for-intermediates” and “directed technical change”. We review applica-
7 tions of the expanding variety framework to the analysis of international technology
8 diffusion, trade, cross-country productivity differences, financial development and fluc-
9 tuations. In many such applications, a key role is played by complementarities in the
10 process of innovation.
11

12
13
14 **Keywords**

15 appropriate technology, complementarity, cycles, convergence, directed technical
16 change, endogenous growth, expanding variety, financial development, imperfect
17 competition, integration, innovation, intellectual property rights, imitation, knowledge,
18 learning, patents, technical change, trade, traps
19

20 *JEL classification:* D92, E32, F12, F15, F43, G22, O11, O16, O31, O33, O41, O47
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

1. Introduction

Endogenous growth theory formalizes the role of technical progress in explaining modern economic growth. Although this is a relatively recent development, many of its ideas were already stressed by authors such as Kuznets, Griliches, Schmookler, Rosenberg and Schumpeter. During the 1950s and 1960s, mainstream economics was dominated by the one-sector neoclassical growth model of Solow (1956) and Swan (1956), whose main focus was on capital accumulation. The model postulated the existence of an aggregate production function featuring constant returns to scale and returns to each input falling asymptotically to zero; given that some inputs cannot be accumulated, the model could not generate sustained growth unless technology was assumed to improve exogenously. This simple treatment of technology as exogenous was considered as unsatisfactory for two main reasons: first, by placing the source of sustained growth outside the model, the theory could not explain the determinants of long-run economic performance and second, empirical evidence pointed out that technical progress often depends on deliberate economic decisions.

The first attempts to endogenize the rate of technical change addressed the first, but not the second, problem. Assuming technical progress to be an unintentional by-product of the introduction of new capital goods through a process named “learning-by-doing”, Arrow (1962) was able to generate sustained growth at a rate that depended on investment decisions. Attempts at explicitly modeling investment in innovation faced another difficulty. A replication argument suggests that, for a given state of technology, production functions should exhibit constant returns to scale. If technical progress is considered as an additional input, however, the technology features increasing returns to scale and inputs cannot be paid their marginal product. Models of learning-by-doing avoided the problem by assuming that increasing returns were external to firms, thereby preserving perfect competition. However, this approach is not viable once investment in technology is recognized as intentional. The solution was to follow the view of Schumpeter (1942), that new technologies provide market power and that investment in innovation is motivated by the prospect of future profits. In this spirit, Shell (1973) studied the case of a single monopolist investing in technical change and Nordhaus (1969a) wrote a growth model with patents, monopoly power and many firms. In neither case did the equilibrium feature sustained growth.¹

A tractable model of imperfect competition under general equilibrium was not available until the analysis of monopolistic competition in consumption goods by Dixit and Stiglitz (1977), later extended to differentiated inputs in production by Ethier (1982). These models also showed how increasing returns could arise from an expansion in the number of varieties of producer and consumer goods, an idea that is at the core of the models studied in this chapter. The first dynamic models of economic growth with monopolistic competition and innovation motivated by profits were built by Judd (1985)

¹ See Levhari and Sheshinski (1969) on necessary and sufficient conditions for the existence of steady-state growth in the presence of increasing returns to scale.

1 and Grossman and Helpman (1989). Yet, these authors were interested in aspects other 1
2 than endogenous growth and none of their models featured long-run growth. Romer 2
3 (1987), who formalized an old idea of Young (1928), was the first to show that models 3
4 of monopolistic competition could generate long-run growth through the increased spe- 4
5 cialization of labor across an increasing range of activities. The final step was taken in 5
6 Romer (1990), which assumed that inventing new goods is a deliberate costly activity 6
7 and that monopoly profits, granted to innovators by patents, motivate discoveries. Since 7
8 then, the basic model of endogenous growth with an expanding variety of products has 8
9 been extended in many direction. 9

10 The distinctive feature of the models discussed in this chapter is “horizontal inno- 10
11 vation”: a discovery consists of the technical knowledge required to manufacture a 11
12 new good that does not displace existing ones. Therefore, innovation takes the form 12
13 of an expansion in the variety of available products. The underlying assumption is that 13
14 the availability of more goods, either for final consumption or as intermediate inputs, 14
15 raises the material well-being of people. This can occur through various channels. Con- 15
16 sumers may value variety per se. For example, having a TV set and a Hi-Fi yields more 16
17 utility than having two units of any one of them. Productivity in manufacturing may in- 17
18 crease with the availability of a larger set of intermediate tools, such as hammers, trucks, 18
19 computers and so on. Similarly, specialization of labor across an increasing variety of 19
20 activities, as in the celebrated Adam Smith example of the pin factory, can make aggre- 20
21 gate production more efficient. The main alternative approach is to model innovation 21
22 as quality improvements on a given array of products (“vertical innovation”), so that 22
23 technical progress makes existing products obsolete. This process of “creative destruc- 23
24 tion” was emphasized by Schumpeter and has been formalized in Aghion and Howitt 24
25 (1992), Grossman and Helpman (1991a, 1991b) and Segerstrom, Anant and Dinopou- 25
26 los (1990). The two approaches naturally complement each other. The main advantage 26
27 of models with horizontal innovation lies in their analytical tractability, making them 27
28 powerful tools for addressing a wide range of questions. However, because of their 28
29 simplistic view on the interaction between innovators, these models are less suited to 29
30 studying the effects of competition between “leaders” and “follower” on the growth 30
31 process. 31
32

33 Section 1 of this chapter describes a simplified version of Romer (1990) and some 33
34 extensions used in the literature. The model exhibits increasing returns to scale and 34
35 steady-state endogenous growth in output per capita and the stock of knowledge. The 35
36 key feature of the theory is the emphasis on investments in technical knowledge as 36
37 the determinant of long-run economic growth. Ideas and technological improvements 37
38 differ from other physical assets, because they entail important public good elements. 38
39 Inventing new technology is typically costly, while reproducing ideas is relatively inex- 39
40 pensive. Therefore, technical knowledge is described as a non-rival good. Nevertheless, 40
41 firms are willing to invest in innovation because there exists a system of intellectual 41
42 property rights (patents) guaranteeing innovators monopoly power over the production 42
43 and sales of particular goods. 43

1 Growth models with an expanding variety of products are a natural dynamic counter- 1
2 part to trade models based on increasing returns and product differentiation. As such, 2
3 they offer a simple framework for studying the effects of market integration on growth 3
4 and other issues in dynamic trade theory. This is the subject of Section 2, which shows 4
5 how trade integration can produce both static gains, by providing access to foreign 5
6 varieties, and dynamic gains, by raising the rate at which new goods are introduced. 6
7 Product-cycle trade and imitation are also considered. 7

8 In many instances, technical progress may be non-neutral towards different factors 8
9 or sectors. This possibility is considered in Section 3, where biased technical change 9
10 is incorporated in the basic growth model. By introducing several factors and sectors, 10
11 the economic incentives to develop technologies complementing a specific factor, such 11
12 as skilled workers, can be studied. These incentives critically depend on the defini- 12
13 tion of property rights over the production of new ideas. The high variability in the 13
14 effectiveness of patent laws across countries has important bearings on the form of 14
15 technical progress. In particular, governments in less developed countries may have an 15
16 incentive not to enforce intellectual property rights in order to speed up the process 16
17 of technology adoption. However, the undesired side effect of free-riding is that in- 17
18 novators in industrialized countries lose incentives to create improvements that are 18
19 most useful in developing countries, but of limited application in industrialized mar- 19
20kets. 20

21 Section 4 introduces complementarity in innovation. While innovation has no effect 21
22 on the profitability of existing intermediate firms in the benchmark model, in reality 22
23 new technologies can substitute or complement existing technologies. Innovation may 23
24 cause technological obsolescence of previous technologies, as emphasized by Schum- 24
25peterian models. In other cases, new technologies complement rather than substitute 25
26 the old ones. For instance, the market for a particular technology tends to be small 26
27 at the time of its introduction, but grows as new compatible applications are devel- 27
28oped. This complementarity in innovation can lead to multiple equilibria and poverty 28
29traps. 29
30

31 Complementarities in the growth process may also arise from financial markets, 31
32 as suggested in Section 5. The progressive endogenous enrichment of asset markets, 32
33 associated with the development of new intermediate industries, may improve the diver- 33
34sification opportunities available to investors. This, in turn, makes savers more prepared 34
35 to invest in high-productivity risky industries, thereby fostering further industrial and 35
36 financial development. As a result, countries at early stages of development go through 36
37 periods of slow and highly volatile growth, eventually followed by a take-off with fi- 37
38nancial deepening and steady growth. 38

39 Finally, Section 6 shows how models with technological complementarities can gen- 39
40erate rich long-run dynamics, including endogenous fluctuations between periods of 40
41high and low growth. Cycles in innovation and growth can either be due to expectational 41
42indeterminacy, or the deterministic dynamics of two-sector models with an endogenous 42
43market structure. 43

1 **2. Growth with expanding variety** 1

2
3 In this section, we present the benchmark model of endogenous growth with expanding 3
4 variety, and some extensions that will be developed in the following sections. 4

5
6 *2.1. The benchmark model* 6

7
8 The benchmark model is a simplified version of Romer (1990), where, for simplicity, 8
9 we abstract from investments in physical capital. The economy is populated by infinitely 9
10 lived agents who derive utility from consumption and supply inelastic labor. The popu- 10
11 lation is constant, and equal to L . Agents' preferences are represented by an isoelastic 11
12 utility function: 12

13
14
$$U = \int_0^\infty e^{-\rho t} \frac{C_t^{1-\theta} - 1}{1-\theta} dt. \quad (1)$$
 14

15
16 The representative household sets a consumption plan to maximize utility, subject to 16
17 an intertemporal budget constraint and a No-Ponzi game condition. The consumption 17
18 plan satisfies a standard Euler equation: 18

19
20
$$\dot{C}_t = \frac{r_t - \rho}{\theta} \cdot C_t. \quad (2)$$
 19

21 There is no physical capital, and savings are used to finance innovative investments. 21

22 The production side of the economy consists of two sectors of activity: a competitive 22
23 sector producing a homogeneous final good, and a non-competitive sector producing 23
24 differentiated intermediate goods. The final-good sector employs labor and a set of in- 24
25 termediate goods as inputs. The technology for producing final goods is represented by 25
26 the following production function: 26

27
28
$$Y_t = L_{y,t}^{1-\alpha} \int_0^{A_t} x_{j,t}^\alpha dj, \quad (3)$$
 28

29
30 where x_j is the quantity of the intermediate good j , A_t is the measure of intermedi- 30
31 ate goods available at t , L_y is labor and $\alpha \in (0, 1)$. This specification follows Spence 31
32 (1976), Dixit and Stiglitz (1977) and Ethier (1982). It describes different inputs as im- 32
33 perfect substitutes, which symmetrically enter the production function, implying that no 33
34 intermediate good is intrinsically better or worse than any other, irrespective of the time 34
35 of introduction. The marginal product of each input is decreasing, and independent of 35
36 the measure of intermediate goods, A_t . 36

37 The intermediate good sector consists of monopolistically competitive firms, each 37
38 producing a differentiated variety j . Technology is symmetric across varieties: the pro- 38
39 duction of one unit of intermediate good requires one unit of final good, assumed to 39
40 be the numeraire.² In addition, each intermediate producer is subject to a sunk cost to 40

41
42 ² In Romer (1990), the variable input is physical capital, and the economy has two state variables, i.e., 42
43 physical capital and knowledge. 43

1 design a new intermediate input variety. New designs are produced instantaneously and 1
2 with no uncertainty. The innovating firm can patent the design, and acquire a perpetual 2
3 monopoly power over the production of the corresponding input. 3

4 In the absence of intellectual property rights, free-riding would prevent any innovative 4
5 activity. If firms could costlessly copy the design, competition would drive ex-post rents 5
6 to zero. Then, no firms would have an incentive, ex-ante, to pay a sunk cost to design a 6
7 new input. 7

8 The research activity only uses labor. An important assumption is that innovation 8
9 generates an intertemporal externality. In particular, the design of a (unit measure of) 9
10 new intermediate good requires a labor input equal to $1/(\delta A_t)$. The assumption that 10
11 labor productivity increases with the stock of knowledge, A_t , can be rationalized by 11
12 the idea of researchers benefiting from accessing the stock of applications for patents, 12
13 thereby obtaining inspiration for new designs. 13

14 The law of motion of technical knowledge can be written as: 14

$$15 \quad \dot{A}_t = \delta A_t L_{x,t}, \quad (4) \quad 15$$

16 where δ is a parameter and L_x denotes the aggregate employment in research. The rate 16
17 of technological change is a linear function of total employment in research.³ Finally, 17
18 feasibility requires that $L \geq L_{x,t} + L_{y,t}$. 18

19 First, we characterize the equilibrium in the final good sector. Let w denote the wage, 19
20 and p_j be the price of the j 'th variety of intermediate input. The price of the final 20
21 product is the numeraire. The representative firm in the competitive final sector takes 21
22 prices as parametric and chooses production and technology so as to maximizes profit, 22
23 given by: 23
24

$$25 \quad \pi_t^Y = L_{y,t}^{1-\alpha} \int_0^{A_t} x_{j,t}^\alpha dj - w_t L_{y,t} - \int_0^{A_t} p_{j,t} x_{j,t} dj. \quad (5) \quad 25$$

26 The first-order conditions yield the following factor demands: 26
27

$$28 \quad p_{j,t} = \alpha L_{y,t}^{1-\alpha} x_{j,t}^{\alpha-1} \quad \forall j \in [0, A_t] \quad (6) \quad 28$$

29 and 29

$$30 \quad w_{y,t} = (1 - \alpha) L_{y,t}^{-\alpha} \int_0^{A_t} x_{j,t}^\alpha dj. \quad (7) \quad 30$$

31
32
33
34
35
36
37 ³ Jones (1995) generalizes this technology and lets

$$38 \quad \dot{A}_t = \delta A_t^{\gamma_A} L_{x,t}^{(1-\gamma_L)}, \quad 38$$

39 where $\gamma_A \leq 1$ is a positive externality through the stock of knowledge and γ_L is a negative externality that 39
40 can be interpreted as coming from the duplication of research effort. Assuming $\gamma_A < 1$ leads to qualitative 40
41 differences in the prediction of the model. In particular, the specification where $\gamma_A = 1$ and $\gamma_L = 0$, which is 41
42 the model discussed here, generates scale effects. See further discussion later in this chapter and, especially, 42
43 in Chapter ? of this Handbook. 43

1 Next, consider the problem of intermediate producers. A firm owning a patent sets 1
2 its production level so as to maximize the profit, subject to the demand function (6). 2
3 The profit of the firm producing the j th variety is $\pi_{j,t} = p_{j,t}x_{j,t} - x_{j,t}$. The optimal 3
4 quantity and price set by the monopolist are 4

$$5 \quad x_{j,t} = x_t = \alpha^{2/(1-\alpha)} L_{y,t} \quad \text{and} \quad p_{j,t} = p = 1/\alpha, \quad (8) \quad 5$$

6 respectively. Hence, the maximum profit for an intermediate producer is 6
7

$$8 \quad \pi_{j,t} = \pi_t = (p - 1)x_t = \frac{1 - \alpha}{\alpha} \alpha^{2/(1-\alpha)} L_{y,t}. \quad (9) \quad 8$$

9 Substitution of x_t into (7) yields the equilibrium wage as: 9
10

$$11 \quad w_t = (1 - \alpha) \alpha^{2\alpha/(1-\alpha)} A_t. \quad (10) \quad 11$$

12 Next, we guess-and-verify the existence of a balanced growth (BG) equilibrium, such 12
13 that consumption, production and technical knowledge grow at the same constant rate, 13
14 γ , and the two sectors employ constant proportions of the workforce.⁴ In BG, both the 14
15 production and the profits of intermediate firms, as given by Equations (8) and (9), are 15
16 constant over time and across industries. Thus, $x_t = x$ and $\pi_t = \pi$. 16
17

18 Free entry implies that the present discounted value (PDV) of profits from innovation 18
19 cannot exceed the entry cost. By the Euler equation, (2), the interest rate is also constant 19
20 in BG. Hence, the PDV of profits equals π/r . The entry cost is given by the wage paid 20
21 to researchers, i.e., $w_t/(\delta A_t)$. Therefore, the free entry condition can be written as: 21
22

$$23 \quad \frac{\pi}{r} \leq \frac{w_t}{\delta A_t}. \quad (11) \quad 23$$

24 We can then use (9) and (10), and substitute the expressions of π and w_t into (11): 24
25

$$26 \quad \frac{(\frac{1-\alpha}{\alpha}) \alpha^{2/(1-\alpha)} L_y}{r} \leq \frac{(1 - \alpha) \alpha^{2\alpha/(1-\alpha)}}{\delta}. \quad (12) \quad 26$$

27 The right-hand side expression is the marginal cost of innovation, independent of A_t , 27
28 due to the cancellation of two opposite effects. On the one hand, labor productivity and, 28
29 hence, the equilibrium wage grow linearly with A_t . On the other hand, the productivity 29
30 of researchers increases with A_t , due to the intertemporal knowledge spillover. Thus, the 30
31 unit cost of innovation is constant over time. Note that, without the externality, the cost 31
32 of innovation would grow over time, and technical progress and growth would come to 32
33 a halt, like in the neoclassical model. 33
34

35 For innovation to be positive, (12) must hold with equality. We can use (i) the resource 35
36 constraint, implying that $L_y = L - L_x$, and (ii) the fact that, from (4) and BG, $L_x = 36$
37 γ/δ , to express (12) as a relationship between the interest rate and the growth rate: 37
38

$$39 \quad r = \alpha(\delta L - \gamma). \quad (13) \quad 39$$

40
41
42 ⁴ The equilibrium that we characterized can be proved to be unique. Moreover, the version of Romer's model 42
43 described here features no transitional dynamics, as in *AK* models [Rebelo (1991)]. 43

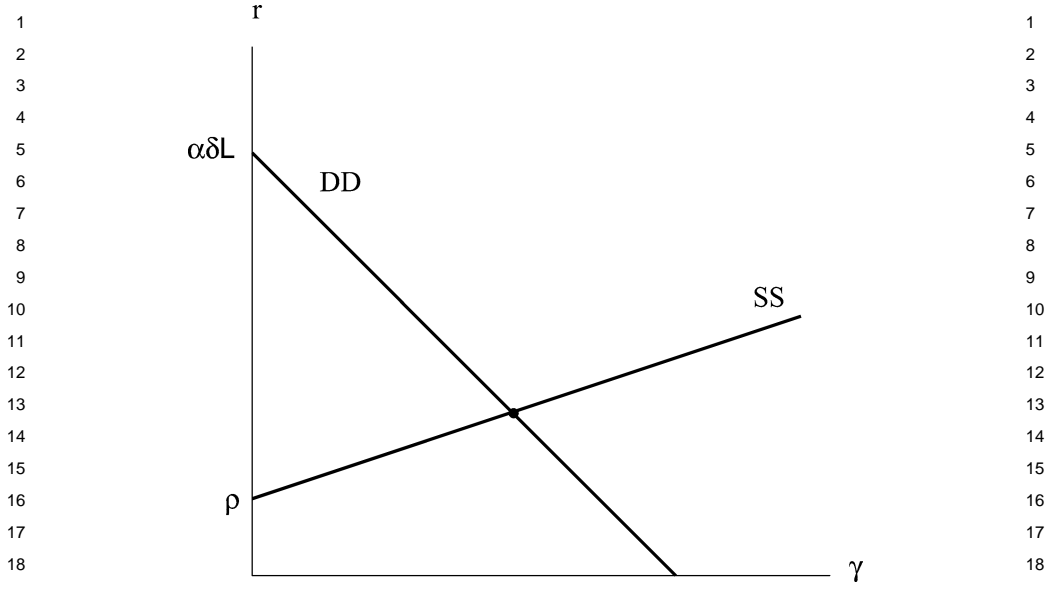


Figure 1.

Equation (13) describes the equilibrium condition on the production side of the economy: the higher is the interest rate that firms must pay to finance innovation expenditure, the lower is employment in research and growth.

Finally, the consumption Euler equation, (2), given BG, yields:

$$r = \rho + \theta\gamma, \quad (14)$$

which is the usual positive relation between interest rate and growth. Figure 1 plots the linear equations (13) and (14), which characterize the equilibrium. The two equations correspond, respectively, to the DD (demand for funds) and SS (supply of savings) linear schedules.

An interior solution exists if and only if $\alpha\delta L > \rho$. When this condition fails to be satisfied, all workers are employed in the production of consumption goods. When it is positive, the equilibrium growth rate is

$$\gamma = \frac{\delta\alpha L - \rho}{\alpha + \theta}, \quad (15)$$

showing that the growth rate is increasing in the productivity of the research sector (δ), the size of the labor force (L) and the intertemporal elasticity of substitution of consumption ($1/\theta$), while it is decreasing in the elasticity of final output to labor, $(1 - \alpha)$, and the discount rate.

The trade-off between final production (consumption), on the one hand, and innovation and growth, on the other hand, can be shown by substituting the equilibrium

1 expression of x into the aggregate production function, (3). This yields: 1

$$2 \quad Y_t = \alpha^{2\alpha/(1-\alpha)} L_y A_t = \alpha^{2\alpha/(1-\alpha)} (L - \gamma/\delta) A_t. \quad (16) \quad 2$$

3 The decentralized equilibrium is inefficient for two reasons:⁵ 3

- 4 1. Intermediate firms exert monopoly power, and charge a price in excess of the 4
- 5 marginal cost of production. This leads to an underproduction of each variety of 5
- 6 intermediate goods. 6
- 7 2. the accumulation of ideas produces externalities not internalized in the laissez- 7
- 8 faire economy. Innovating firms compare the private cost of innovation, $w_t/(\delta A_t)$, 8
- 9 with the present discounted value of profits, π/r . However, they ignore the 9
- 10 spillover on the future productivity of innovation. 10
- 11 11

12 Contrary to Schumpeterian models, innovation does not cause “creative destruction”, 12

13 i.e., no rent is reduced by the entry of new firms. As a result, growth is always sub- 13

14 optimally low in the laissez-faire equilibrium. Policies aimed at increasing research 14

15 activities (e.g., through subsidies to R&D or intermediate production) are both growth- 15

16 and welfare-enhancing. This result is not robust, however. Benassy (1998) shows that 16

17 in a model where the return to specialization is allowed to vary and does not depend on 17

18 firms’ market power (α), research and growth in the laissez-faire equilibrium may be 18

19 suboptimally too high. 19

20 2.2. *Two variations of the benchmark model: “lab-equipment” and “labor-for* 20

21 *intermediates”* 21

22 22

23 We now consider two alternative specifications of the model that have been used in the 23

24 literature, and that will be discussed in the following sections. The first specification 24

25 is the so-called “lab-equipment” model, where the research activity uses final output 25

26 instead of labor as a productive input.⁶ More formally, Equation (4) is replaced by the 26

27 condition $\dot{A}_t = Y_x/\mu$, where Y_x denotes the units of final output devoted to research 27

28 (hence, consumption is $C = Y - Ax - Y_x$) and μ the output cost per unit of innovation. 28

29 In the lab-equipment model, there is no research spillover of the type discussed in the 29

30 benchmark model. Labor is entirely allocated to final production ($L_y = L$), and the 30

31 free-entry condition (12) is replaced by 31

$$32 \quad \frac{\frac{1-\alpha}{\alpha} \alpha^{2/(1-\alpha)} L}{r} \leq \mu. \quad (17) \quad 32$$

33 33

34 34

35 35

36 36

37 ⁵ There is an additional reason why, in general, models with a Dixit–Stiglitz technology can generate 37

38 inefficient allocations in laissez-faire, namely that the range of intermediate goods produced is endogenous. 38

39 The standard assumption of complete markets is violated in Dixit–Stiglitz models, because there is no market 39

40 price for the goods not produced. This issue is discussed in Matsuyama (1995, 1997). A dynamic example 40

41 of such a failure is provided by the model of Acemoglu and Zilibotti (1997), which is discussed in detail in 41

42 Section 6. 42

43 ⁶ The “lab-equipment” model was first introduced by Rivera-Batiz and Romer (1991a); see also Barro and 43

Sala-i-Martin (1995).

Hence, using the Euler condition, (14), we obtain the following equilibrium growth rate:

$$\gamma = \frac{(1 - \alpha)\alpha^{(1+\alpha)/(1-\alpha)}L/\mu - \rho}{\theta}.$$

Sustained growth is attained by allocating a constant share of production to finance the research activity.

The second specification assumes that labor is not used in final production, but is used (instead of final output) as the unique input in the intermediate goods production.⁷ More formally, the final production technology is

$$Y_t = Z^{1-\alpha} \int_0^{A_t} x_{j,t}^\alpha dj, \quad (18)$$

where Z is a fixed factor (e.g., land) that is typically normalized to unity and ignored. In this model, $1/A_t$ units of labor are required to produce one unit of any intermediate input, with constant marginal costs. Therefore, in this version of the model, innovation generates a spillover on the productivity of both research and intermediate production.⁸ We refer to this version as the “labor-for-intermediates” model.

It immediately follows that, in equilibrium, the production of each intermediate firm equals $x = L - L_x$. The price of intermediates is once more a mark-up over the marginal cost, $p_t = w_t/(\alpha A_t)$. In a BG equilibrium, wages and technology grow at the same rate, hence their ratio is constant. Let $\omega \equiv (w_t/A_t)$. The maximum profit is, then:

$$\pi = \left(\frac{1 - \alpha}{\alpha}\right)\omega x = \frac{1 - \alpha}{\alpha}\omega(L - L_x).$$

The free entry condition can be expressed as:

$$\frac{1 - \alpha}{r\alpha}(L - L_x) \leq \frac{1}{\delta},$$

hence,

$$\gamma = \left(\frac{1 - \alpha}{\alpha}L\delta - \rho\right) / \left(\frac{1 - \alpha}{\alpha} + \theta\right).$$

Clearly, both the “lab-equipment” and “labor-for-intermediates” model yield solutions qualitatively similar to that of the benchmark model.

⁷ We follow the specification used by Young (1993). A related approach, treating the variety of inputs as consumption goods produced with labor, is examined in Grossman and Helpman (1991a, 1991b).

⁸ The spillover on the productivity of intermediate production is not necessary to have endogenous growth. Without it, an equilibrium can be found in which production of each intermediate falls as A grows: $\gamma_A = -\gamma_x$. In this case, employment in production, Ax , is constant and the growth rate of Y is $(1 - \alpha)\gamma_A$.

1 2.3. *Limited patent protection* 1

2
3 In this section, we discuss the effects of limited patent protection. For simplicity, we 3
4 focus on the lab-equipment version discussed in the previous section. The expectation 4
5 of monopoly profits provides the basic incentive motivating investment in innovation; at 5
6 the same time, monopoly rights introduce a distortion in the economy that raises prices 6
7 above marginal costs and causes the underprovision of goods. Since the growth rate of 7
8 knowledge in the typical decentralized equilibrium is below the social optimum, the 8
9 presence of monopoly power poses a trade-off between dynamic and static efficiency, 9
10 leading to the question, first studied by Nordhaus (1969a, 1969b), of whether there 10
11 exists an optimal level of protection of monopoly rights. In the basic model, we assumed 11
12 the monopoly power of innovators to last forever. Now, we study how the main results 12
13 change when agents cannot be perfectly excluded from using advances discovered by 13
14 others. A tractable way of doing this is to assume monopoly power to be eroded at a 14
15 constant rate, so that in every instant, a fraction m of the monopolized goods becomes 15
16 competitive.⁹ Then, for a given range of varieties in the economy, A_t , the number of 16
17 “imitated” intermediates that have become competitive, A_t^* , follows the law of motion: 17
18

$$19 \quad \dot{A}_t^* = m(A_t - A_t^*). \quad (19) \quad 19$$

20 Stronger patent protection can be considered as a reduction in the imitation rate m . Note 20
21 that the model now has two state variables, A_t and A_t^* , and will exhibit transitional 21
22 dynamics. In general, from any starting point, the ratio A_t^*/A_t will converge to the 22
23 steady-state level.¹⁰ 23
24

$$25 \quad \frac{A^*}{A} = \frac{m}{\gamma + m}, \quad (20) \quad 25$$

26 where $\gamma \equiv \dot{A}/A$. 27

28 Once a product is imitated, the monopoly power of the original producer is lost and 28
29 its prices is driven down to the marginal cost by competition. Thus, at each point in 29
30 time, intermediates still produced by monopolists are sold as before at the markup price 30
31 $1/\alpha$, while for the others, the competitive price is one. Substituting prices into demand 31
32 functions yields the quantity of each intermediate sold in equilibrium: 32
33

$$34 \quad x_j = \alpha^{1/(1-\alpha)} L \equiv x^* \quad \text{for } j \in (0, A_t^*), \quad 34$$

$$35 \quad x_j = \alpha^{2/(1-\alpha)} L \equiv x \quad \text{for } j \in (A_t^*, A_t). \quad (21) \quad 35$$

36 Note that $x^* > x$, because the monopolized goods have a higher price. 36
37
38
39

40 ⁹ A growth model with limited patent life is developed by Judd (1985). Here, we follow Barro and Sala-i 40
41 Martin (1995). An alternative way of introducing limited patent protection is to assume monopolies to have a 41
42 deterministic lifetime T . In this case, the PDV of an innovation is $(1 - e^{-rT})\pi/r$ (assuming balanced growth). 42

43 ¹⁰ This can be seen imposing $\dot{A}^*/A^* = \gamma$ in (19). 43

1 Free entry requires the PDV of profits generated by an innovation, V , to equal its
2 cost μ . Along the balanced growth path, where the interest rate is constant, arbitrage
3 in asset markets requires the instantaneous return to innovation, π/μ , to equal the real
4 interest rate adjusted for imitation risk: $r + m$.¹¹ Since prices and quantities of the
5 monopolized goods are identical to those in the basic model, π is not affected by im-
6 itation. Imitation only affects the duration of the profit flow, which is reflected in the
7 effective interest rate. Therefore, limiting patent lives introduces a new inefficiency: al-
8 though the benefit from a discovery is permanent for the economy, the reward for the
9 innovator is now only temporary. Using the Euler equation for consumption growth,
10 $\gamma = (r - \rho)/\theta$, and the adjusted interest rate in (17), we get the growth rate of the
11 economy:

$$12 \quad \gamma = \frac{1}{\theta} \left[(1 - \alpha) \alpha^{(1+\alpha)(1-\alpha)} \frac{L}{\mu} - m - \rho \right].$$

15 As expected, the growth rate is decreasing in the imitation rate, as the limited dura-
16 tion of the monopoly effectively reduces the private value of an innovation. If we were
17 concerned about long-run growth only, it would then be clear that patents should al-
18 ways be fully and eternally protected. However, for a given level of technology, A_t ,
19 output is higher the shorter is the patent duration (higher m), as can be seen by substi-
20 tuting equilibrium quantities (21) and the ratio of imitated goods (20) in the production
21 function (3):

$$22 \quad Y_t = \alpha^{2\alpha/(1-\alpha)} A_t L \left[1 + \left(\frac{m}{\gamma + m} \right) (\alpha^{-\alpha/(1-\alpha)} - 1) \right].$$

25 Therefore, a reduction in the patent life entails a trade-off between an immediate con-
26 sumption gain and future losses in terms of lower growth, and its quantitative analysis
27 requires the calculation of welfare along the transition. Kwan and Lai (2003) perform
28 such an analysis, both numerically and by linearizing the BG equilibrium in the neigh-
29 borhood of the steady-state, and show the existence of an optimum patent life. They
30 also provide a simple calibration, using US data on long-run growth, markups and plau-
31 sible values for ρ and θ , to suggest that over-protection of patents is unlikely to happen,
32 whereas the welfare cost of under-protection can be substantial.

33 Alternatively, the optimal patent length can be analytically derived in models with a
34 simpler structure. For example, Grossman and Lai (2004) construct a modified version
35 of the model described above, where they assume quasi-linear functions. They show the
36

37
38 ¹¹ A simple way of seeing this is through the following argument. In a time interval dt , the firm provides a
39 profit stream $\pi \cdot dt$, a capital gain of $\dot{V} \cdot dt$ if not imitated and a capital loss V if imitated (as the value of the
40 patent would drop to zero). In the limit $dt \rightarrow 0$, the probability of being imitated in this time interval is $m \cdot dt$
41 and the probability of not being imitated equals $(1 - m \cdot dt)$. Therefore, the expected return for the firm is
42 $\pi \cdot dt + (1 - m \cdot dt) \dot{V} \cdot dt - mV \cdot dt$. Selling the firm and investing the proceeds in the capital market would
43 yield an interest payment of $rV \cdot dt$. Arbitrage implies that the returns from these two forms of investment
should be equal and in a steady state $\dot{V} = 0$, implying $\pi/V = r + m$.

1 optimal patent length to be an increasing function of the useful life of a product, of con- 1
2 sumers' patience and the ratio of consumers' and producers' surplus under monopoly to 2
3 consumers' surplus under competition. In addition, they derive the optimal patent length 3
4 for noncooperative trading countries and find that advanced economies with a higher in- 4
5 novative potential will, in general, grant longer patents. A similar point is made in Lai 5
6 and Qiu (2003). 6

9 3. Trade, growth and imitation 9

11 Growth models with an expanding variety of products are a natural dynamic counterpart 11
12 to the widely-used trade models based on increasing returns and product differentiation 12
13 developed in the 1980s [e.g., Helpman and Krugman (1985)]. As such, they offer a 13
14 simple framework for studying the effects of market integration on growth and other 14
15 issues in dynamic trade theory. Quality-ladder models have also been proposed in this 15
16 literature, but they are a less natural counterpart to the static new trade theory, as they 16
17 do not focus on the number of varieties available in an economy and their growth rate. 17
18 As we shall see, economic integration can provide both static gains, through the access 18
19 to a wider range of goods, and dynamic gains, through an increase in the rate at which 19
20 new varieties are introduced. However, the results may vary when integration is limited 20
21 to commodity markets with no international diffusion of knowledge [Rivera-Batiz and 21
22 Romer (1991a)] and when countries differ in their initial stock of knowledge [Devereux 22
23 and Lapham (1994)]. 23
24

25 Finally, the analysis in this section is extended to product-cycle trade: the introduction 25
26 of new products in advanced countries and their subsequent imitation by less developed 26
27 countries. An important result will be to show that, contrary to the closed economy 27
28 case, imitation by less developed countries may spur innovation and growth [Helpman 28
29 (1993)]. 29

31 3.1. Scale effects, economic integration and trade 31

32 In this section, we use the benchmark model to discuss the effects of trade and integra- 32
33 tion. The model features scale effects. Take two identical countries with identical labor 33
34 endowment, $L = L^*$. In isolation, both countries would grow at the same rate, as given 34
35 by (15). But if they merge, the growth rate of the integrated country increases to: 35
36

$$37 \gamma^I = \frac{\delta\alpha(L + L^*) - \rho}{\alpha + \theta} = \frac{2\alpha\delta L - \rho}{\alpha + \theta}. \quad 37$$

38 Therefore, the model predicts that economic integration boosts growth. 38

39 Integration, even if beneficial, may be difficult to achieve. However, in many in- 39
40 stances, trade operates as a substitute for economic integration. Rivera-Batiz and Romer 40
41 42
43

(1991a) analyze under which condition trade would attain the same benefits as economic integration. To this aim, they consider two experiments:¹²

1. The economies can trade at no cost in goods and assets, but knowledge spillovers remain localized within national borders;
2. In addition, knowledge spillovers work across borders after trade.

In both cases, to simplify the analysis, the two economies are assumed to produce, before trade, disjoint subsets of intermediate goods. This assumption avoids complications arising from trade turning monopolies into duopolies in those industries which exist in both countries. Clearly, after trade, there would be no incentive for overlap in innovation, and the importance of inputs that were historically produced in both countries would decline to zero over time.

We start from the case analyzed by Rivera-Batiz and Romer (1991a), where the two countries are perfectly identical before trade. Namely, $L = L^*$ and $A_0 = A_0^*$, where the star denotes the *foreign* economy, and time zero denotes the moment when trade starts. Since, in a BG equilibrium, $\gamma = \delta L_x$, trade can only affect growth via the split of the workforce between production and research. Such a split, however, is not affected by trade, for in the symmetric equilibrium, trade increases by the same proportion the productivity of workers in production and the profitability of research. Since both the cost and private benefit of innovation increase by the same factor, investments in innovation remain unchanged.

More formally, the after trade wage is

$$w_{\text{trade}} = (1 - \alpha)L_y^{-\alpha}x^\alpha(A + A^*), \quad (22)$$

which is twice as large as in the pre-trade equilibrium since at the moment of trade liberalization, $A = A^*$. Higher labor costs are a disincentive to research. But trade also increases the market for intermediate goods. Each monopolist can now sell its product in two markets. Since the demand elasticity is the same in both markets, the monopoly price equals $1/\alpha$ in both markets. Thus, the after trade profit is

$$\pi_{\text{trade}} = (p - 1)(x + x^*) = 2 \frac{1 - \alpha}{\alpha} \alpha^{2/(1-\alpha)} L_y.$$

The free-entry condition becomes, for both countries:

$$2 \frac{\frac{1-\alpha}{\alpha} \alpha^{2/(1-\alpha)} L_y}{r} \leq 2 \frac{(1 - \alpha) \alpha^{2\alpha/(1-\alpha)}}{\delta}, \quad (23)$$

which, after simplifying, is identical to (12). Therefore, the split of the workforce between production and research remains unchanged, and trade has no permanent effects on growth. Opening up to free trade, however, induces a once-and-for-all gain: both output and consumption increase in both countries, similarly to an unexpected increase

¹² The original article considers two versions of the model, one using the benchmark set-up and the other using the “lab-equipment” version. For the sake of brevity, we restrict the attention to the first. Romer (1994) extends the analysis to the case when a tariff on imports is imposed.

1 in the stock of knowledge, since final producers in both countries can use a larger set of 1
2 intermediate goods. 2

3 This result is not robust to asymmetric initial conditions. Devereux and Lapham 3
4 (1994) show that if, initially, the two countries have different productivity levels, trade 4
5 leads to specialization and a rise in the world growth rate.¹³ Consider the economies 5
6 described above, but assume that $A_0 < A_0^*$. Recall that free-entry implies: 6
7

$$8 \quad V \leq \frac{w}{\delta A} \quad \text{and} \quad V^* \leq \frac{w^*}{\delta A^*}, \quad 8$$

9 where V, V^* denote the PDV of profits for an intermediate firm located at home and 9
10 abroad, respectively. First, trade in intermediate goods and free capital markets equalize 10
11 the rate of return to both financial assets (r) and labor (w).¹⁴ Second, monopoly profits 11
12 are independent of firms' locations, thereby implying that the value of firms must be the 12
13 same all over the world: $V = V^* = V^w$. Therefore, at the time of trade liberalization, 13
14 we must have: 14
15

$$16 \quad \frac{w_{\text{trade}}}{\delta A} > \frac{w_{\text{trade}}}{\delta A^*} \geq V^w, \quad 16$$

17 implying that no innovation is carried out in equilibrium in the (home) country, starting 17
18 from a lower productivity. Moreover, the productivity gap in R&D widens over time: 18
19 indeed, trade forever eliminates the incentives to innovate in the initially poorer country. 19
20

21 In the richer (foreign) country, however, trade boosts innovation.¹⁵ The value of for- 21
22 eign firms must satisfy the following Bellman equation: 22
23

$$24 \quad rV^* = \dot{V}^* + \frac{1-\alpha}{\alpha} \alpha^{2/(1-\alpha)} (L_y^* + L), \quad 24$$

25 where we note that $L_y = L$. The free-entry condition implies that: 25
26

$$27 \quad V^* = \frac{(1-\alpha)\alpha^{2\alpha/(1-\alpha)} A^* + A}{\delta A^*}. \quad 27$$

28 Since knowledge only accumulates in the foreign country, the value of intermediate 28
29 firms must decline over time, and in the long-run tend to its pre-trade value, i.e., $V^* =$ 29
30 $(1-\alpha)\alpha^{2\alpha/(1-\alpha)}/\delta$. Therefore, in the long run, the free-entry condition is 30
31

$$32 \quad \frac{\frac{1-\alpha}{\alpha} \alpha^{2/(1-\alpha)} (L + L_y^*)}{r} \leq \frac{(1-\alpha)\alpha^{2\alpha/(1-\alpha)}}{\delta}. \quad 32$$

33
34
35
36
37
38 ¹³ See also Rivera-Batiz and Romer (1991b) on the effects of trade restrictions with asymmetric countries. 38

39 ¹⁴ Recall Equation (22). The equalization of wages descends from a particular feature of the equilibrium, i.e., 39
40 that the marginal product of labor is independent of the level of employment in production (since x is linear 40
41 in L_y). This feature is not robust. If the production technology had land as an input, for instance, wages would 41
42 not be equalized across countries; see Devereux and Lapham (1994) for an analysis of the more general case. 42

43 ¹⁵ Our discussion focuses on a world where no economy becomes fully specialized in research, since this 43
seems to be the empirically plausible case.

1 Comparing (24) with (12) shows that trade reduces employment in production and, consequently, increases the long-run research activity in the foreign country, which implies
2 that trade increases growth. In terms of Fig. 1, trade creates an outward shift in the DD
3 schedule, leading to a higher interest rate and faster growth in equilibrium.

4
5 The result can be interpreted as trade leading to specialization. The home country
6 specializes in final production, while the foreign country diversifies between manufac-
7 turing and innovation.¹⁶ This is efficient, since there are country-wide economies of
8 scale in innovation. Although trade leads to zero innovation in the home country, mar-
9 kets are integrated: final good producers, in both countries, can use the same varieties
10 of intermediates and all consumers in the world can invest in the innovative firms of the
11 foreign economy. Therefore, the location of innovation and firms has no impact on the
12 relative welfare of the two countries.

13 Consider now the case when trade induces cross-country flows of ideas, i.e., if the
14 knowledge spillover is determined, after trade, by the world stock of ideas contained in
15 the union of A and A^* . When free trade is allowed, the accumulation of knowledge in
16 each country is given by

$$\dot{A} = \delta L_x (A + A^*) \quad \text{and} \quad \dot{A}^* = \delta L_x^* (A + A^*).$$

17
18
19 Even if trade did not affect the allocation of the workforce between production and
20 research, the rate of growth of technology would increase. But there is an additional
21 effect; the larger knowledge spillover increases labor productivity in research, inducing
22 an increase of employment in research. Formally, the total effect is equivalent to an
23 increase in parameter δ . In terms of Figure 1, trade in goods plus flow of ideas imply an
24 upward shift of the DD locus for both countries. Hence, trade attains the same effect as
25 economic integration (increasing δ is equivalent to increasing L). This result is robust
26 to asymmetric initial conditions.

27 28 3.2. *Innovation, imitation and product cycles*

29
30 The model just presented may be appropriate for describing trade integration between
31 similar countries, but it misses important features of North–South trade. In a seminal
32 article, Vernon (1966) argued that new products are first introduced in rich countries (the
33 North), where R&D capabilities are high and the proximity to large and rich markets
34 facilitates innovation. After some time, when a product reaches a stage of maturity and
35 manufacturing methods become standardized, the good can easily be imitated and then,
36 the bulk of production moves to less developed countries (the South), to take advantage
37 of low wages. The expanding variety model provides a natural framework for studying
38 the introduction of new goods and their subsequent imitation (product cycle trade).¹⁷

39
40
41 ¹⁶ Home-country patent holders will still produce intermediates, but as compared to the world's stock of
intermediates, they will be of measure zero.

42 ¹⁷ Quality ladder models of innovation have been used to study product-cycles by, among others, Grossman
43 and Helpman (1991a, 1991b), Segerstrom et al. (1990) and Dinopoulos and Segerstrom (2003).

1 We have already discussed imitation within the context of a closed economy. Here, we 1
2 extend the analysis to the case where a richer North innovates, while a poorer South only 2
3 engages in imitation. The analysis yields new results that modify some of the previous 3
4 conclusions on the effect of imitation on innovation. The key questions are, first, how the 4
5 transfer of production to the South through imitation affects the incentives to innovate 5
6 and, second, how it affects the income distribution between North and South. 6

7 Following Helpman (1993), consider a two-region model of innovation, imitation and 7
8 trade. Assume that R&D, producing new goods, is performed in the North only and that 8
9 costless imitation takes place in the South at a constant rate m .¹⁸ The imitation rate can 9
10 be interpreted as an inverse measure of protection of Intellectual Property Rights (IPRs). 10
11 Once a good is copied in the South, it is produced by competitive firms. Therefore, 11
12 at every point in time, there is a range A_t^N of goods produced by monopolists in the 12
13 North and a range A_t^S of goods that have been copied and are produced in the South 13
14 by competitive firms. Given that the rate of introduction of new good is $\gamma = \dot{A}_t/A_t$, 14
15 where $A_t = A_t^N + A_t^S$, and that monopolized goods are copied at the instantaneous rate 15
16 m , $\dot{A}_t^S = mA_t^N$, it follows that a steady-state where the ratio A_t^N/A_t^S is constant must 16
17 satisfy: 17

$$18 \frac{A^N}{A} = \frac{\gamma}{\gamma + m} \quad \text{and} \quad \frac{A^S}{A} = \frac{m}{\gamma + m}. \quad (25) \quad 19$$

20 We use the “labor-for-intermediates” version of the growth model, so that the price 20
21 of a single variety depends on the prevailing wage rate in the country where it is man- 21
22 ufactured. This is an important feature of product cycle models, allowing the North to 22
23 benefit from low production costs in the South for imitated goods. Therefore, we define 23
24 the aggregate production function as in (18): 24
25

$$26 Y_t = \int_0^{A_t} x_i^\alpha di, \quad (26) \quad 27$$

28 where A_t is the (growing) range of available products x_i and $\varepsilon = 1/(1-\alpha)$ is the elastic- 28
29 ity of substitution between any two varieties. Intermediates are manufactured with $1/A_t$ 29
30 units of labor per unit of output in both regions. Northern firms charge a monopoly 30
31 price, as long as their products have not been imitated, equal to a constant markup $1/\alpha$ 31
32 over the production cost, given by the wage rate. On the contrary, Southern firms pro- 32
33 duce imitated goods that have become competitive and sell them at a price equal to the 33
34 marginal cost. To summarize: 34
35

$$36 p_t^N = \frac{w_t^N}{\alpha A_t} \quad \text{and} \quad p_t^S = \frac{w_t^S}{A_t}, \quad (27) \quad 37$$

38 where p_t^N and p_t^S are the prices of any variety of intermediates produced in the North 38
39 and South, respectively. 39
40
41
42
43

¹⁸ The rate of imitation is made endogenous in Grossman and Helpman (1991b).

As in the benchmark model, innovation requires labor: the introduction of new products per unit of time \dot{A}_t equals $\delta A_t L_x$, where L_x is the (Northern) labor input employed in R&D, δ is a productivity parameter and A_t captures an externality from past innovations. This implies that the growth rate of the economy is a linear function of the number of workers employed in R&D, $\gamma = L_x \delta$. As usual, profits generated by the monopoly over the sale of the new good are used to cover the cost of innovation. Since the profits per product are a fraction $(1 - \alpha)$ of total revenue $p^N x$ and the labor market clears, $A_t^N x / A_t + \gamma / \delta = L^N$, profits can be written as:

$$\pi^N = \frac{1 - \alpha}{\alpha} \frac{w_t^N}{A_t^N} \left(L^N - \frac{\gamma}{\delta} \right). \quad (28)$$

Arbitrage in asset markets implies that $(r + m)V^N = \pi^N + \dot{V}^N$, where V^N is the PDV of a new good and the effective interest rate is adjusted by the imitation risk. Along a BG path, $\dot{V}^N = 0$ and free entry ensures that the value of an innovation equals its cost, $w_t^N / \delta A_t$. Combining these considerations with (25) and (28) yields:

$$\frac{1 - \alpha}{\alpha} (\delta L^N - \gamma) \frac{\gamma + m}{\gamma} = r + m. \quad (29)$$

Together with the Euler equation for consumption growth, (29) provides an implicit solution for the long-run growth rate of innovation. Note that the left-hand side is the profit rate (i.e., instantaneous profits over the value of the innovation) and the right-hand side represents the effective cost of capital, inclusive of the imitation risk.

To see the effect of a tightening of IPRs (a reduction of m), consider how an infinitesimal change in m affects the two sides of (29). Taking a log linear approximation, the impact of m on the profit rate is $1/(\gamma + m)$, whereas the effect on the cost of capital is $1/(r + m)$. In the case of log preferences, studied by Helpman (1993), $r > \gamma$. Hence, a reduction of m has a larger impact on the profit rate than on the effective cost of capital, thereby reducing the profitability of innovation and growth. What is the effect on the fraction of goods produced in the North? Rewriting (29) with the help of (25) as:

$$\frac{A^N}{A} = \frac{1 - \alpha}{\alpha} (\delta L^N - \gamma) \frac{1}{r + m}, \quad (30)$$

it becomes apparent that a reduction of m increases the share of goods manufactured in the North, both through its direct effect and by reducing γ and r .

To understand these results, note that stronger IPRs have two opposite effects. First, a lower imitation rate prolongs the expected duration of the monopoly on a new product developed in the North, thereby increasing the returns to innovation. Second, since firms produce for a longer time in the North, it rises the demand for Northern labor, w^N , and hence, the cost of innovation. For the specification with log utility, the latter effect dominates and innovation declines. More generally, the link between the rate of imitation and innovation can go either way [as in Grossman and Helpman (1991a)]. However, the important result here is that tighter IPRs does not necessarily stimulate innovation in the long run.

The effect of IPRs on the North–South wage ratio can be found using (27), together with the relative demand for intermediates:¹⁹

$$\frac{w^N}{w^S} = \alpha \left[\frac{A^N}{A^S} \frac{L_S}{L_N - \gamma/\delta} \right]^{1/\varepsilon} \quad (31)$$

Given that a decline in the imitation rate m raises $(A^N/A^S)/(L_N - \gamma/\delta)$ (see Equation (30)), a tightening of IPRs raises the relative wage of the North. Helpman (1993) computes welfare changes in the North and in the South (including transitional dynamics) after a change in the imitation rate m , and concludes that the South is unambiguously hurt by a decline in imitation. Moreover, if the imitation rate is not too high, the North can also be worse-off.

More recent papers on product cycles, incorporating the notion that stronger IPRs make relocation of production to the South a more attractive option, have come to different conclusions. For example, by assuming that Northern multinationals can produce in the South and that Southern firms can only imitate after production has been transferred to their country, Lai (1998) shows that stronger IPRs increase the rate of product innovation and the relative wage of the South. Similarly, Yang and Maskus (2001) find that if Northern firms can license their technology to Southern producers, being subject to an imitation risk, stronger IPRs reduce the cost of licensing, free resources for R&D and foster growth, with ambiguous effects on relative wages. Finally, the literature on appropriate technology [e.g., Diwan and Rodrik (1991), Acemoglu and Zilibotti (2001), Gancia (2003)] has shown that, when the North and the South have different technological needs, the South has an incentive to protect IPRs in order to attract innovations more suited to their technological needs. Some of these results are discussed in the next sections.

4. Directed technical change

So far, technical progress has been modeled as an increase in total factor productivity (A) that is neutral towards different factors and sectors. For many applications, however, this assumption is not realistic. For example, there is evidence that technical progress has been skill-biased during the last century and that this bias accelerated during the 1980s. Similarly, the fact that the output shares of labor and capital have been roughly constant in the US while the capital–labor ratio has been steadily increasing

¹⁹ Relative demand for intermediates is:

$$\frac{p^N}{p^S} = \left(\frac{x^N}{x^S} \right)^{\alpha-1}$$

Using $x^N = AL_y^N/A^N$, $x^S = AL_y^S/A^S$ and the pricing formula (27) yields the expression in the text.

1 suggests that technical change has mainly been labor-augmenting.²⁰ Further, industry 1
2 studies show R&D intensity to vary substantially across sectors. In order to build a the- 2
3 ory for the direction of technical change, a first step is to introduce more sectors into the 3
4 model. Then, studying the economic incentives to develop technologies complementing 4
5 a specific factor or sector can help understand what determines the shape of technology. 5

6 An important contribution of this new theory will be to shed light on the determi- 6
7 nants of wage inequality [Acemoglu (1998, 2003a)]. Another application studies under 7
8 which circumstances technologies developed by profit-motivated firms are appropriate 8
9 for the economic conditions of the countries where they are used. The analysis will 9
10 demonstrate that, since IPRs are weakly protected in developing countries, new tech- 10
11 nologies tend to be designed for the markets and needs of advanced countries. As a 11
12 result, these technologies yield a low level of productivity when adopted by developing 12
13 countries [Acemoglu and Zilibotti (2001)]. Trade can reinforce this problem and create 13
14 interesting general equilibrium effects. 14

15 Although most of the results discussed in this section can be derived using models of 15
16 vertical innovation, the expanding variety approach has proved to be particularly suited 16
17 for addressing these issues because of its analytical tractability and simple dynamics. 17
18 For instance, creative destruction, a fundamental feature of quality-ladder models, is 18
19 not a crucial element for the problems at hand, and abstracting from it substantially 19
20 simplifies the analysis. 20
21

22 4.1. *Factor-biased innovation and wage inequality* 22 23

24 Directed technical change was formalized by Acemoglu (1998), and then integrated by 24
25 Acemoglu and Zilibotti (2001) into a model of growth with expanding variety, to ex- 25
26 plain the degree of skill-complementarity of technology.²¹ In this section, we discuss 26
27 the expanding variety version [following the synthesis of Acemoglu (2002)] by extend- 27
28 ing the “lab-equipment” model to two sectors employing skilled and unskilled labor, 28
29 respectively. Consider the following aggregate production function: 29
30

$$31 \quad Y = [Y_L^{(\varepsilon-1)/\varepsilon} + Y_H^{(\varepsilon-1)/\varepsilon}]^{\varepsilon/(\varepsilon-1)}, \quad (32) \quad 31$$

32 where Y_L and Y_H are goods produced with unskilled labor, L , and skilled labor, H , 32
33 respectively. Y represents aggregate output, used for both consumption and investment, 33
34 as a combination of the two goods produced in the economy, with an elasticity of sub- 34
35 stitution equal to ε . Maximizing Y under a resource constraint gives constant elasticity 35
36 demand functions, implying a negative relationship between relative prices and relative 36
37

38
39
40 ²⁰ Unless the production function is Cobb–Douglas, in which case the direction of technical progress is 40
41 irrelevant. Empirical estimates suggest that the elasticity of substitution between labor and capital is likely to 41
42 be less than one. See Hamermesh (1993) for a survey of early estimates and Krusell et al. (2000) and Antras 42
43 (2004) for more recent contributions. 42

43 ²¹ Important antecedents are Kennedy (1964), Samuelson (1965) and Atkinson and Stiglitz (1969). 43

1 quantities:

$$2 \quad \frac{P_H}{P_L} = \left[\frac{Y_L}{Y_H} \right]^{1/\varepsilon}, \quad (33)$$

3 where P_L and P_H are the prices of Y_L and Y_H , respectively. Aggregate output is chosen
4 as the numeraire, hence:

$$5 \quad (P_L^{1-\varepsilon} + P_H^{1-\varepsilon})^{1/(1-\varepsilon)} = 1. \quad (34)$$

6 The distinctive feature of this model is that the two goods are now produced using
7 different technologies:

$$8 \quad Y_L = L^{1-\alpha} \int_0^{A_L} x_{L,j}^\alpha dj, \quad (35)$$

$$9 \quad Y_H = H^{1-\alpha} \int_0^{A_H} x_{H,j}^\alpha dj,$$

10 where $x_{L,j}$, $j \in [0, A_L]$, are intermediate goods complementing unskilled labor,
11 whereas $x_{H,j}$, $j \in [0, A_H]$ complement skilled labor. This assumption captures the fact
12 that different factors usually operate with different technologies and that a new tech-
13 nology may benefit one factor more than others.²² For example, it has been argued that
14 computers boosted the productivity of skilled more than that of unskilled labor, whereas
15 the opposite occurred after the introduction of the assembly line. As before, technical
16 progress takes the form of an increase in the number of intermediate goods, $[A_L, A_H]$,
17 but now an innovator must decide which technology to expand. The profitability of the
18 two sectors pins down, endogenously, the direction of technical change. In a steady-
19 state equilibrium, there is a constant ratio of the number of intermediates used by each
20 factor, A_H/A_L , and this can be interpreted as the extent of the “endogenous skill-bias”
21 of the technology.

22 The analysis follows the same steps as in model with a single factor. Final good pro-
23 ducers take the price of their output (P_L, P_H), the price of intermediates ($p_{L,j}, p_{H,j}$)
24 and wages (w_L, w_H) as given. Consider a variety j used in the production of Y_L . Profit
25 maximization gives the following isoelastic demand:

$$26 \quad x_{L,j} = \left[\frac{\alpha P_L}{p_{L,j}} \right]^{1/(1-\alpha)} L, \quad (36)$$

27 and an equivalent expression for $x_{H,j}$.

28 The intermediate good sector is monopolistic, with each producer owning the patent
29 for a single variety. The cost of producing one unit of any intermediate good is one

30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

²² The analysis can be generalized to specifications where, in the spirit of Heckscher–Ohlin models, each sector uses all productive factors, but factor intensities differ across sectors. The model can also be generalized to more than two factors and sectors.

1 unit of the numeraire. The symmetric structure of demand and technology implies that
2 all monopolists set the same price, $p_{Lj} = p_L$. In particular, given the isoelastic de-
3 mand, they set $p_L = 1/\alpha$ and sell the quantity $x_{L,j} = (\alpha^2 P_L)^{1/(1-\alpha)} L$. The profit flow
4 accruing to intermediate producers can therefore be expressed as

$$5 \quad \pi_L = (1 - \alpha)\alpha^{(1+\alpha)/(1-\alpha)} (P_L)^{1/(1-\alpha)} L. \quad (37)$$

6 Similar conclusions are reached for varieties used in the production of Y_H , leading to

$$7 \quad \pi_H = (1 - \alpha)\alpha^{(1+\alpha)/(1-\alpha)} (P_H)^{1/(1-\alpha)} H. \quad (38)$$

8 From (37)–(38), it immediately follows that the relative profitability in the two sectors
9 is given by

$$10 \quad \frac{\pi_H}{\pi_L} = \left(\frac{P_H}{P_L} \right)^{1/(1-\alpha)} \frac{H}{L}, \quad (39)$$

11 which, since profits are used to finance innovation, is also the relative profitability of
12 R&D directed to the two sectors. The first term in (39) represents the “price effect”:
13 there is a greater incentive to invent technologies producing more expensive goods.²³
14 The second term is the “market size” effect: the incentive to develop a new technology
15 is proportional to the number of workers that will be using it.²⁴

16 Next, using the price of intermediates in (36) and (35) gives final output in each
17 sector:

$$18 \quad Y_L = \alpha^{2\alpha/(1-\alpha)} P_L^{\alpha/(1-\alpha)} A_L L, \quad (40)$$

$$19 \quad Y_H = \alpha^{2\alpha/(1-\alpha)} P_H^{\alpha/(1-\alpha)} A_H H.$$

20 Note the similarity with (16). As in the benchmark model, output – in each sector –
21 is a linear function of technology and labor. But sectoral output now also depends on
22 sectoral prices, P_L and P_H , since a higher price of output increases the value of produc-
23 tivity of intermediates, but not their costs, and therefore encourages firms to use more of
24 them, thereby raising labor productivity. Note that this is not the case in the one-sector
25 model since there, the price of output is proportional to the price of intermediates.

26 We can now solve for prices and wages as functions of the state of technology and
27 endowments. Using (40) into (33) and noting that the wage bill is a constant fraction of
28 sectoral output, yields:

$$29 \quad \frac{P_H}{P_L} = \left[\frac{A_H H}{A_L L} \right]^{-(1-\alpha)/\sigma}, \quad (41)$$

$$30 \quad \frac{w_H}{w_L} = \left[\frac{A_H}{A_L} \right]^{1-1/\sigma} \left[\frac{H}{L} \right]^{-1/\sigma}, \quad (42)$$

31 ²³ The price effect, restated in terms of factor prices, was emphasized by Hicks (1932) and Habakkuk (1962).

32 ²⁴ Market size, although in the context of industry- and firm-level innovation, was emphasized as a determi-
33 nant of technical progress by Griliches and Schmookler (1963), Schmookler (1966) and Schumpeter (1950).

1 where $\sigma \equiv 1 + (1 - \alpha)(\varepsilon - 1)$ is, by definition, the elasticity of substitution between 1
2 H and L .²⁵ Note that the skill premium, w_H/w_L , is decreasing in the relative supply of 2
3 skilled labor (H/L) and increasing in the skill-bias (A_H/A_L). 3

4 The final step is to find the equilibrium for technology. We assume, as in the lab- 4
5 equipment model of Section 2.2, that the development of a new intermediate good to 5
6 require a fixed cost of μ units of the numeraire. Free entry and an arbitrage condition 6
7 require the value V_z of an innovation directed to factor $Z \in \{L, H\}$ to equal its cost. 7
8 Since the value of an innovation is the PDV of the infinite stream of profits it generates, 8
9 an equilibrium with a positive rate of innovation in both types of intermediates such 9
10 that the ratio A_H/A_L remains constant, i.e., a BG path where P_H/P_L , w_H/w_L and 10
11 π_H/π_L are also constant, requires profit equalization in the two sectors, $\pi_H = \pi_L = \pi$. 11
12 Imposing this restriction yields the equilibrium skill-bias of technology, 12

$$13 \frac{A_H}{A_L} = \left[\frac{H}{L} \right]^{\sigma-1} . \quad (43) \quad 14$$

15 Equation (43) shows that, as long as workers of different skill levels are gross substitutes 15
16 ($\sigma > 1$), an increase in the supply of one factor will induce more innovation directed 16
17 to that specific factor. This is the case because, with $\sigma > 1$, the market size effect 17
18 dominates the price effect, and technology is biased towards the abundant factor. The 18
19 opposite is true if $\sigma < 1$. As usual, the growth rate of the economy can be found from 19
20 the free-entry condition $\pi_Z/r = \mu$, $Z \in \{L, H\}$. Using (34), (41) and (43) to substitute 20
21 for prices and the interest rate from the Euler equation, yields: 21
22

$$22 \gamma = \frac{1}{\theta} \left[\frac{(1 - \alpha)\alpha^{(1+\alpha)/(1-\alpha)}}{\mu} (L^{\sigma-1} + H^{\sigma-1})^{1/(\sigma-1)} - \rho \right]. \quad 23$$

24 If we only had one factor (e.g., $H = 0$), the growth rate would reduce to that of the 24
25 benchmark model. 25

26 Directed technical change has interesting implications on factor prices. Using (43), 26
27 the skill-premium becomes: 27

$$28 \frac{w_H}{w_L} = \left[\frac{H}{L} \right]^{\sigma-2} . \quad (44) \quad 29$$

30 Equation (44) shows that the slope of the labor demand curve, i.e., the relationship 30
31 between relative wages and relative labor supply, can be either positive or negative and is 31
32 the result of two opposite forces. On the one hand, a large supply of one factor depresses 32
33 the price of its product while, on the other hand, it induces a technology bias in its 33
34 favor, thereby raising its productivity. A high substitutability between H and L implies 34
35 a weak price effect of an increase in relative supply, which makes a positive relationship 35
36 more likely. In particular, if $\sigma > 2$, the market size effect is sufficiently strong to not 36
37

37 38 39 40 41 42 43
25 This is the short-run elasticity of substitution between L and H , for a given technology A_L and A_H . 43

1 only dominate the price effect on technical change (see Equation (43)), but also the
2 substitution effect between skilled and unskilled workers at a given technology.

3 This result can help rationalize several facts. First, it suggests that technical change
4 has been skill biased during the past 60 years, because of the steady growth in the
5 supply of skilled labor. Second, the case $\sigma > 2$ offers an explanation for the fall and
6 rise in the US skill premium during the 1970s and 1980s. In the 1970s, there was a large
7 increase in the supply of skilled labor (H/L). Assuming this shock to be unexpected, the
8 model predicts an initial fall in the skill premium (recall that A_H/A_L is a state variable
9 that does not immediately adjust), followed by its rise due to the induced skill biased
10 technical change, a pattern broadly consistent with the evidence.

11 In Acemoglu (2003b), this set-up is used to study the direction of technical progress
12 when the two factors of production are capital and labor. Beyond the change of notation,
13 the resulting model has an important qualitative difference, as capital can be accumu-
14 lated. The main finding is that, when both capital and labor augmenting innovations
15 are allowed, a balanced growth path still exists and features labor-augmenting technical
16 progress only. The intuition is that, while there are two ways of increasing the produc-
17 tion of capital-intensive goods (capital-augmenting technical change and accumulation),
18 there is only one way of increasing the production of labor-intensive goods (labor-
19 augmenting technical progress). Therefore, in the presence of capital accumulation,
20 technical progress must be more labor-augmenting than capital-augmenting. Further,
21 Acemoglu shows that, if capital and labor are gross complements (i.e., the elasticity of
22 substitution between the two is less than one), which seems to be the empirically rel-
23 evant case [see, for example, Antras (2004)], the economy converges to the balanced
24 growth path.

25 Finally, the theory of directed technical change can be used to study which industries
26 attract more innovation and why R&D intensity differs across sectors. In this exercise,
27 following a modified version of Klenow (1996), we abstract from factor endowments as
28 determinants of technology, by assuming there to be a single primary input, which we
29 call labor. Instead, other characteristics can make one sector more profitable than others.
30 Major explanations put forward in the literature on innovation are industry differences
31 in technological opportunities, market size and appropriability of rents, all factors that
32 can easily be embedded in the basic model with two sectors. In particular, to capture the
33 market size hypothesis, we introduce a parameter η defining the relative importance of
34 industry i in aggregate consumption:

$$Y = [\eta Y_i^{(\varepsilon-1)/\varepsilon} + (1 - \eta) Y_j^{(\varepsilon-1)/\varepsilon}]^{\varepsilon/(\varepsilon-1)}.$$

35
36
37
38
39 Differences in technological opportunities can be incorporated by allowing the cost of
40 an innovation, μ_i , to vary across sectors. Finally, we assume that an inventor in in-
41 dustry i can only extract a fraction λ_i of the profits generated by his innovation. The
42 previous analysis carries over almost unchanged, with the main difference that we now
43 need to solve for the allocation of labor across industries. This can be done requiring all

1 industries to pay the same wage, i.e., setting (42) equal to one: 1

$$\frac{L_i}{L_j} = \left(\frac{A_i}{A_j} \right)^{\sigma-1}.$$

2
3
4
5 Solving the new arbitrage condition stating that innovation for the two industries should 5
6 be equally profitable in BG, $\lambda_i \pi_i / \mu_i = \lambda_j \pi_j / \mu_j$, yields the relative industry-bias of 6
7 technology: 7
8

$$\frac{A_i}{A_j} = \left(\frac{\lambda_i \mu_j}{\lambda_j \mu_i} \right)^{1/(2-\sigma)} \left(\frac{\eta}{1-\eta} \right)^{\varepsilon/(2-\sigma)}.$$

9
10
11 As expected, industries with a larger market size, better technological opportunities 12
13 and higher appropriability attract more innovations.²⁶ Empirical estimates surveyed by 13
14 Cohen and Levin (1989) suggest that about one half of the industry differences in re- 14
15 search intensity can be attributed to the available measures of these three factors. 15
16

17 4.2. *Appropriate technology and development* 17

18
19 Directed technical change has interesting implications for the analysis of some devel- 19
20 opment issues. Acemoglu and Zilibotti (2001) show that technologies resulting from 20
21 directed technical change are optimal for the economic conditions of the markets where 21
22 they are sold. They analyze the implications of this finding in a two-country world where 22
23 technological innovation takes place in the North, and the South does not enforce (or 23
24 imperfectly enforce) IPRs. In this environment, innovators in the North can only extract 24
25 rents from selling technologies (embodied in new varieties of intermediate goods) in 25
26 the Northern market, since new technologies can be copied and locally produced in the 26
27 South. Thus, innovation does not respond to the factor endowment of the South: the 27
28 equilibrium skill-bias of technical change (see Equation (43) in the previous section) 28
29 is determined by the factor endowment of the North only. In this sense, technological 29
30 development tends to be “inappropriate” for the South: there is too much investment 30
31 in inventing new technologies augmenting the productivity of skilled workers, and too 31
32 little in inventing new technologies augmenting the productivity of unskilled workers. 32
33 Such excessive skill-bias prevents the South from fully profiting from technological im- 33
34 provements. The theory can explain North–South productivity differences, even when 34
35 the technology is identical and there are no significant barriers to technology adop- 35
36 tion.²⁷ 36
37

38
39 ²⁶ This is true as long as $\sigma < 2$. This restriction is required to have balanced growth. If violated, e.g., if goods 39
40 are highly substitutable, it would be profitable to direct innovation to one sector only. 40

41 ²⁷ Evidence on cross-country TFP differences is provided by, among others, Klenow and Rodriguez-Clare 41
42 (1997), Hall and Jones (1999), Caselli, Esquivel and Lefort (1996) and Prescott (1998). The view that tech- 42
43 nological differences arise from barriers to technology adoption is expressed by, among others, Parente and 43
Prescott (1994) and Prescott (1998).

We start by studying the set of advanced countries, called North. A continuum of measure one of final goods is produced by competitive firms. Final goods, indexed by $i \in [0, 1]$, are aggregated to give a composite output, $Y = \exp(\int_0^1 \log y_i di)$, which is the numeraire. There are two differences with respect to the model of the previous section: first, there is a continuum of sectors, not just two, and second, the elasticity of substitution between sectors is unity.²⁸ Each good i can be produced with both skilled and unskilled labor using two sets of intermediate goods: intermediates $[0, A_L]$ used by unskilled workers only and intermediates $[0, A_H]$ used by skilled workers only. Therefore, despite the continuum of sectors, there are only two types of technologies, as in the basic model of directed technical change. The production function takes the following form:

$$y_i = [(1-i)l_i]^{1-\alpha} \int_0^{A_L} x_{L,v,i}^\alpha dv + [ih_i]^{1-\alpha} \int_0^{A_H} x_{H,v,i}^\alpha dv, \quad (45)$$

where l_i and h_i are the quantities of unskilled and skilled labor employed in sector i , respectively, and $x_{z,v,i}$ is the quantity of intermediate good of type v used in sector i together with the labor of skill level $z = L, H$. Note that sectors differ in labor-augmenting productivity parameters, $(1-i)$ for the unskilled technology and i for the skilled technology, so that unskilled labor has a comparative advantage in sectors with a low index. Producers of good i take the price of their product, P_i , the price of intermediates ($p_{L,v}, p_{H,v}$) and wages (w_L, w_H) as given. Profit maximization gives the following demands for intermediates:

$$x_{L,v,i} = (1-i)l_i[\alpha P_i/p_{L,v}]^{1/(1-\alpha)} \quad \text{and} \quad x_{H,v,i} = ih_i[\alpha P_i/p_{H,v}]^{1/(1-\alpha)}. \quad (46)$$

The intermediate good sector is monopolistic. Each producer holds the patent for a single type of intermediate good v , and sells its output to firms in the final good sectors. The cost of producing one unit of any intermediate is conveniently normalized to α^2 units of the numeraire. Profit maximization by monopolists implies that prices are a constant markup over marginal costs, $p = \alpha$. Using the price of intermediates together with (46) and (45) gives the final output of sector i as a linear function of the number of intermediate goods and labor:

$$y_i = P_i^{\alpha/(1-\alpha)} [A_L(1-i)l_i + A_Hih_i]. \quad (47)$$

From (47), it is easily seen that all sectors whose index i is below a threshold level J will use the unskilled technology only and the remaining sectors will employ the skilled technology only. This happens because of the comparative advantage of unskilled workers in low index sectors and the linearity of the production function (there is no incentive to combine the two technologies and, for a given i , one always dominates the other).

²⁸ The composite output Y can be interpreted as a symmetric Cobb–Douglas over the measure of final goods $i \in [0, 1]$.

1 The total profits earned by monopolists are: 1

$$\begin{aligned}
 2 \quad \pi_{L,v} &= (1 - \alpha)\alpha \int_0^1 P_i^{1/(1-\alpha)} (1 - i)l_i di \quad \text{and} & 2 \\
 3 & & 3 \\
 4 & & 4 \\
 5 \quad \pi_{H,v} &= (1 - \alpha)\alpha \int_0^1 P_i^{1/(1-\alpha)} ih_i di. & 5 \\
 6 & & 6 \\
 7 & & 7
 \end{aligned} \tag{48}$$

8 Note that, by symmetry, $\pi_{L,v} = \pi_{L,j}$ and $\pi_{H,v} = \pi_{H,j}$. Given the Cobb–Douglas
specification in (45), the wage bill in each sector is a fraction $(1 - \alpha)$ of sectoral output.
9 Therefore, Equation (47) can be used to find wages:²⁹ 9

$$10 \quad w_L = (1 - \alpha)P_i^{1/(1-\alpha)}A_L(1 - i) \quad \text{and} \quad w_H = (1 - \alpha)P_i^{1/(1-\alpha)}A_Hi. \tag{49}$$

11 Defining $P_L \equiv P_0$, $P_H \equiv P_1$ and dividing equations in (49) by their counterparts 11
in sectors 0 and 1, respectively, it is possible to derive the following pattern of prices: 12
13 for $i \leq J$, $P_i = P_L(1 - i)^{-(1-\alpha)}$ and for $i \geq J$, $P_i = P_Hi^{-(1-\alpha)}$. Intuitively, the 13
14 price of a good produced with skilled (unskilled) labor is decreasing in the sectoral 14
15 productivity of skilled (unskilled) workers. Next, note that to maximize Y , expenditures 15
16 across goods must be equalized, i.e., $P_iy_i = P_Hy_1 = P_Ly_0$ (as for a symmetric Cobb– 16
17 Douglas). This observation, plus the given pattern of prices and full employment, imply 17
18 that labor is evenly distributed among sectors: $l_i = L/J$, $h_i = H/(1 - J)$, as prices 18
19 and sectoral productivity compensate each other. Finally, in sector $i = J$, it must be the 19
20 case that both technologies are equally profitable or $P_L(1 - J)^{-(1-\alpha)} = P_HJ^{-(1-\alpha)}$; 20
21 this condition, using $P_Hy_1 = P_Ly_0$ and (47), yields: 21
22 22

$$23 \quad \frac{J}{1 - J} = \left(\frac{P_H}{P_L}\right)^{1/(1-\alpha)} = \left(\frac{A_H H}{A_L L}\right)^{-1/2}. \tag{50}$$

24 The higher the relative endowment of skill (H/L) and the skill-bias of technology 24
25 (A_H/A_L), the larger the fraction of sectors using the skill-intensive technology $(1 - J)$. 25
26 Finally, integrating P_iy_i over $[0, 1]$, using (47), (50) and the fact that the consumption 26
27 aggregate is the numeraire (i.e., $\exp[\int_0^1 \ln P_i di] = 1$) gives a simple representation for 27
28 aggregate output: 28

$$29 \quad Y = \exp(-1)[(A_L L)^{1/2} + (A_H H)^{1/2}]^2, \tag{51}$$

30 which is a CES function of technology and endowments, with an elasticity of substitu- 30
31 tion between factors equal to two. 31

32 So far, the analysis defines an equilibrium for a given technology. Next, we need to 32
33 study innovation and characterize the equilibrium skill-bias of technology, (A_H/A_L) . 33
34 As before, technical progress takes the form of an increase in A_L and A_H and is the 34
35 35

36 ²⁹ In Acemoglu and Zilibotti (2001), there is an additional parameter ($Z > 1$), which is here omitted for 36
37 simplicity, which augments the productivity of skilled workers, ensuring that the skill premium is positive in 37
38 equilibrium. 38
39 39
40 40
41 41
42 42
43 43

1 result of directed R&D investment. The cost of an innovation (of any type) is equal to 1
2 μ units of the numeraire, and R&D is profitable as long as the PDV of the infinite flow 2
3 of profits that a producer of a new intermediate expects to earn covers the fixed cost 3
4 of innovation. Finally, free entry ensures that there are no additional profits. Using the 4
5 price pattern, instantaneous profits can be simplified as: 5
6

$$7 \quad \pi_H = \alpha(1 - \alpha)P_H^{1/(1-\alpha)}H. \quad (52) \quad 7$$

8 A parallel expression gives π_L . Balanced growth requires $\pi_L = \pi_H$; in this case, A_H 8
9 and A_L grow at the same rate, the ratio A_H/A_L is constant as are J , P_L and P_H . 9
10 Imposing $\pi_L = \pi_H$ in (52) and using (50) yields: 10
11

$$12 \quad \frac{A_H}{A_L} = \frac{1 - J}{J} = \frac{H}{L}. \quad (53) \quad 12$$

13 Note that the equilibrium skill-bias is identical to that of (43) in the special case when 13
14 $\sigma = 2$. Further, (53) shows that the higher is the skill endowment of a country, the larger 14
15 is the range of sectors using the skilled technology. This is a complete characterization 15
16 of the equilibrium for fully integrated economies developing and selling technologies 16
17 in their markets with full protection of IPRs and can be interpreted as a description of 17
18 the collection of rich countries, here called the North. 18
19

20 Consider now Southern economies, where skilled labor is assumed to be relatively 20
21 more scarce: $H^S/L^S < H^N/L^N$. Assume that intellectual property rights are not 21
22 enforced in the South and that there is no North–South trade. It follows that intermediate 22
23 producers located in the North cannot sell their goods or copyrights to firms located 23
24 in the South, so that the relevant market for technologies is the Northern market only. 24
25 Nonetheless, Southern producers can copy Northern innovations at a small but positive 25
26 cost. As a consequence, no two firms in the South find it profitable to copy the same 26
27 innovation and all intermediates introduced in the North are immediately copied (pro- 27
28 vided that the imitation cost is sufficiently small) and sold to Southern producers by a 28
29 local monopolist. Under these assumptions, firms in the South take the technologies de- 29
30 veloped originating in the North as given and do not invest in innovation.³⁰ This means 30
31 that both the North and the South use the same technologies, but $A_H/A_L = H^N/L^N$, 31
32 i.e., the skill-bias is determined by the factor endowment of the North, since this is the 32
33 only market for new technologies. Except for this, the other equilibrium conditions also 33
34 apply to the South after substituting the new endowments, H^S and L^S . 34
35

36 We are now ready to answer the following questions: are technologies appropriate 36
37 for the skill endowment of the countries where they are developed? What happens to 37
38 aggregate productivity if they are used in a different economic environment? 38
39

40
41 ³⁰ Imitation can be explicitly modelled as an activity similar to innovation, but less costly. Assuming the cost 41
42 of an innovation of type z to decrease with the distance from the relevant technology frontier A_Z^N , as in Barro 42
43 and Sala-i-Martin (1997), would yield very similar results. 43

1 Simple differentiation on (51) establishes that Y is maximized for $A_H/A_L = H/L$. 1
2 This is exactly condition (53), showing that the equilibrium skill-bias is optimally cho- 2
3 sen for the Northern skill composition. On the contrary, since factor abundance in the 3
4 South does not affect the direction of technical change, new technologies developed 4
5 in the North are inappropriate for the needs of the South. As a consequence, output 5
6 per capita, $Y/(L + H)$ is greater in the North than in the South. The reason for these 6
7 productivity differences is a technology–skill mismatch. To understand why, note that, 7
8 from Equation (50), $J^S > J^N$. Rewriting (53) as $A_H J^N = A_L(1 - J^N)$ and inspecting 8
9 Equation (47) reveals that unskilled workers are employed in the North up to sector J^N , 9
10 where they become as productive as skilled workers. This basic efficiency condition is 10
11 violated in the South, where $A_H J^S > A_L(1 - J^S)$. Because of its smaller skill endow- 11
12 ment, the South is using low-skill workers in some sectors where high-skill workers 12
13 would be more productive. 13

14 This result can help understand the existence of substantial differences in TFP across 14
15 countries, even when the technology is common. In particular, Acemoglu and Zilibotti 15
16 (2001) compare the predictive power of their model in explaining cross-country out- 16
17 put differences with that of a comparable neoclassical model, where all countries have 17
18 access to the same technologies and output is Cobb–Douglas in labor, human and phys- 18
19 ical capital. Their computations suggest that the proposed mechanism can account for 19
20 one-third to one half of the total factor productivity gap between the United States and 20
21 developing countries. Predictions on the pattern of North–South, cross-industry, produc- 21
22 tivity differences are also tested. Since the South uses the same technology [A_L, A_H] as 22
23 the rest of the world, but it has a higher relative price for skill-intensive goods, it follows 23
24 that the value of productivity in LDCs relative to that of the North should be higher in 24
25 skill-intensive sectors. The empirical analysis supports this prediction. 25

26 The view that countries adopt different technologies out of a world “menu”, and that 26
27 the choice of the appropriate technology depends on factor endowments, particularly on 27
28 the average skill of the labor force, finds support in the analysis of Caselli and Coleman 28
29 (2000). However, these authors also find that many poor countries choose technolo- 29
30 gies inside the world technology frontier, thereby suggesting that barriers to technology 30
31 adoption may also be important to explain the low total factor productivity of these 31
32 countries. 32

33 34 *4.3. Trade, inequality and appropriate technology* 34

35 36 We have seen that directed technical change can help understand inequality, both within 36
37 and between countries. Several authors have stressed that international trade is another 37
38 important determinant of income distribution. For example, Wood (1994) argues that 38
39 the higher competition with imports from LDCs may be responsible for the deterio- 39
40 ration in relative wages of low-skill workers in the US in the past decades. Further, 40
41 there is a widespread concern that globalization may be accompanied by a widening 41
42 of income differences between rich and poor countries. Although the analysis of these 42
43 issues goes beyond the scope of this paper, we want to argue that R&D-driven endoge- 43

nous growth models can fruitfully be used to understand some of the links between trade and inequality. In particular, we now show that trade with LDCs can have a profound impact on income distribution, beyond what is suggested by static trade theory, through its effect on the direction of technical change. By changing the relative prices and the location of production, international trade can change the incentives for developing innovations targeted at specific factors or sectors, systematically benefiting certain groups or countries more than others. A key assumption in deriving these results is that, as in the previous paragraph, LDCs do not provide an adequate protection of IPRs.

First, consider the effect of trade in the benchmark model of directed technical change. The analysis follows Acemoglu (2002, 2003a). Recall that the profitability of an innovation depends on its market size and the price of the goods it produces, as in Equation (39). What happens to technology if we allow free trade in Y_L and Y_H between a skill-abundant North and a skill-scarce South? The market size for innovations does not change, because inventors continue to sell their machines in the North only. But trade, at first, will increase the relative price of skill-intensive goods in the North. To see this, note that trade generates a single world market with a relative price depending on the world supply of goods. Since skills are scarcer in the world economy than in the North alone, trade will increase the relative price of skill-intensive goods in the North (the opposite will happen in the South). In particular, world prices are now given by Equation (41) using world endowments:

$$\frac{P_H}{P_L} = \left(\frac{A_H}{A_L} \frac{H^W}{L^W} \right)^{-(1-\alpha)/\sigma}. \quad (54)$$

This change in prices, for a given technology, makes skill-complement innovations more profitable and accelerates the creation of skill-complementary machines. Since, along the BG path, both types of innovations must be equally profitable and hence $\pi_H = \pi_L$, Equation (39) shows that this process continues until the relative price of goods has returned to the pre-trade level in the North. Substituting Equation (54) into (39) and imposing $\pi_H = \pi_L$, yields the new equilibrium skill bias of technology:

$$\frac{A_H}{A_L} = \frac{L^W}{H^W} \left[\frac{H^N}{L^N} \right]^\sigma. \quad (55)$$

Given that $H^N/L^N > H^W/L^W$, the new technology is more skill-biased and skilled workers in the North earn higher wages. The effect on the skill premium can be seen by substituting (55) into (42):

$$\frac{w_H}{w_L} = \left[\frac{H^N}{L^N} \frac{L^W}{H^W} \right] \left[\frac{H^N}{L^N} \right]^{\sigma-2}. \quad (56)$$

The effect of a move from autarky to free trade can be approximated by the elasticity of the skill premium to a change in L^W/H^W computed at $L^W/H^W = L^N/H^N$ (that is, starting from the pre-trade equilibrium). Equation (56) shows this elasticity to be unity.

1 Thus, if, for example, L^W/H^W were 4% higher than L^N/H^N , the model would predict 1
2 trade to raise the skill premium by the same 4%.³¹ 2

3 Without technical change, instead, the reaction of the skill premium to a change in 3
4 the perceived scarcity of factors due to trade depends on the degree of substitutability of 4
5 skilled and unskilled workers. From Equation (42), the elasticity of the skill premium to 5
6 a change in L/H would be $1/\sigma$, less than in the case of endogenous technology as long 6
7 as $\sigma > 1$, i.e., when skilled and unskilled workers are gross substitutes. Therefore, with 7
8 directed technical change and $\sigma > 1$, trade increases the skill premium in the North 8
9 by more than would otherwise be the case: for example, if the elasticity of substitution 9
10 is 2, the endogenous reaction of technical progress doubles the impact of trade on wage 10
11 inequality. 11

12 Note that another direct channel through which trade can affect factor prices in mod- 12
13 els of endogenous technical change is by affecting the reward to innovation. If trade 13
14 increases the reward to innovation (for example, through the scale effect) and the R&D 14
15 sector is skill-intensive relative to the rest of the economy, trade will naturally spur 15
16 wage inequality. This mechanism is studied by Dinopoulos and Segerstrom (1999) in a 16
17 quality-ladder growth model with no scale effects.³² 17

18 What are the implications of trade opening for cross-country income differences? 18
19 We have seen that trade induces a higher skill bias in technology; given the result of 19
20 Acemoglu and Zilibotti (2001) that the excessive skill-complementarity of Northern 20
21 technologies is a cause of low productivity in Southern countries, it may seem natural 21
22 to conclude that trade would then increase productivity differences. However, this con- 22
23 clusion would be premature. In the absence of any barriers, trade equalizes the price 23
24 of goods; given that the production functions adopted so far rule out complete special- 24
25 ization, this immediately implies that factor prices and sectoral productivity are also 25
26 equalized. This does not mean that trade equalizes income levels; because of their dif- 26
27 ferent skill-composition, the North and the South will still have differences in income 27
28 per capita, but nothing general can be said.³³ 28

29 The fact that trade generates productivity convergence crucially depends on factor 29
30 prices being equalized by trade. Since factor price equalization is a poor approximation 30
31 of reality, it is worth exploring the implications of models with endogenous technolo- 31
32 gies when this property does not hold. A simple way of doing this is to add Ricardian 32
33 productivity differences, so that trade opening leads to complete specialization. In this 33
34 34

35 ³¹ Borjas, Freeman and Katz (1997) show that 4% is a plausible estimate of the increase in the unskilled labor 35
36 content of US trade with LDCs between 1980 and 1995. Therefore, this simple exercise may give a sense of 36
37 how much of the roughly 20% increase in the US skill premium in the same period can be attributed to trade. 37

38 ³² Recently, other papers have suggested that trade between identical countries may as well increase skill 38
39 premia through its effect on technology. See, for example, Epifani and Gancia (2002), Neary (2003) and 39
40 Thoenig and Verdier (2003). 40

41 ³³ A general result is that the endogenous response of technology makes trade less beneficial for LDCs than 41
42 would otherwise be the case. This occurs because, after trade opening, the skill premium rises as a result of the 42
43 induced skill-biased technical change. Given that the North is more skilled-labor abundant, it proportionally 43
44 benefits more from a higher skill premium. 43

1 case, the endogenous response of technology to weak IPRs in LDCs becomes a force 1
2 promoting productivity divergence.³⁴ Further, trade with countries providing weak pro- 2
3 tection for IPRs may have an adverse effect on the growth rate of the world economy. 3
4 These results, shown by Gancia (2003), can be obtained by modifying Acemoglu and 4
5 Zilibotti (2001) as follows. First, we allow the elasticity of substitution between final 5
6 goods to be larger than one: $Y = [\int_0^1 y_i^{(\varepsilon-1)/\varepsilon} di]^{\varepsilon/(\varepsilon-1)}$, with $\varepsilon > 1$. Then, we assume 6
7 that each good y_i can be produced by competitive firms both in the North and the South, 7
8 using sector-specific intermediates and labor: 8

$$9 \quad y_i = [(1-i)l_i^S]^{1-\alpha} \int_0^{A_i} (x_{i,v}^S)^\alpha dv + [il_i^N]^{1-\alpha} \int_0^{A_i} (x_{i,v}^N)^\alpha dv. \quad (57)$$

10 There are three important differences with respect to (45). First, $(1-i)$ and i now cap- 10
11 ture Ricardian productivity differences between the North and South, implying that the 11
12 North is relatively more productive in high index sectors. Second, intermediate goods 12
13 are sector specific, not factor specific (there is now a continuum $[0, 1]$ of technologies, 13
14 not only two). Third, there is only one type of labor. Given that the endogenous com- 14
15 ponent of technology (A_i) is still assumed to be common across countries, the sectoral 15
16 North–South productivity ratio only depends on the Ricardian elements. The new im- 16
17 plication is that countries specialize completely under free trade, as each good is only 17
18 manufactured in the location where it can be produced at a lower cost. 18
19

20 The equilibrium can be represented by the intersection of two curves, as in Dornbusch 20
21 et al. (1977). For any relative wage, the first curve gives the range $[0, J]$ of goods effi- 21
22 ciently produced in the South: $\frac{J}{1-J} = \frac{w^N}{w^S}$. The second curve combines trade balance 22
23 and a BG research arbitrage condition, requiring profits to be equalized across sectors 23
24 and countries. To find this, the model assumes that the owner of a patent can only ex- 24
25 tract a fraction $\lambda < 1$ of the profits generated by its innovation in the South, so that λ 25
26 can be interpreted as an index of the strength of international IPRs protection. The trade 26
27 balance plus the research arbitrage condition turn out to be [see Gancia (2003)]: 27
28

$$29 \quad \frac{w^N}{w^S} = \lambda^{-\tilde{\sigma}} \left[\frac{L^S \int_J^1 (i)^{\tilde{\sigma}/(1-\tilde{\sigma})} di}{L^N \int_0^J (1-i)^{\tilde{\sigma}/(1-\tilde{\sigma})} di} \right]^{1-\tilde{\sigma}}, \quad (58)$$

30 with $\tilde{\sigma} \equiv (1-\alpha)(\varepsilon-1) \in (0, 1)$.³⁵ As long as $\tilde{\sigma} > 0$ (i.e., $\varepsilon > 1$), the wage gap is 30
31 decreasing in the degree of protection of IPRs in the South, λ . The reason is that weaker 31
32 protection of IPRs shifts innovations out of Southern sectors and increases the relative 32
33 productivity of the North. From the condition $\frac{J}{1-J} = \frac{w^N}{w^S}$, it is easily seen that a weaker 33
34

34 The idea that trade may magnify cross-country inequality was put forward by several economists. Some 40
41 examples are Stiglitz (1970), Young (1991), Krugman and Venables (1995), Matsuyama (1996), Rodriguez- 41
42 Clare (1996) and Ventura (1997).

42 ³⁵ $\tilde{\sigma} < 1$ guarantees balanced growth across sectors. $\tilde{\sigma} > 0$, i.e., an elasticity of substitution between goods 42
43 greater than one, rules out immiserizing growth. 43

1 protection of IPRs in the South, by raising w^N/w^S , is accompanied by a reduction in 1
2 sectors $[1 - J]$ located in the North, because higher wages make the North less com- 2
3 petitive. A second result emerges by calculating the growth rate of the world economy. 3
4 In particular, Gancia (2003) shows the growth rate of the world economy to fall with λ 4
5 and approach zero if λ is sufficiently low. The reason is that a lower λ shifts innovation 5
6 towards Northern sectors and, at the same time, induces the relocation of more sectors 6
7 to the South, where production costs become lower. This, in turn, implies that a wider 7
8 range of goods becomes subject to weak IPRs and hence, to a low innovation incentive. 8
9

11 5. Complementarity in innovation 11

12
13 In the models described so far, innovation has no effect on the profitability of existing 13
14 intermediate firms. This is a knife-edge property which descends from the specifica- 14
15 tion of the final production technology, (3). In general, however, new technologies can 15
16 substitute or complement existing technologies. 16

17 Innovation often causes technological obsolescence of previous technologies. Substi- 17
18 tution is emphasized, in an extreme fashion, by Schumpeterian models such as Aghion 18
19 and Howitt (1992). In such models, innovation provides “better of the same”, i.e., more 19
20 efficient versions of the pre-existing inputs. Growth is led by a process of creative de- 20
21 struction, whereby innovations do not only generate but also destroy rents over time. 21
22 This has interesting implications for dynamics: the expectation of future innovations 22
23 discourages current innovation, since today’s innovators expect a short life of their rents 23
24 due to rapid obsolescence. More generally, substitution causes a decline in the value 24
25 of intermediate firms over time, at a speed depending on the rate of innovation in the 25
26 economy. 26

27 There are instances, however, where new technologies complement rather than sub- 27
28 stitute old technologies. The market for a particular technology is often small at the 28
29 moment of its first introduction. This limits the cash-flow of innovating firms, which 29
30 initially pose little threat to more established technologies. However, the development 30
31 of new compatible applications expands the market for successful new technologies 31
32 over time, thereby increasing the profits earned by their producers. Rosenberg (1976) 32
33 discusses a number of historical examples, where such complementarities were impor- 33
34 tant. A classical example is the steam engine. This had been invented in the early part 34
35 of the XVIIIth Century, but its diffusion remained very sporadic before a number of 35
36 complementary innovations (e.g., Watt’s separate condenser) made it competitive with 36
37 the waterwheels, which remained widespread until late in the XIXth Century. 37

38 Complementarity in innovation raises interesting issues concerning the enforcement 38
39 and design of intellectual property rights. For instance, what division of the surplus be- 39
40 tween basic and secondary innovation maximizes social welfare? This issue is addressed 40
41 by Scotchmer and Green (1995) who construct a model where innovations are sequen- 41
42 tially introduced, and the profits of major innovators can be undermined by subsequent 42
43 derivative innovations. In this case, the threat of derivative innovations can reduce the 43

1 incentive for firms to invest in major improvements in the first place. However, too 1
2 strong a defense of the property right of basic innovators may reduce the incentive to 2
3 invest in socially valuable derivative innovations. Scotchmer and Green (1995) show 3
4 that the optimal policy in fact consists of a combination of finite breath and length of 4
5 patents. Scotchmer (1996) instead argues that it may be optimal to deny patentability 5
6 to derivative innovations, instead allowing derivative innovations to be developed under 6
7 licensing agreements with the owner of the basic technology. More recently, Bessen and 7
8 Maskin (2002) show that when there is sufficient complementarity between innovations 8
9 (as in the case of the software industry), weak patent laws may be conducive to more 9
10 innovation than strong patent laws. The reason is that while the incumbent's current 10
11 profit is increased by strong patent laws, its prospect of developing future profitable 11
12 innovation is reduced when patent laws inhibit complementary innovations. 12

13 While this literature focuses on the partial equilibrium analysis of single industries, 13
14 complementarity in innovation also has implications on broader development questions. 14
15 Multiple equilibria originating from coordination failures [of the type emphasized, in 15
16 different contexts, by Murphy, Shleifer and Vishny (1989), and Cooper and John (1988)] 16
17 can arise when there is complementarity in innovation. Countries can get locked-in into 17
18 an equilibrium with no technology adoption, and temporary big-push policies targeting 18
19 incentives to adopt new technologies may turn out to be useful.³⁶ One such example is 19
20 Ciccone and Matsuyama (1996). In their model, multiple equilibria and poverty traps 20
21 may arise from the two-way causality between the market size of each intermediate 21
22 good and their variety: when the availability of intermediates is limited, final good 22
23 producers are forced to use a labor intensive technology which, in turn, reduces the 23
24 incentive to introduce new intermediates. 24

25 Young (1993) constructs a model where innovation expands the variety of both in- 25
26 termediate and final goods. New intermediate inputs are not used by mature final in- 26
27 dustries, and their market is initially thin. The expansion of the market for technologies 27
28 over time creates complementarity in innovation. The details of this model are dis- 28
29 cussed in the remainder of this section. To this aim, we augment the benchmark model 29
30 of Section 2 with the endogenous expansion in the variety of final goods.³⁷ Over time, 30
31 innovative investments make new intermediate inputs available to final producers, as in 31
32 Romer's model. However, as a by-product (spillover), they also generate an equivalent 32
33 expansion of the set of final goods that can be produced. There are no property rights 33
34 defined on the production of new final goods, and these are produced by competitive 34
35 firms extraneous to the innovation process. 35
36 36
37 37
38 38

39 ³⁶ Interestingly, in models with complementarity in innovation, market economies may be stuck in no-growth 39
40 traps that are inefficient in the sense that the optimal intertemporal allocation would require positive invest- 40
41 ment and growth. See, for example, Ciccone and Matsuyama (1999). 41

42 ³⁷ Models featuring an expanding variety of final products include Judd (1985), Grossman and Helpman 42
43 (1989, Chapter 3) and, more recently, Xie (1998) and Funke and Strulik (2000). Here, we follow Young 43
(1993) which, in turn, is close to Judd's paper. 43

A_t will now denote the measure of both final goods and intermediate goods available in the economy at t . Final products are imperfect substitutes in consumption, and the instantaneous utility function is:³⁸

$$V_t = \int_0^{A_t} \ln(C_{s,t}) ds,$$

with total utility being

$$U = \int_0^\infty e^{-\rho t} V_t dt.$$

This specification implies that consumers' needs grow as new goods become available. Suppose, for instance, that a measure ε of new goods is introduced between time t and $t + j$. At time t , consumers are satisfied with not consuming the varieties yet to be invented. However, at time $t + j$, the same consumers' utility would fall to minus infinity if they did not consume the new goods.

The productive technology for the s 'th final good is given by

$$Q_s = \left(\int_0^{\min[s\Theta, A]} x_{j,s}^\alpha dj \right)^{1/\alpha}, \quad (59)$$

where $\Theta \geq 1$ is a parameter. Note that labor is not used in the final goods production. First, to build the intuition in the simplest case, we maintain that all final goods are produced with the same technology employing all available varieties of intermediate inputs. More formally, we characterize the equilibrium in the limit case where $\Theta \rightarrow \infty$, so that $\min[s\Theta, A] = A$. This assumption will be relaxed later.

We use the "labor-for-intermediates" model introduced in Section 2.2, where labor is used for research and intermediate production and the productivity of labor in intermediate production equals A_t . We choose the nominal wage as the numeraire.³⁹ Hence, the profit of an intermediate producer can be expressed as:

$$\pi = \frac{1 - \alpha}{\alpha} \frac{x}{A} = \frac{1 - \alpha}{\alpha} \frac{L - L_x}{A}. \quad (60)$$

Note that profits fall over time at the rate at which knowledge grows. In a BG equilibrium, the interest rate is constant and A grows at the constant rate γ . Free entry implies:

$$\int_t^\infty e^{-r\tau} \pi_\tau d\tau = \frac{1 - \alpha}{\alpha} \frac{L - \gamma/\delta}{A_t(r + \gamma)} \leq \frac{1}{\delta A_t},$$

³⁸ This is the benchmark specification in Young (1993), where it is then extended to general CES preferences across goods. The logarithmic specification is analytically convenient because of the property that consumers spend an equal income share on all existing goods.

³⁹ Note that we cannot simply set the price of the final good as the numeraire, as there is an increasing variety of final goods.

1 where we have used the fact that $\gamma = \delta L_x$. Simplifying terms yields 1

$$2 \frac{1 - \alpha \delta L - \gamma}{\alpha r + \gamma} \leq 1 \quad (61) \quad 3$$

4 The intertemporal optimality condition for consumption also differs from the bench- 5
6 mark model. In particular, if E denotes the total expenditure in final goods, the Euler 6
7 condition is:⁴⁰ 7

$$8 \frac{\dot{E}}{E} = r - \rho + \frac{\dot{N}_t}{N_t}. \quad 8$$

9 In a BG equilibrium, the total expenditure on consumption goods is constant. Hence, 9
10 10

$$11 r = \rho - \gamma. \quad (62) \quad 11$$

12 This expression can be substituted into (61) to give a unique solution for γ . As long as 12
13 growth is positive, we have 13

$$14 \gamma = \delta L - \frac{\alpha \rho}{1 - \alpha}, \quad 14$$

15 which is almost identical to (15), except for the constant term $\alpha/(1 - \alpha)$ being replaced 15
16 by $1/\alpha$. In the limit case considered so far ($\Theta \rightarrow \infty$), the model is isomorphic to Romer 16
17 (1990). 17
18 18

19 Next, we move to the general case where $\Theta \geq 1$ is finite. This implies that final pro- 19
20 ducers cannot use the entire range of intermediate goods. In particular, an intermediate 20
21 good indexed by s cannot be used by “mature” final industries having an index j , such 21
22 that $j < s/\Theta$. This assumption captures the idea that a technology mismatch develops 22
23 over time between mature final good industries and new technologies.⁴¹ 23
24 24

25 An important implication of this assumption is that, when introduced, a new technol- 25
26 ogy (intermediate input) is only required by a limited number of final industries. Thus, 26
27 the monopolist producing a new variety has a small cash-flow. This is especially true 27
28 when the parameter Θ is small: as $\Theta \rightarrow 1$, there is no demand for a new intermediate 28
29 good at the time of its first appearance. However, the market for technologies expands 29
30 over time, as new final goods using “modern” technologies appear. This dynamic mar- 30
31 ket size effect generates complementarity in the innovation process. An innovator is 31
32 eager to see rapid technical progress, as this expands the number of users of the new 32
33 technology. 33
34 34

35 Countering this effect, there is a process of “expenditure diversion” that reduces, *ce-* 35
36 *teris paribus*, the demand for each intermediate good. Over time, technical progress 36
37 expands the number of intermediate inputs over which final producers spread their 37
38 38

39 39
40 ⁴⁰ See Young (1993, p. 783) for the derivation of this Euler equation. 40

41 ⁴¹ In principle, it would seem natural to assume that new final goods do not use very old intermediate goods. 41
42 Young (1993, p. 780) argues that allowing for this possibility would not change the main results, but would 42
43 make the analysis more involved. 43

1 demand. As noted above, the total expenditure on final goods is constant in a BG equi- 1
2 librium. Since final good firms make zero profits at all times, and intermediates are 2
3 the only inputs, the total expenditure on the intermediate goods must also be constant. 3
4 Therefore, an increase in A dilutes the expenditure over a larger mass of intermediate 4
5 goods, and reduces the profit of each existing intermediate firm. This effect generates 5
6 substitution rather than complementarity in innovation. 6

7 The dynamic market size effect may dominate for young intermediate firms. But as 7
8 a technology becomes more mature, the expenditure dilution effect takes over. Thus, 8
9 firms can go through a life-cycle: their profit flow increases over time at an earlier stage 9
10 and decreases at a later stage. 10

11 We denote by $\pi(A_\tau, A_t)$ the profit realized at time τ by an intermediate producer 11
12 who entered the market in period $t < \tau$. Solving the profit maximization problem for 12
13 the intermediate monopolist, subject to the demand from final industries, leads to the 13
14 following expression: 14

$$15 \pi(A_\tau, A_t) = \frac{1 - \alpha}{\alpha A_\tau} (L - L_x) \left(1 + \frac{\gamma(\tau - t) - 1}{\Theta} \right). \quad (63) \quad 16$$

17 It is easily verified that as $\Theta \rightarrow \infty$, the solution becomes identical to (60), where 17
18 nominal profits fall at the same rate as A_t . 18

19 Free-entry implies: 19

$$20 \int_t^\infty e^{-r(\tau-t)} \pi(A_\tau, A_t) d\tau \leq \frac{1}{\delta A_t}. \quad (64) \quad 21$$

22 Solving the integral on the left-hand side, using the Euler condition, $r + \gamma = \rho$, and 22
23 simplifying terms yields the following equilibrium condition: 23

$$24 f_{FE}(\gamma) = \frac{1 - \alpha}{\alpha \rho^2 \Theta} (\delta L - \gamma)(\gamma + \rho(\Theta - 1)) \leq 1, \quad (65) \quad 25$$

26 where all terms but γ are parameters. For sufficiently large values of Θ , i.e., when 26
27 the market for new technology is large, $f'_{FE}(\gamma) < 0$ and the equilibrium is unique. 27
28 However, if $\Theta < 1 + \delta L/\rho$, $f_{FE}(\gamma)$ is non-monotonic, and multiple equilibria are 28
29 possible. 29

30 Figure 2 describes the three possible cases. As long as $\rho > \gamma$, which is a necessary 30
31 and sufficient condition for the interest rate to be positive, $f_{FE}(\gamma)$ is increasing in Θ . 31
32 For a range of small Θ 's, there is no equilibrium with positive innovation (lower curve). 32
33 The only equilibrium is a point such as X' , featuring zero growth. 33

34 For an intermediate range of Θ , we have $f_{FE}(\gamma) = 1$ in correspondence of two 34
35 values of γ (intermediate curve). This implies that (for generic economies), there exist 35
36 three equilibria, where equilibria such as point X feature zero innovation and growth. 36
37 Firms contemplating entry expect no expansion of the market size for new technologies. 37
38 Furthermore, such market size is too small to warrant profitable deviations, and the 38
39 expectation of no innovation is fulfilled in equilibrium. Equilibria such as point Y are 39
40 41
42
43

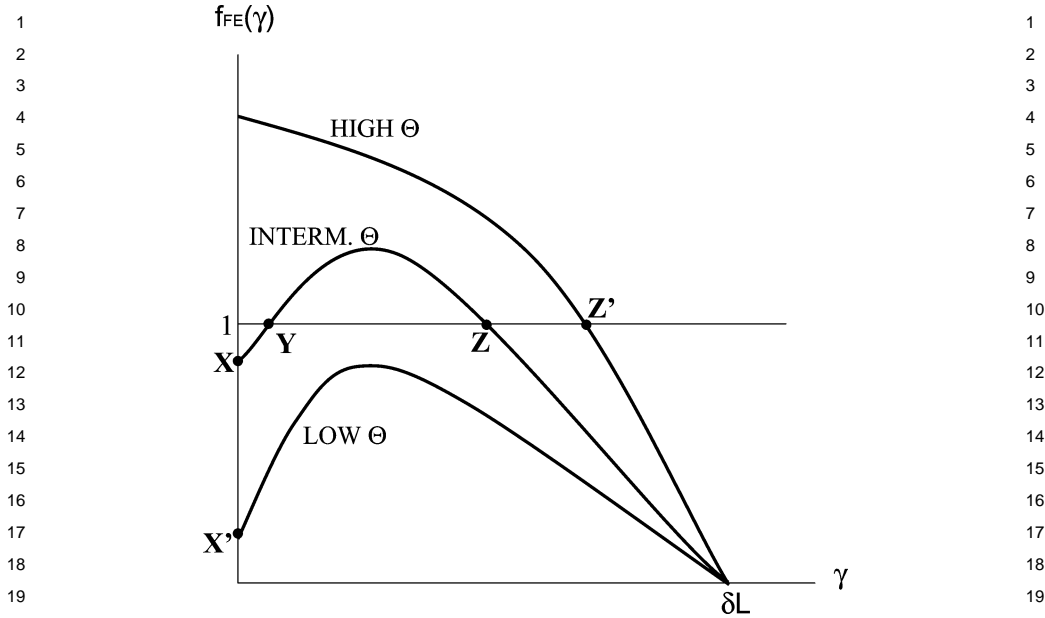


Figure 2.

characterized by local complementarity in innovation: the expectation of higher future innovation and growth increases the value of new firms, stimulating current entry and innovation. In steady-state (BG), this implies a positive slope of the locus $f_{FE}(\gamma)$.⁴² Eventually, for sufficiently high growth rates, the diversion effect dominates. Thus, in an equilibrium like Z , the value of innovating firms depends negatively on the speed of innovation.⁴³

Finally, for a range of large Θ 's, substitution dominates throughout (upper curve). The initial market for new technologies is sufficiently large to make the expenditure diversion effect dominate the market size effect, even at low growth rates. The equilibrium is unique, and the solution is isomorphic to that of the benchmark model of expanding variety.

⁴² As mentioned above, firms go through a life-cycle here. When a new technology is introduced, the profit flow of an innovating firm is small. As time goes by, the expenditure diversion effect becomes relatively more important. The value of a firm upon entry is the PDV of its profit stream. Local complementarity occurs if, for a particular γ , profits increase at a sufficiently steep rate in the earlier part of the firm's life-cycle.

⁴³ If the expectational stability of the equilibria in the sense of Evans and Honkapohja (2001) is tested, equilibria such as point Y are not found to be E-stable, while equilibria such as X and Z are stable. See the discussion in Section 7.2.

6. Financial development

A natural way in which the expansion of the variety of industries can generate complementarities in the growth process is through its effects on financial markets. Acemoglu and Zilibotti (1997) construct a model where the introduction of new securities, associated with the development of new intermediate industries, improves the diversification opportunities available to investors. Investors react by supplying more funds, which fosters further industrial and financial development, generating a feedback.⁴⁴ The model offers a theory of development. At early stages of development, a limited number of intermediate industries are active (due to technological nonconvexities), which limits the degree of risk-spreading that the economy can achieve. To avoid highly risky investments, agents choose inferior but safer technologies. The inability to diversify idiosyncratic risks introduces a large amount of uncertainty in the growth process. In equilibrium, development proceeds in stages. First, there is a period of “primitive accumulation” with a highly variable output, followed by take-off and financial deepening and finally, steady growth. Multiple equilibria and poverty traps are possible in a generalized version of the model.

The theory can explain why the growth process is both slow and highly volatile at early stages of development, and stabilizes as an economy grows richer. Evidence of this pattern can be found in the accounts of pre-industrial growth given by a number of historians, such as Braudel (1979), North and Thomas (1973) and DeVries (1990). For instance, in cities such as Florence, Genoa and Amsterdam, prolonged periods of prosperity and growth have come to an end after episodes of financial crises. Interestingly, these large set-backs were not followed (as a neoclassical growth model would instead predict) by a fast recovery but, rather, by long periods of stagnation. Similar phenomena are observed in the contemporary world. Acemoglu and Zilibotti (1997) document robust evidence of increases in GDP per capita being associated with large decreases in the volatility of the growth process. It has also been documented that higher volatility in GDP is associated with lower growth [Ramey and Ramey (1995)].

We here describe a simplified version of the model. Time is discrete. The economy is populated by overlapping generations of two-period lived households. The population is constant, and each cohort has a unit mass ($L = 1$). There is uncertainty in the economy, which we represent by a continuum of equally likely states $s \in [0, 1]$. Agents are assumed to consume only in the second period of their lives.⁴⁵ Their preferences are parameterized by the following (expected) utility function, inducing unit relative risk

⁴⁴ This paper is part of a recent literature on the two-way relationship between financial development and growth. This includes Bencivenga and Smith (1991), Greenwood and Jovanovic (1990) and Zilibotti (1994). In none of these other papers does financial development take the form of an expansion in the “variety” of assets.

⁴⁵ This is for simplicity. Acemoglu and Zilibotti (1997) assume that agents consume in both periods. It is also possible to study the case of a general CRRA utility function.

1 aversion:

$$2 \quad E_t U(c_{t+1}) = \int_0^1 \log(c_{t+1}^s) ds. \quad (66) \quad 3$$

4
5 The production side of the economy consists of a unique final good sector, and a
6 continuum of intermediate industries. The final good sector uses intermediate inputs
7 and labor to produce final output. Output in state s is given by the following production
8 function:

$$9 \quad Y_{s,t} = (x_{s,t-1} + x_{\phi,s,t-1})^\alpha L^{1-\alpha}. \quad (67) \quad 10$$

11 The term in brackets is “capital”, and it is either produced by a continuum of inter-
12 mediate industries, each producing some state-contingent amount of output (x_s), or a
13 separate sector using a “safe technology” (x_ϕ). The measure of the industries with a
14 state-contingent production, A_t , is determined in equilibrium, and A_t can expand over
15 time, like in Romer’s model, but it can also fall. Moreover, $A_t \in [0, 1]$, i.e., the set of
16 inputs is bounded.

17 In their youth, agents work in the final sector and earn a competitive wage, $w_{s,t} =$
18 $(1 - \alpha)Y_{s,t}$. At the end of this period, they take portfolio decisions: they can place their
19 savings in a set of risky securities ($\{F_i\}_{i \in [0, A_t]}$), consisting of state-contingent claims to
20 the output of the intermediate industries or, in a safe asset (ϕ), consisting of claims to the
21 output of the safe technology. After the investment decisions, the uncertainty unravels,
22 the security yields its return and the amount of capital brought forward to the next period
23 is determined. The capital is then sold to final sector firms and fully depreciates after
24 use. Old agents consume their capital income and die.

25 Intermediate industries use final output for production. An intermediate industry $i \in$
26 $[0, A_t]$ is assumed to produce a positive output only if state $s = i$ occurs. In all other
27 states of nature, the firm is not productive. Moreover, the i th industry is only productive
28 if it uses a minimum amount of final output, M_i , where

$$29 \quad M_i = \max \left\{ 0, \frac{D}{(1-x)}(i-x) \right\}, \quad 30$$

31
32 with $x \in (0, 1)$. This implies that some intermediate industries require a certain mini-
33 mum size, M_i , before being productive. In particular, industries $i \leq x$ have no minimum
34 size requirement, and for the rest of the industries, the minimum size requirement in-
35 creases linearly with the index i .

36 To summarize, the intermediate technology is described by the following production
37 function:

$$38 \quad x_{i,s} = \begin{cases} RF_i & \text{if } i = s \text{ and } F_i \geq M_i, \\ 0 & \text{otherwise.} \end{cases} \quad 39$$

40 Since there are no start-up costs, all markets are competitive. Thus, firms retain no
41 profits, and the product is entirely distributed to the holders of the securities. The j th
42 security entitles its owner to a claim to R units of capital in state j (as long as the mini-
43 mum size constraint is satisfied, which is always the case in equilibrium), and otherwise

1 to nothing. Savings invested in the “safe technology” give the return

$$2 \quad x_{\phi,s} = r\phi, \quad \forall s \in [0, 1], \quad 3$$

4 where $r < R$. Thus, one unit of the safe asset is a claim to r units of capital in all states
5 of nature.

6 Since the risky securities yield symmetric returns, and there is safety in numbers, it
7 is optimal for risk-averse agents to hold a portfolio containing all available securities in
8 equal amounts. More formally, the optimal portfolio decision features $F_i = F$, for all
9 $i \in [0, A_t]$. We refer to this portfolio consisting of an equal amount of all traded risky
10 securities as a *balanced portfolio*.

11 If $A_t = 1$, a balanced portfolio of risky securities bears no risk, and first-order
12 dominates the safe investment. However, due to the presence of technological non-
13 convexities (minimum size requirements), not all industries are in general activated.
14 When $A_t < 1$, the inferior technology is safer, and there is a trade-off between risk and
15 productivity. In this case, the optimal investment decision of the representative saver
16 can be written as:

$$17 \quad \max_{\phi_t, F_t} A_t \log[\rho_{G,t+1}(RF_t + r\phi_t)] + (1 - A_t) \log[\rho_{B,t+1}(r\phi_t)], \quad 18 \quad (68)$$

19 subject to

$$20 \quad \phi_t + A_t F_t \leq w_t. \quad 21 \quad (69)$$

22 $\rho_{\bar{s},t+1}$ denotes the rate of return of capital, which is taken as parametric by agents,
23 and does not affect the solution of the program.⁴⁶ Agents also take A_t , i.e., the set of
24 securities offered, as parametric.

25 Simple maximization yields:

$$26 \quad \phi_t^* = \frac{(1 - A_t)R}{R - rA_t} w_t, \quad 27 \quad (70)$$

$$28 \quad F_{i,t}^* = \begin{cases} F(A_t) \equiv \frac{R-r}{R-rA_t} w_t, & \forall i \leq A_t, \\ 0 & \forall i > A_t. \end{cases} \quad 29 \quad (71)$$

30 Figure 3 expresses the demand for each risky asset, $F(A_t)$ (FF schedule), as a func-
31 tion of the measure of intermediate industries which are active. The FF schedule is

32
33
34
35
36 ⁴⁶ In equilibrium:

$$37 \quad \rho_{G,t+1} = \alpha(RF_t + r\phi_t)^{\alpha-1} \quad 38$$

39 and

$$40 \quad \rho_{B,t+1} = \alpha(r\phi_t)^{\alpha-1}. \quad 41$$

42 $\rho_{G,t+1}$ applies in the “good state”, i.e., when the realized state is $i \leq A_t$, while $\rho_{B,t+1}$ is the marginal
43 product of capital in the “bad” state, when the realized state is $i > A_t$ and no risky investment pays off.

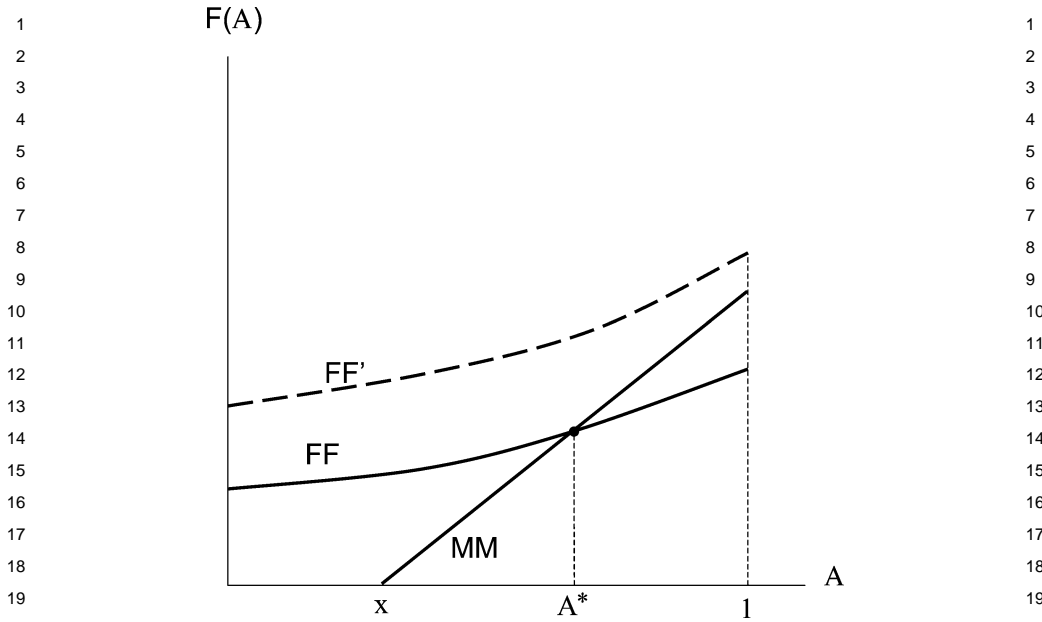


Figure 3.

upward sloping, implying that there is complementarity in the demand for risky assets: the demand for *each asset* grows with the variety of intermediate industries.

Complementarity arises because the more active are intermediate industries, the better is risk-diversification. Thus, as A_t increases, savers shift their investments away of the safe asset into high-productivity risky projects (the “stock market”). Such complementarity hinges on risk aversion being sufficiently high.⁴⁷ In general, similar to Young (1993), an increase in A creates two effects. On the one hand, investments in the stock market become safer because of better diversification opportunities, which induces complementarity. On the other hand, investments are spread over a larger number of assets, inducing substitution. With sufficiently high risk aversion, including the unit CRRA specification upon which we focus, the first effect dominates.

The equilibrium measure of active industries, A_t^* , is determined (as long as $A^* < 1$) by the following condition:

$$F(A_t^*) = M_{A_t^*}.$$

⁴⁷ Suppose agents were risk averse, but only moderately so. Suppose, in particular, that they were so little risk-averse that they would decide not to hold any safe asset in their portfolio. Then, an expansion in the set of risky securities would induce agents to spread their savings (whose total amount is predetermined) over a larger number of assets. In this case, assets would be substitutes rather than complements.

1 In Figure 3, the equilibrium is given by the intersection of schedules FF and MM, 1
2 where the latter represents the distribution of minimum size requirements across 2
3 industries. Intuitively, A_t^* is the largest number of industries for which the technological 3
4 non-convexity can be overcome, subject to the demand of securities being given by 4
5 (71).⁴⁸ 5

6 Growth increases wage income and the stock of savings over time. In equilibrium, 6
7 this induces an expansion of the intermediate industries, A_t^* . This can once more be 7
8 seen in Figure 3: growth creates an upward shift of the FF schedule, causing the equilib- 8
9 rium to move to the left. Therefore, growth triggers financial development. In particular, 9
10 when the stock of savings becomes sufficiently large, the financial market is sufficiently 10
11 thick to allow all industries to be active. In the case described by the dashed curve, 11
12 FF', the economy is sufficiently rich to afford $A_t^* = 1$. The inferior safe technology is 12
13 then abandoned. Financial development, speeds up growth by channelling investments 13
14 towards the more productive technology. 14

15 The stochastic equilibrium dynamics of GDP can be explicitly derived: 15

$$16 Y_{t+1} = \begin{cases} F_B(Y_t) = \left((1 - \alpha) \frac{r(1 - A_t^*)}{R - rA_t^*} RY_t \right)^\alpha & \text{prob. } 1 - A_t^*, \\ F_G(Y_t) = ((1 - \alpha)RY_t)^\alpha & \text{prob. } A_t^*, \end{cases} \quad (72) \quad 17$$

18 where $A_t^* = A(Y_{t-1}) \leq 1$ is the equilibrium measure of intermediate industries, such 18
19 that $A' \geq 0$.⁴⁹ The first line corresponds to the case of a “bad realization” at time t , 19
20 such that $s \in (A_t^*, 1]$. In this case, none of the active intermediate industries turned out 20
21 to pay-off at time t , and capital at time $t + 1$ is only given by the return of the safe 21
22 technology. The second line corresponds to the case of a “good realization” at t , such 22
23 that $s \in [0, A_t^*]$. In this case, the risky investment paid off at time t , and capital and 23
24 output are relatively large at time $t + 1$. Note that the probability of a good realization 24
25 increases with the level of development, since $A'(Y_{t-1}) \geq 0$ (with strict inequality as 25
26 long as $A^* < 1$). 26

27 Figure 4 describes the dynamics. The two schedules represent output at time $t + 1$ 27
28 as a function of output at time t conditional on good news ($F_G(Y_t)$) and bad news 28
29 ($F_B(Y_t)$), respectively. At low levels of capital ($Y \leq Y_L$), the marginal product of capital 29
30 is very high, which guarantees that growth is positive, even conditional on bad news. 30
31 In the intermediate range where $Y \in [Y_L, Y_M]$, growth only occurs if news is good, 31
32 since $F_B(Y_t) < Y_t < F_G(Y_t)$. The threshold Y_L is not a steady-state; however, it is a 32
33 point around which the economy will spend some time. When the initial output is below 33
34 Y_L , the economy necessarily grows towards it. When it is above Y_L , output falls back 34
35 whenever bad news occurs. So, in this region, the economy is still exposed to undiversified 35
36 risks, and experiences fluctuations and set-backs. Finally, for $Y \geq Y_M$, there are 36
37 37
38 38
39 39

40 ⁴⁸ Acemoglu and Zilibotti (1997) show the laissez-faire portfolio investment to be inefficient. Efficiency 40
41 would require more funds to be directed to industries with large non-convexities, i.e., agents not holding a 41
42 balanced portfolio. The inefficiency is robust to the introduction of a rich set of financial institutions. 42

43 ⁴⁹ Acemoglu and Zilibotti (1997) derive a closed-form solution for A_t^* that we do not report here. 43

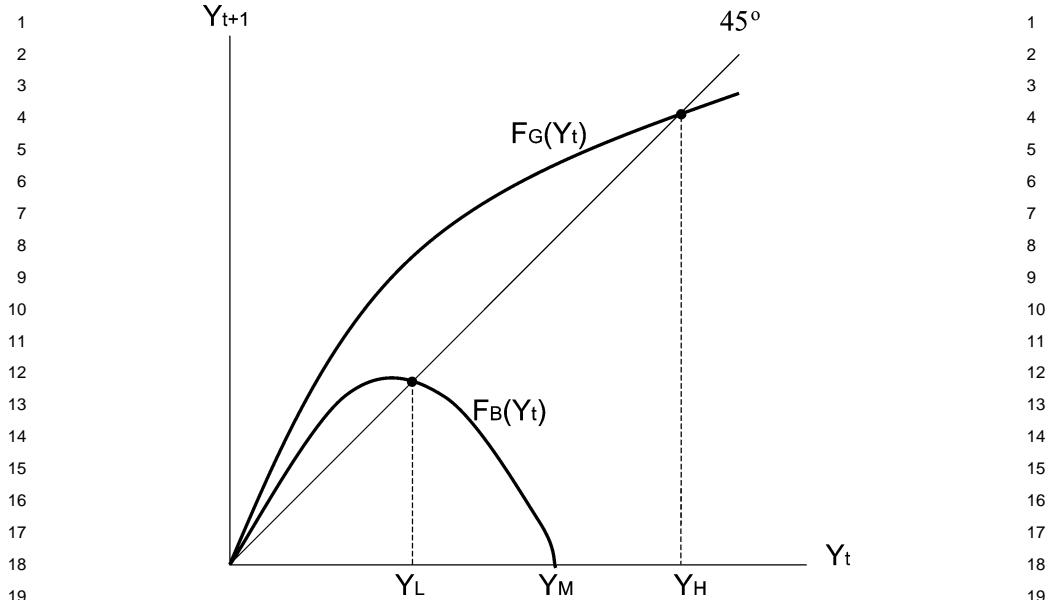


Figure 4.

enough savings in the economy to overcome all technological non-convexities. When the economy enters this region, all idiosyncratic risks are removed, and the economy deterministically converges to Y_H .⁵⁰

Note that it may appear as if, in the initial stage, countries striving to take off do not grow at a sustained rate during long periods. The demand for insurance takes the form of investments in low-productivity technologies, and poor economies tend to have low total factor productivity and slow growth.

In the case described by Figure 4, the economies “almost surely” converge to a unique steady-state. Different specifications of the model can, however, lead to less optimistic predictions. With higher risk aversion, for instance, traps can emerge, as in the example described in Figure 5. An economy starting with a GDP in the region $[0, Y_{MM})$ would never attain the high steady-state Y_H , and would instead perpetually wander in the trapping region $[0, Y_{LL}]$. Conversely, an economy starting above Y_M would certainly converge to the high steady-state, Y_H . Finally, the long-run fate of an economy starting in the region $[Y_{MM}, Y_M]$ would be determined by luck: an initial set of positive

⁵⁰ That the economy converges “almost surely” to a steady-state where all risk is diversified away only occurs under parameter restrictions ensuring that $Y^{SS} > Y^1$. Although the model presented here is neoclassical and features zero growth in the long run, it is possible to augment it with spillover of the learning-by-doing type, as in Romer (1986), and make it generate self-sustained growth.

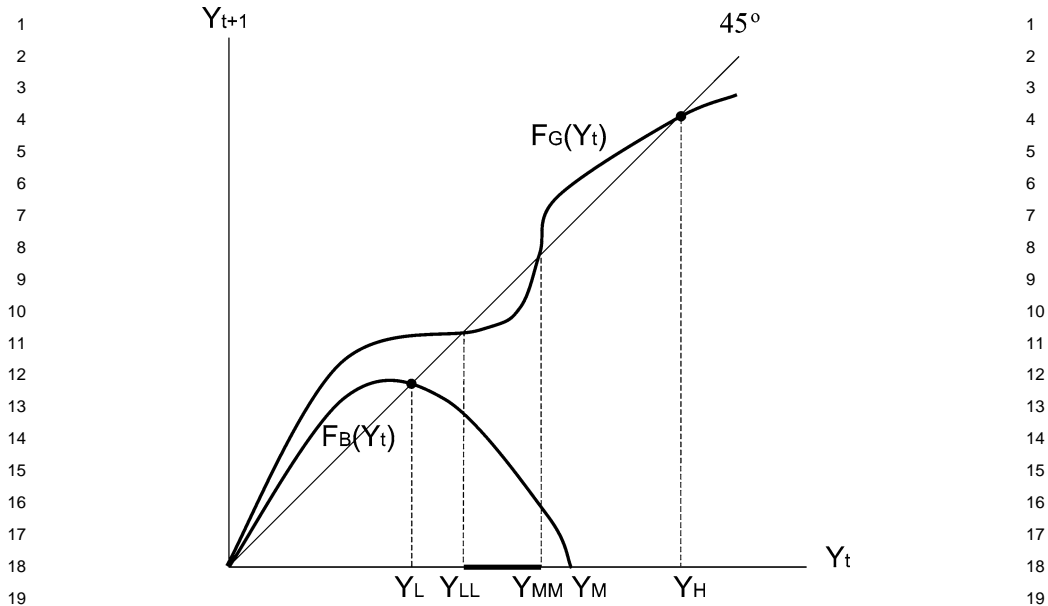


Figure 5.

draws would bring this economy into the basin of attraction of the good equilibrium. A single set-back, however, would forever jeopardize its future development.⁵¹

The model can be extended in a number of directions. A two-country extension shows that international capital flows may lead to divergence, rather than convergence between economies. This result is due to the interplay between two forces: first, decreasing returns to capital would tend to direct foreign investments towards poorer countries, as in standard neoclassical models. Second, the desire to achieve better diversification pushes investments towards thicker markets. The latter force tends to prevail at some earlier stages of the development process. So, poor countries suffer an outflow of capital, which spills over to lower income and wages for the next generation, thereby slowing down the growth process. The analysis of capital flows, financial integration and financial crises in the context of similar models is further developed in recent papers by Martin and Rey (2000, 2001 and 2002). A different extension of the model is pursued by Cetorelli (2002) who shows that the theory can account for phenomena such as “club convergence”, economic miracles, growth disasters and reversals of fortune.

⁵¹ Consider, for instance, the limit case where agents are infinitely risk-averse. In this case, agents refuse to invest in the stock-market as long as this entails some uncertainty, i.e., as long as there are not enough savings in the economy to open all industries. Thus, an economy starting above Y_M converges to Y_H , while an economy starting below Y_M converges to Y_L , and is stuck in a poverty trap.

1 Recent empirical studies analyze implications of the theory about the patterns of 1
2 risk-sharing and diversification. Kalemli-Ozcan et al. (2001, 2003) document that 2
3 regions with access to better insurance through capital markets can afford a higher degree 3
4 of specialization. Using cross-country data at different levels of disaggregation, Imbs 4
5 and Wacziarg (2003) find robust evidence of sectoral diversification increasing in GDP. 5
6 However, their findings also suggest that, at a relatively late stage of the development 6
7 process, the pattern reverts and countries once more start to specialize. This tendency for 7
8 advanced countries to become more specialized as they grow can be explained by fac- 8
9 tors emphasized by the “new economic geography” literature [Krugman (1991)], from 9
10 which the theory described in this section abstracts, such as agglomeration externalities 10
11 and falling transportation costs. 11
12
13

14 **7. Endogenous fluctuations** 14

15
16 In the models reviewed so far, the economies converge in the long run to balanced- 16
17 growth equilibria characterized by linear dynamics. Growth models with expanding 17
18 variety and technological complementarities can, however, generate richer long-run dy- 18
19 namics, including limit cycles. In this section, we review two such models. 19

20 In the former, based on Matsuyama (1999), cycles in innovation and growth arise 20
21 from the deterministic dynamics of two-sector models with an endogenous market struc- 21
22 ture. The theory can explain some empirical observations about low-frequency cycles, 22
23 and their interplay with the growth process. In particular, it predicts that waves of rapid 23
24 growth mainly driven by “factor accumulation” are followed by spells of innovation- 24
25 driven growth. Interestingly, these latter periods are characterized by lower investments 25
26 and slower growth. This is consistent with the findings of Young (1995) that the growth 26
27 performance of East-Asian countries was mainly due to physical and human capital 27
28 accumulation, while there was little total factor productivity (TFP) growth. According 28
29 to Matsuyama’s theory, the observation of low TFP growth should not lead to the pes- 29
30 simistic conclusions that growth is destined to die-off. Rather, rapid factor accumulation 30
31 could set the stage for a new phase of growth characterized by more innovative activity. 31
32 The predictions of this theory bear similarities to those of models with General Purpose 32
33 Technologies (GPT), e.g., Helpman (1998) and Aghion and Howitt (1998, Chapter 8). 33
34 For instance, they predict that a period of rapid transformation and intense innovation 34
35 (e.g., the 1970’s) can be associated with productivity slowdowns. However, GPT-based 35
36 theories rely on the exogenous arrival of new “fundamental” innovations generating 36
37 downstream complementarities. In contrast, cycles in Matsuyama (1999) are entirely 37
38 endogenous.⁵² 38
39

40
41 ⁵² Cyclical equilibria can also emerge in Schumpeterian models, due to the dynamic relationship between 41
42 innovative investments and creative destruction. An example is the seminal contribution of Aghion and Howitt 42
43 (1992). More recently, Francois and Lloyd-Ellis (2003) construct a Schumpeterian model where entrepreneurs 43

1 In the latter model, based on Evans et al. (1998), cycles in innovation and growth are 1
2 instead driven by expectational indeterminacy. The mechanism in this paper is different, 2
3 as cycles hinge on multiple equilibria and sunspots. Some main predictions are also dif- 3
4 ferent: contrary to Matsuyama (1999), the equilibrium features a positive comovement 4
5 of investments and innovation. The main contribution of the paper is to show that cy- 5
6 cles can be learned by unsophisticated agents holding adaptive expectations. Thus, the 6
7 predictive power of the theory does not rest on the assumption that agents' expectations 7
8 are rational and that agents can compute complicated dynamic equilibria. 8
9

10 7.1. *Deterministic cycles* 10

11
12 Matsuyama (1999) presents a model of expanding variety where an economy can per- 12
13 petually oscillate in equilibrium between periods of innovation and periods of no inno- 13
14 vation. Cycles arise from the deterministic periodic oscillations of two state variables 14
15 (physical capital and knowledge). Unlike the model that will be discussed in the next 15
16 section, the equilibrium is determinate and there are no multiple steady-states. 16
17

18 More specifically, the source of the oscillatory dynamics is the market structure of the 17
19 intermediate goods market. Monopoly power is assumed to be eroded after one period. 18
20 The loss of monopoly power is due to the activity of a competitive fringe which can copy 19
21 the technology with a one-period lag. In every period, new industries are monopolized, 20
22 while mature industries are competitive. The profits of innovators depend on the mar- 21
23 ket structure of the intermediate sector. The larger is the share of competitive industries 22
24 in the intermediate sector, the lower is the profit of innovative firms, since competitive 23
25 industries sell larger quantities and charge lower prices. In periods of high innovation, 24
26 a large share of industries are monopolized, which increases the profitability of innova- 25
27 tion, thereby generating a feedback. In these times, investment in physical capital is low 26
28 due to the crowding out from the research activity. Conversely, times of low innovation 27
29 are times of high competition, since old monopolies lose power and there are few new 28
30 firms. Thus, the rents accruing to innovative firms are small. In these periods, savings 29
31 are invested in physical capital, and while innovation is low, the high accumulation of 30
32 physical capital creates the conditions for future innovation to be profitable. 31
32

33 Time is discrete. The production of final goods is as in (3), where we set $L_y = L = 1$. 33
34 Intermediate goods are produced using physical capital, with one unit of capital pro- 34
35 ducing one unit of intermediate product, x . Innovation also requires capital, with a 35
36 requirement of μ units of capital per innovation. Monopoly power is assumed to last one 36
37 period only. Therefore, in period t , all intermediate inputs with an index $z \in [0, A_{t-1}]$ 37
38

39
40 can decide to time the implementation of innovations [similarly to Shleifer (1986)]. In this model, agents time 40
41 the implementation so as to profit from buoyant demand and maximize the duration of their leadership. This 41
42 mechanism leads to a clustering of innovations and endogenous cycles. While this model can explain some 42
43 features of fluctuations at business cycle frequencies, Matsuyama's model is better suited for the analysis of 43
44 long waves.

are competitively priced, whereas all those with an index $z \in [A_{t-1}, A_t]$ are monopolistically priced. The prices of competitive and non-competitive varieties are $p_t^c = r_t$ and $p_t^m = r_t/\alpha$, respectively, where the superscript $h \in \{c, m\}$ denotes the market structure. The relative demand for two varieties x_t^c and x_t^m must be

$$\frac{x_t^c}{x_t^m} = \left(\frac{p_t^c}{p_t^m} \right)^{-1/(1-\alpha)} = \alpha^{-1/(1-\alpha)}. \quad (73)$$

The one-period monopoly profit is $\pi_t = p_t^m x_t^m - r_t x_t^m = x_t^m r_t (1 - \alpha)/\alpha$. Since patents expire after one period, π_t is also the value of a monopolistic firm at the beginning of period t . Therefore, free-entry implies:

$$\frac{1 - \alpha}{\alpha} x_t^m \leq \mu, \quad (74)$$

with equality holding when innovation is positive.

Capital is assumed to fully depreciate after each period. The stock of capital can be allocated to research or intermediate production, subject to the following resource constraint:

$$K_{t-1} = A_{t-1} x_t^c + (A_t - A_{t-1})(x_t^m + \mu),$$

implying

$$A_t - A_{t-1} = A_{t-1} \max \left\{ 0, \frac{(1 - \alpha)K_{t-1}}{\mu A_{t-1}} - \alpha^{-\alpha/(1-\alpha)} \right\}, \quad (75)$$

where we have used (73) and (74) to eliminate x_t^c and x_t^m . As shown by (75), there exists a threshold to the capital-knowledge ratio that triggers positive innovation. In particular, innovation occurs if $K_{t-1}/A_{t-1} \geq \alpha^{-\alpha/(1-\alpha)}(1 - \alpha)^{-1}\mu \equiv k_L$. If $K_{t-1}/A_{t-1} < k_L$, then, all capital is allocated to intermediate production, all intermediate industries are competitive and final production is given by the standard neoclassical Cobb–Douglas technology:

$$Y_t = A_{t-1}^{1-\alpha} K_{t-1}^\alpha. \quad (76)$$

In this case, an economy is said to be in a ‘‘Solow regime’’, with decreasing returns to capital. Since there is no investment in innovation, A is constant and the dynamics has a neoclassical character. In contrast, if $K_{t-1}/A_{t-1} \geq k_L$, then a positive share of the capital stock is allocated to innovation and final production equals:

$$Y_t = A_{t-1} \left[\alpha^{-1/(1-\alpha)} \frac{\alpha}{1-\alpha} \mu \right]^\alpha + (A_t - A_{t-1}) \left[\frac{\alpha}{1-\alpha} \mu \right]^\alpha.$$

Using (75) and simplifying terms, this equation can be written as

$$Y_t = DK_{t-1}, \quad (77)$$

1 where $D \equiv (k_L)^{-(1-\alpha)}$. In this case, the returns to capital are constant, like in endoge- 1
2 nous growth models, and the economy is said to be in a “Romer regime”.⁵³ 2

3 For tractability, we assume a constant savings rate, implying that $K_t = sY_t$.⁵⁴ Define 3
4 $k_t = k_L^{-1} \cdot K_t/A_t$ as the (adjusted) capital-to-knowledge ratio. Then, standard algebra 4
5 using (75), (76) and (77) establishes the following equilibrium law of motion: 5

$$6 \quad k_t = f(k_{t-1}) = \begin{cases} sDk_{t-1}^\alpha & \text{if } k_{t-1} < 1, \\ \frac{sDk_{t-1}}{1 + \alpha^{-\alpha/(1-\alpha)}(k_{t-1} - 1)} & \text{if } k_{t-1} \geq 1. \end{cases} \quad (78) \quad 6$$

7
8
9
10 The mapping $k_t = f(k_{t-1})$ has two fixed points. The first is $k = 0$, the second can 10
11 either be $k = (sD)^{1/(1-\alpha)} \equiv \hat{k}_1$, if $sD \leq 1$, or $k = 1 + \alpha^{\alpha/(1-\alpha)}(sD - 1) \equiv \hat{k}_2$, if 11
12 $sD \geq 1$. In the former case, the fixed point lies in the range of the “Solow regime”, 12
13 while in the latter, it lies in the range of the “Romer regime”.⁵⁵ 13

14 Three cases are possible: 14

- 15 1. If $sD \leq 1$, the economy converges monotonically to $k = \hat{k}_1$. In this case, the 15
16 economy never leaves the Solow regime, and there are no innovative investments. 16
17 The neoclassical dynamics converge to a stagnating level of GDP per capita. 17
- 18 2. If $sD > \max\{1, \alpha^{-\alpha/(1-\alpha)} - 1\}$, then capital first monotonically accumulates in 18
19 the Solow regime, with no innovation. The economy overcomes the development 19
20 threshold, $k = 1$ in finite time, and the process of innovation starts thereafter. 20
21 Eventually, the economy converges to the BG equilibrium \hat{k}_2 in an oscillatory 21
22 fashion. In the BG, capital and knowledge are accumulated at the same positive 22
23 rate, and income per capita grows over time. 23
- 24 3. If $sD \in (1, \alpha^{-\alpha/(1-\alpha)} - 1]$, the economy does not converge asymptotically to any 24
25 BG equilibrium, and perpetually oscillates in the long run between the Solow- 25
26 and the Romer-regime. This case is described by Figure 6. On the one hand, there 26
27 is no steady-state in the Solow-regime, which rules out that the economy can be 27
28 trapped in a stable equilibrium with no innovation. On the other hand, the steady- 28
29 state \hat{k}_2 is locally unstable and cannot be an attractor of the dynamics in itself. 29
30 Instead, there exists a period-2 cycle, such that one of the periodic points lies in the 30
31 Solow regime (k_S), while the other lies in the Romer regime (k_R).⁵⁶ The period-2 31

32
33 ⁵³ Zilibotti (1995) finds similar dichotomic equilibrium dynamics in a one-sector model with learning-by- 33
34 doing spillovers. Economies may converge to a stationary steady-state with “Solow dynamics” or embark on 34
35 a virtuous path of “Romer dynamics” with self-sustained growth. Cycles cannot arise in equilibrium, while 35
36 multiple self-fulfilling prophecies exist. 36

37 ⁵⁴ Matsuyama (2001) relaxes this restriction and characterizes equilibrium by a second-order difference equa- 37
38 tion. Some of the main results, like the existence of a period-2 cycle, survive this generalization. 38

39 ⁵⁵ It is easily verified that $f'(0) > 1$, $f'(\hat{k}_1) = \alpha \in (0, 1)$, and $f'(\hat{k}_2) = -(\alpha^{-\alpha/(1-\alpha)} - 1)/(sD)$, where 39
40 $f'(\hat{k}_2) \in (-1, 0)$ if $sD > \alpha^{-\alpha/(1-\alpha)} - 1$ and $f'(\hat{k}_2) < -1$ if $sD \in (1, \alpha^{-\alpha/(1-\alpha)} - 1)$. These properties 40
41 are used to establish the results discussed below. 41

42 ⁵⁶ A period-2 cycle exists if, given a mapping $x_{t+1} = f(x_t)$, $f(f(\cdot))$ has fixed points other than the fixed 42
43 point of $f(\cdot)$. A sufficient condition is that (i) $f(\cdot)$ is continuous, (ii) there exists a closed, finite interval, 43
44 I , such that $f(I) \subset I$ and (iii) $f(\cdot)$ has an unstable fixed point. (i) and (iii) are clearly satisfied; (ii) is 44
45 established in the next footnote for the interval I_{abs} . 45

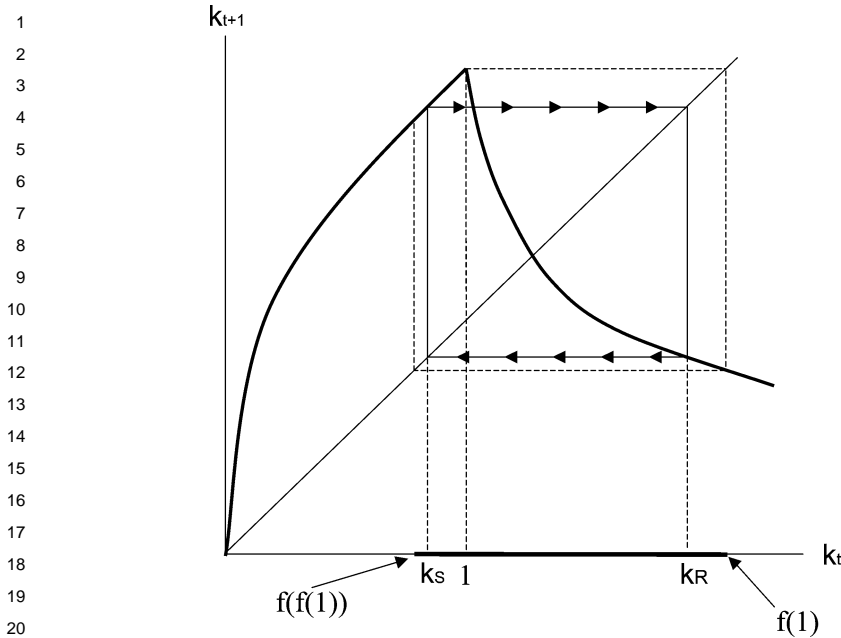


Figure 6.

cycle is not necessarily stable, and if it is unstable, the economy can converge to cycles of higher periodicity or feature chaos. A general property of the dynamics is that the economy necessarily enters the region $I_{abs} = [f(f(1)), f(1)]$ (shaded in Figure 6), and never escapes from it.⁵⁷

In the case described by Figure 6, the model predicts that a poor economy would first grow through capital accumulation, and eventually enter the absorbing region I_{abs} . Then, there is an alternance of periods of innovation and periods of no innovation. GDP per capita grows on average, but at a non-steady rate, and there are cycles in the innovative activity. Interestingly, output and capital grow more quickly in periods of no innovation (Solow regime) than in periods of high innovation (Romer regime). Another implication is that if an economy grows quickly, but has a low TFP growth, this does not imply that growth will die-off. Rather, fast capital accumulation can create the conditions for future waves of innovation, and vice versa.

⁵⁷ To prove this result, two properties of the mapping need to be shown [see Azariadis (1993)]. First, $f(\cdot)$ must be unimodal, i.e., (i) $f(\cdot)$ must be continuous; (ii) $f(\cdot)$ must be increasing in some left-hand neighborhood of 1 and decreasing in some right-hand neighborhood of 1. Second, it must be the case that $f(f(1)) < 1$. That f is unimodal is immediate by inspection. After some algebra, it can also be proved that $f(f(1)) < 1$.

1 7.2. Learning and sunspots 1

2
3 Evans et al. (1998) propose the following generalization of the technology (3) for final 3
4 production: 4

$$5 \quad Y = L^{1-\alpha} \left[\int_0^A x_j^\zeta dj \right]^\phi, \quad (79) \quad 6$$

7
8 where $\zeta\phi = \alpha$. This specification encompasses the technology (3), in the case of $\phi = 1$, 8
9 and allows intermediate inputs to be complements or substitutes. They focus on the case 9
10 of complementarity ($\phi > 1$), and show that in this case, the equilibrium can feature 10
11 multiple steady states, expectational indeterminacy and sunspots. They emphasize the 11
12 possibility of equilibria where the economy can switch stochastically between periods 12
13 of high and low growth. 13

14 Time is discrete, and intermediate firms rent physical capital from consumers to pro- 14
15 duce intermediate goods. One unit of capital is required per unit of intermediate good 15
16 produced. Capital is assumed not to depreciate. The resource constraint of this economy 16
17 is: 17

$$18 \quad Y_t = C_t + K_t \cdot \chi \left(\frac{K_{t+1} - K_t}{K_t} \right), \quad 18$$

19 where $\chi(\cdot)$ is a function such that $\chi' > 0$, $\chi'' \geq 0$. If there are no costs of adjustment, 19
20 then, $\chi(x) = x$. If $\chi'' > 0$, there are convex costs of adjustments. 20

21 By proceeding as in Section 2, we can characterize the equilibrium of the intermediate 21
22 industry.⁵⁸ The profit of intermediate producers, in particular, turns out to be: 22

$$23 \quad \pi = \Omega A^\xi (r p_K)^{\alpha/(\alpha-1)}, \quad (80) \quad 23$$

24 where $\xi \equiv (\phi - 1)/(1 - \alpha)$ and $\Omega \equiv (1 - \zeta)\zeta^{(1+\alpha)/(1-\alpha)}\phi^{1/(1-\alpha)}L$ are two posi- 24
25 tive constant. We denote by p_K the relative price of capital, expressed in terms of the 25
26 consumption good numeraire. If there are no adjustment costs, then, $p_K = 1$ while, in 26
27 general, $p_K = \chi'(\cdot)$. Note that profits increase with A , as long as $\phi > 1$. 27

28 Two technical assumptions ensure that the model has BG properties. First, the design 28
29 of a new good requires A_t^ξ units of output. Second, innovative investments incurred at t 29
30 only give the first profit in period $t + 1$. Free entry then implies: 30

$$31 \quad \sum_{s=0}^{\infty} \frac{\pi_{t+s}}{(1+r)^{s+1}} \leq p_{K,t} A_t^\xi. \quad (81) \quad 31$$

32 In a BG equilibrium, consumption and capital grow at the common rate, γ . When $\phi > 1$, 32
33 this rate exceeds the growth rate of technical knowledge, $\gamma_A \equiv A_{t+1}/A_t$. In particular, 33
34 34

35
36
37
38
39
40
41 ⁵⁸ Note that firms rent, and do not own, their capital stock. Adjustment costs are borne at the aggregate level, 41
42 not at the level of each decision unit. Therefore, it continues to be legitimate to write the profit maximization 42
43 problem for intermediate producers as a sequence of static maximization problems. 43

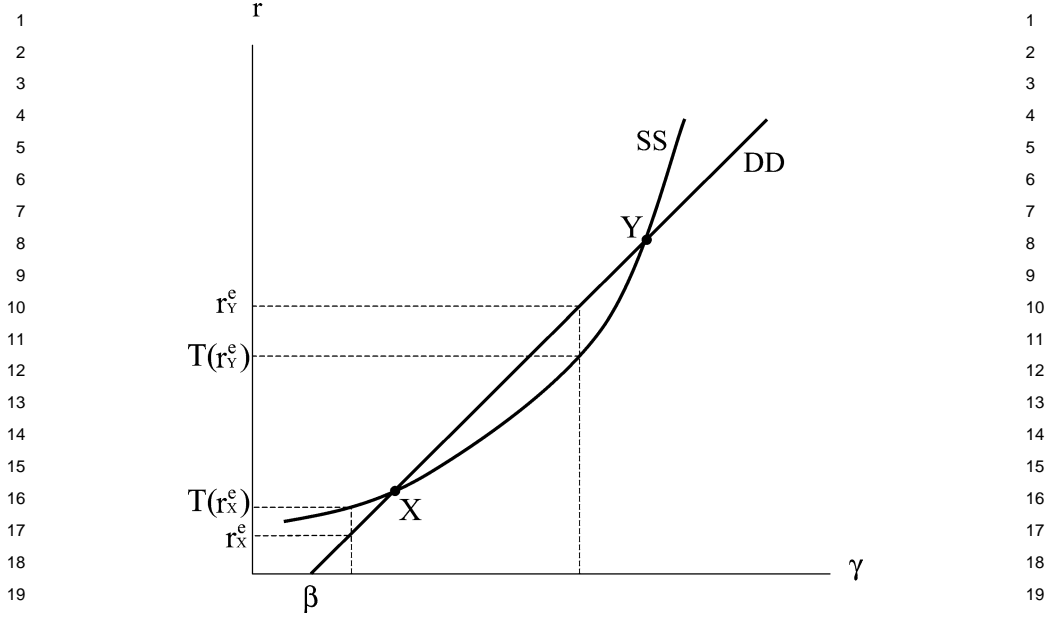


Figure 7.

it can be shown that $\gamma = \gamma_A^{1+\xi}$. Substituting (80) into (81), and solving, yields:

$$\gamma = (1 + r - \Omega(p_K)^{-1/(1-\alpha)} r^{-\alpha/(1-\alpha)})^{(\phi-\alpha)/(\phi-1)}, \quad (82)$$

which is the analogue of Equation (13) in the benchmark model.

The model is closed by the (discrete time) Euler equation for consumption:

$$\gamma = [\beta(1 + r)]^{1/\sigma}, \quad (83)$$

where β is the discount factor. Equations (82) and (83) fully characterize the equilibrium.

Figure 7 provides a geometric representation for the case of logarithmic preferences and zero adjustment cost ($p_K = 1$). The SS curve is linear, with the slope β^{-1} . The DD curve is also positively sloped. In the case represented, the two curves cross twice, thereby implying that there are two BG equilibria featuring positive innovation and growth (points X and Y).

Standard stability analysis is inappropriate for dynamic models with perfect foresight. It is possible, however, to analyze the expectational stability (E-stability) of the BG equilibria. E-stability is tested as follows. Set an arbitrary initial level for the expected interest rate r^e , and let agents choose their optimal savings plan according to (83). This implies a notional growth rate of consumption and capital, as determined by the SS curve. Next, firms take action. At the notional growth rate, there is a unique interest

1 rate consistent with the no-arbitrage condition implied by (82), as shown by the DD 1
2 curve. The composition of these two operations define a mapping from an expected to 2
3 a realized interest rate: 3

$$4 \quad r_t = T(r_t^e). \quad (84) \quad 4$$

5 A perfect foresight BG equilibrium is a fixed point to the mapping, $r = T(r)$. After 5
6 consumers have observed the realized interest rate, they update their expectations about 6
7 next period's interest rate using adaptive learning, i.e.: 7
8

$$9 \quad r_{t+1}^e = r_t^e + \psi_t(r_t - r_t^e), \quad (85) \quad 9$$

10 where $\xi_t = \psi/t$. The sequence $\{\psi_t\}$ determines how sensitive the expectations are to 10
11 past errors, and it is known as the gain sequence. Substituting (84) into (85) defines a 11
12 dynamic system, whose stability can be analyzed by linearization techniques. In general, 12
13 expectational stability occurs whenever $T'(r) < 1$, where r is the steady-state interest 13
14 rate.⁵⁹ 14
15

16 An inspection of Figure 7 shows the equilibrium X to be E-stable, while the equi- 16
17 librium Y is not. Let r_X^e and r_Y^e denote two expected interest rates which are below the 17
18 equilibria X and Y , respectively. Then, in the case of the equilibrium X , $T(r_X^e) > r_X^e$, 18
19 and the adaptive adjustment moves the economy towards the equilibrium, inducing con- 19
20 vergence. In contrast, in the case of the equilibrium Y , $T(r_Y^e) < r_Y^e$, and the adaptive 20
21 adjustment moves the economy away from the equilibrium, thereby inducing diver- 21
22 gence. 22

23 In the case analyzed so far, only one BG is E-stable, and E-stability can be used as 23
24 a selection criteria. It is possible, however, that multiple E-stable BG equilibria exist in 24
25 the general model with convex adjustment costs. 25

26 Figure 8 describes a case with four steady-states, two of them being E-stable. Equi- 26
27 libria such as X and Z are E-stable (note that $T(r_X^e) > r_X^e$ and $T(r_Z^e) > r_Z^e$). Moreover, 27
28 in the neighborhood of these equilibria, there exist stationary sunspot equilibria. In one 28
29 such equilibrium, the economy switches stochastically between two points in the neigh- 29
30 borhood of X and Z , respectively, with switching probabilities given by a time-invariant 30
31 transition probability matrix. The fact that both X and Z are E-stable is sufficient for 31
32 any stationary sunspot equilibrium in their neighborhood to be E-stable in itself.⁶⁰ 32
33

34 We conclude that a modified version of the model of growth with expanding vari- 34
35 ety can generate endogenous fluctuations. The key assumptions are complementarity 35
36 between capital goods and convex adjustment costs to capital. The former assumption 36
37 guarantees the existence of multiple BG equilibria, around which sunspot equilibria can 37
38 be constructed. The latter assumption guarantees that the sunspot equilibrium is expect- 38
39 ationally stable, i.e., it can be learned through adaptive expectations. 39

40
41 ⁵⁹ See Evans and Honkapohja (2001) for a state-of-art analysis of expectational indeterminacy. 41

42 ⁶⁰ For general discussion of sunspot equilibria, see Azariadis and Guesnerie (1986), Grandmont (1986) and 42
43 Azariadis (1993). 43

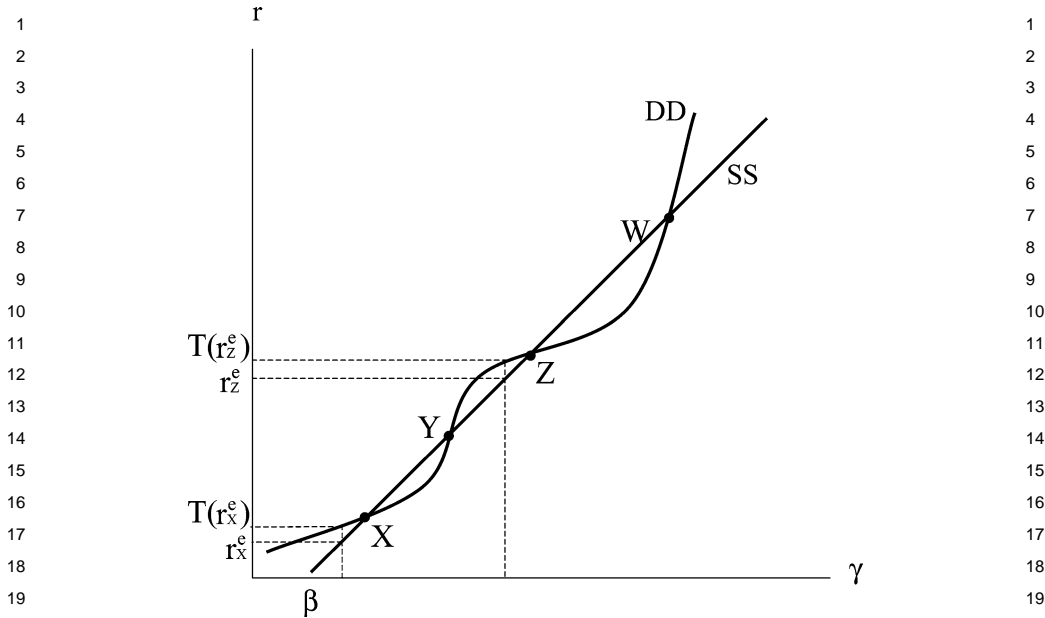


Figure 8.

The model assumes increasing returns to physical and knowledge capital. The reduced form representation of the final good technology is:

$$Y = A\phi^{-\alpha} K_F^\alpha,$$

where $K_F = Ax$ denotes aggregate capital used in intermediate production. Empirical estimates suggest that $\alpha \leq 0.4$, which implies that the lower bound to the output elasticity of knowledge to generate multiplicity is $\phi - \alpha = 0.6$. Evans et al. (1998) provide a numerical example of an E-stable sunspot equilibrium, assuming $\phi = 4$. Recent estimations from Porter and Stern (2000) using patent numbers report $\phi - \alpha$ to be around 0.1, however. Therefore, the model seems to require somehow extreme parameters to generate endogenous fluctuations.

Augmenting the model with other accumulated assets, such as human capital, may help obtain the results under realistic parameter configurations. This is complicated by the presence of scale effects in the expanding variety model. However, in a recent paper, Dalgaard and Kreiner (2001) formulate a version of the model with human capital accumulation and without scale effects. In their model, both human capital (embodied knowledge) and technical change (disembodied knowledge) are used to produce final goods. The scale effect is avoided by congestion effects in the accumulation of human capital. An interesting feature of this model is that, unlike other recent models without scale effects, positive long-run growth in income per capita does not hinge on positive population growth.

1 **8. Conclusions** 1

2
3 In this chapter, we have surveyed recent contributions to growth theory inspired by 3
4 Romer's (1990) expanding variety model. Key features of the theory are increasing 4
5 returns through the introduction of new products that do not displace existing ones and 5
6 the existence of monopoly rents providing an incentive for firms to undertake costly 6
7 innovative investments. This model has had a tremendous impact on the literature, and 7
8 we could only provide a partial review of its applications. Then, we decided to focus 8
9 on a few major themes: trade and biased technical change, with their effects on growth 9
10 and inequality, financial development, complementarity in the process of innovation and 10
11 endogenous fluctuations. 11

12 While only being a limited selection, these applications give a sense of the suc- 12
13 cess of the model in providing a tractable framework for analyzing a wide array of 13
14 issues in economic growth. In fact, we have shown how the model can incorporate a 14
15 number of general equilibrium effects that are fundamental in the analysis of trade, 15
16 wage inequality, cross-country productivity differences and other topics. Further, while 16
17 the original model has linear AK-dynamics, we have surveyed recent generalizations 17
18 featuring richer dynamics, which can potentially be applied to the study of financial de- 18
19 velopment and innovations waves. Given its longevity, flexibility and simplicity, we are 19
20 convinced that the growth model with horizontal innovation will continue to be useful 20
21 in future research. 21

22
23
24 **Acknowledgements** 24

25
26 We thank Philippe Aghion, Jeremy Greenwood, Kiminori Matsuyama and Jerome Van- 26
27 dennbusche for helpful comments, Zheng Song for excellent research assistance, and 27
28 Christina Lönnblad for editorial assistance. 28

29
30
31 **Uncited references** 31

32
33 [Jones (1999)] [Jones (2001)] 33
34

35
36 **References** 36

- 37
38 Acemoglu, D. (1998). "Why do new technologies complement skills? Directed technical change and wage 38
39 inequality". *Quarterly Journal of Economics* 113, 1055–1090. 39
40 Acemoglu, D. (2002). "Directed technical change". *Review of Economic Studies* 69, 781–809. 40
41 Acemoglu, D. (2003a). "Patterns of skill premia". *Review of Economic Studies* 70, 199–230. 41
42 Acemoglu, D. (2003b). "Labor- and capital-augmenting technical change". *Journal of the European Economic* 42
43 Association 1, 1–37. 43
44 Acemoglu, D., Zilibotti, F. (2001). "Productivity differences". *Quarterly Journal of Economics* 116, 563–606. 43

- 1 Acemoglu, D., Zilibotti, F. (1997). "Was Prometheus unbounded by chance? Risk, diversification and
2 growth". *Journal of Political Economy* 105, 710–751.
- 3 Aghion, P., Howitt, P. (1998). *Endogenous Growth Theory*. MIT Press, Cambridge, MA.
- 4 Aghion, P., Howitt, P. (1992). "A model of growth through creative destruction". *Econometrica* 60 (2), 323–
5 351.
- 6 Antras, P. (2004). "Is the U.S. aggregate production function Cobb–Douglas? New estimates of the elasticity
7 of substitution". *Contributions to Macroeconomics* 4 (1).
- 8 Arrow, K.J. (1962). "The economic implications of learning-by-doing". *Review of Economic Studies* 29 (1),
9 155–173.
- 10 Atkinson, A.B., Stiglitz, J.E. (1969). "A new view of technological change". *Economic Journal*, 573–578.
- 11 Azariadis, C. (1993). *Intertemporal Macroeconomics*. Blackwell.
- 12 Azariadis, C., Guesnerie, R. (1986). "Sunsspots and cycles". *Review of Economic Studies* 53, 725–737.
- 13 Barro, R.J., Sala-i-Martin, X. (1995). *Economic Growth*. McGraw-Hill.
- 14 Barro, R.J., Sala-i-Martin, X. (1997). "Technological diffusion, convergence, and growth". *Journal of Eco-
15 nomic Growth* 2, 1–26.
- 16 Benassy, J.-P. (1998). "Is there always too little research in endogenous growth with expanding product vari-
17 ety?". *European Economic Review* 42, 61–69.
- 18 Bencivenga, V., Smith, B. (1991). "Financial intermediation and endogenous growth". *Review of Economic
19 Studies* 58, 195–209.
- 20 Borjas, G.J., Freeman, R.B., Katz, L.F. (1997). "How much do immigration and trade affect labor market
21 outcomes". *Brookings Papers on Economic Activity*, 1–67.
- 22 Braudel, F. (1979). *Civilization and Capitalism*. Harper and Row, New York.
- 23 Caselli, F., Esquivel, G., Lefort, F. (1996). "Reopening the convergence debate: A new look at cross-country
24 growth empirics". *Journal of Economic Growth* 1, 363–389.
- 25 Caselli, F., Coleman, J. (2000). "The world technology frontier". NBER Working Paper 7904.
- 26 Cetorelli, N. (2002). "Could prometheus be bound again? A contribution to the convergence controversy".
27 *Journal of Economic Dynamics and Control* 27, 29–50.
- 28 Ciccone, A., Matsuyama, K. (1999). "Efficiency and equilibrium with dynamic increasing returns due to
29 demand complementarities". *Econometrica* 67, 499–525.
- 30 Ciccone, A., Matsuyama, K. (1996). "Start-up costs and pecuniary externalities as barriers to economic de-
31 velopment". *Journal of Development Economics* 49, 33–59.
- 32 Cohen, W.M., Levin, R.C. (1989). "Empirical studies of innovation and market structure". In: *Handbook of
33 Industrial Organization*, vol. II. Elsevier, Amsterdam, pp. 1059–1107.
- 34 Cooper, R.W., John, A. (1988). "Coordinating coordination failures in Keynesian models". *Quarterly Journal
35 of Economics* 103, 441–463.
- 36 Dalgaard, C.-J., Kreiner, C.T. (2001). "Is declining productivity inevitable?". *Journal of Economic Growth* 6,
37 187–203.
- 38 Dinopoulos, E., Segerstrom, P. (1999). "A Schumpeterian model of protection and relative wages". *American
39 Economic Review* 89, 450–472.
- 40 Dinopoulos, E., Segerstrom, P. (2003). "A theory of North–South trade and globalization". Mimeo. Stockholm
41 School of Economics.
- 42 Devereux, M.B., Lapham, B.J. (1994). The stability of economic integration and endogenous growth. *Quar-
43 terly Journal of Economics* 109, 299–305.
- 44 DeVries, J. (1990). *The Economy of Europe in an Age of Crisis; 1600–1750*. Cambridge University Press,
45 Cambridge.
- 46 Dixit, A.K., Stiglitz, J.E. (1977). "Monopolistic competition and optimum product diversity". *American Eco-
47 nomic Review* 67 (3), 297–308.
- 48 Diwan, I., Rodrik, D. (1991). "Patents, appropriate technology, and North–South trade". *Journal of Interna-
49 tional Economics* 30, 27–48.
- 50 Dornbusch, R., Fischer, S., Samuelson, P.A. (1977). "A continuum Ricardian model of comparative advantage,
51 trade and payments". *American Economic Review* 67, 823–839.

- 1 Epifani, P., Gancia, G. (2002). "The skill bias of world trade". IIES Seminar Paper 707. 1
- 2 Ethier, W.J. (1982). "National and international returns to scale in the modern theory of international trade". 2
American Economic Review 72 (3), 389–405. 3
- 3 Evans, G., Honkapohja, S., Romer, P. (1998). "Growth cycles". American Economic Review 88 (3), 495–515. 4
- 4 Evans, G., Honkapohja, S. (2001). Learning and Expectations in Macroeconomics. Princeton University 5
Press, Princeton. 5
- 6 Francois, P., Lloyd-Ellis, H. (2003). "Animal spirits through creative destruction". American Economic Re- 6
view 93, 530–550. 7
- 7 Funke, M., Strulik, H. (2000). "On endogenous growth with physical capital, human capital and product 8
variety". European Economic Review 44, 491–515. 9
- 8 Gancia, G. (2003). "Globalization, divergence and stagnation". IIES Seminar Paper 720. 9
- 9 Grandmont, J.-M. (1986). "Periodic and aperiodic behavior in discrete one-dimensional dynamical system". 10
In: Hildenbrand, W., Mas-Colell, A. (Eds.), Contributions to Mathematical Economics: In Honor of Ger- 11
ard Debreu. Elsevier, Amsterdam. 12
- 12 Greenwood, J., Jovanovic, B. (1990). "Financial development, growth and the distribution of income". Journal 13
of Political Economy 98, 1067–1107. 14
- 13 Griliches, Z., Schmookler, J. (1963). "Inventing and maximizing". American Economic Review 53, 725–729. 14
- 14 Grossman, G., Helpman, E. (1989). "Product development and international trade". Journal of Political Econ- 15
omy 97, 1261–1283. 16
- 15 Grossman, G., Helpman, E. (1991a). Innovation and Growth in the World Economy. MIT Press, Cambridge, 17
MA. 18
- 16 Grossman, G., Helpman, E. (1991b). "Endogenous product cycles". Economic Journal 101, 1214–1229. 19
- 17 Grossman, G., Lai, E. (2004). "International protection of intellectual property", American Economic Review, 20
in press. 20
- 18 Habakkuk, H.J. (1962). American and British Technology in the Nineteenth Century: Search for Labor Saving 21
Inventions. Cambridge University Press, Cambridge. 22
- 19 Hall, R., Jones, C.I. (1999). "Why do some countries produce so much more output per workers than others?". 23
Quarterly Journal of Economics 114, 83–116. 24
- 20 Hamermesh, D.S. (1993). Labor Demand. Princeton University Press, Princeton. 25
- 21 Helpman, E. (1993). "Innovation, imitation and intellectual property rights". Econometrica 61, 1247–1280. 26
- 22 Helpman, E. (1998). General Purpose Technology and Economic Growth. MIT Press, Cambridge, MA. 27
- 23 Helpman, E., Krugman, P. (1985). Market Structure and Foreign Trade. MIT Press, Cambridge, MA. 28
- 24 Hicks, J. (1932). The Theory of Wages. Macmillan, London. 28
- 25 Imbs, J., Wacziarg, R. (2003). "Stages of diversification". American Economic Review 93, 63–86. 29
- 26 Jones, C.I. (1995). "R&D-based models of economic growth". Journal of Political Economy 103 (August), 30
759–784. 31
- 27 Jones, C.I. (1999). "Growth: With or without scale effects". American Economic Review 89, 139–144. 31
- 28 Jones, C.I. (2001). "Was the industrial revolution inevitable? Economic growth over the very long run". Ad- 32
vances in Macroeconomics 1 (2). Article 1. 33
- 29 Judd, K.L. (1985). "On the performance of patents". Econometrica 53 (3), 567–585. 34
- 30 Kalemli-Ozcan, S., Sorensen, B.E., Yosha, O. (2001). "Economic integration, industrial specialization, and 35
the asymmetry of macroeconomic fluctuations". Journal of International Economics 55, 107–137. 36
- 31 Kalemli-Ozcan, S., Sorensen, B.E., Yosha, O. (2003). "Risk sharing and industrial specialization: Regional 37
and international evidence". American Economic Review 93, 903–918. 38
- 32 Kennedy, C. (1964). "Induced bias in innovation and the theory of distribution". Economic Journal 74, 541– 38
547. 39
- 33 Klenow, P.J. (1996). "Industry innovation: Where and why". Carnegie–Rochester Conference Series on Public 40
Policy 44, 125–150. 41
- 34 Klenow, P.J., Rodriguez-Clare, A. (1997). "The neoclassical revival in growth economics: Has it gone too 41
far?". In: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual 1997. MIT Press, Cam- 42
bridge, MA, pp. 73–102. 43

- 1 **Krugman, P.** (1991). "Increasing returns and economic geography". *Journal of Political Economy* 99, 483–
2 499. 1
- 3 **Krugman, P., Venables, A.** (1995). "Globalization and the inequality of nations". *Quarterly Journal of Eco-*
4 *nomics* 110, 857–880. 3
- 5 **Krusell, P., Ohanian, L., Rios-Rull, V., Violante, G.** (2000). "Capital skill complementarity and inequality".
6 *Econometrica* 68, 1029–1053. 5
- 7 **Kwan, F.Y.K., Lai, E.L.C.** (2003). "Intellectual property rights protection and endogenous economic growth".
8 *Journal of Economic Dynamics and Control* 27, 853–873. 6
- 9 **Lai, E.L.C.** (1998). "International intellectual property rights protection and the rate of product innovation".
10 *Journal of Development Economics* 55, 133–153. 8
- 11 **Lai, E.L.C., Qiu, L.D.** (2003). "The North's intellectual property rights standard for the South?". *Journal of*
12 *International Economics* 59, 183–209. 9
- 13 **Levhari, D., Sheshinski, E.** (1969). "A theorem on returns to scale and steady-state growth". *Journal of Polit-*
14 *ical Economy* 77, 60–65. 11
- 15 **Martin, P., Rey, H.** (2000). "Financial integration and asset returns". *European Economic Review* 44, 1327–
16 1350. 13
- 17 **Martin, P., Rey, H.** (2001). "Financial super-markets: Size matters for asset trade". NBER Working Paper
18 8476. 14
- 19 **Martin, P., Rey, H.** (2002). "Financial globalization and emerging markets: With or without crash?". NBER
20 Working Paper 9288. 16
- 21 **Matsuyama, K.** (1995). "Complementarities and cumulative processes in models of monopolistic competi-
- 22 tion". *Journal of Economic Literature* 33, 701–729. 18
- 23 **Matsuyama, K.** (1996). "Why are there rich and poor countries? Symmetry-breaking in the world economy".
24 *Journal of the Japanese and International Economies* 10, 419–439. 19
- 25 **Matsuyama, K.** (1997). "Complementarity, instability and multiplicity". *The Japanese Economic Review* 48,
26 240–266. 21
- 27 **Matsuyama, K.** (1999). "Growing through cycles". *Econometrica* 67, 335–348. 22
- 28 **Matsuyama, K.** (2001). "Growing through cycles in an infinitely lived agent economy". *Journal of Economic*
29 *Theory* 100, 220–234. 23
- 30 **Murphy, K.M., Shleifer, A., Vishny, R.W.** (1989). "Industrialization and the big push". *Quarterly Journal of*
31 *Economics* 106 (2), 503–530. 25
- 32 **Neary, P.** (2003). "Globalisation and market structure". *Journal of the European Economic Association* 1,
33 245–271. 27
- 34 **Nordhaus, W.D.** (1969a). "An economic theory of technological change". *American Economic Review* 59
35 (2), 18–28. 28
- 36 **Nordhaus, W.D.** (1969b). *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*.
37 MIT Press, Cambridge, MA. 30
- 38 **North, D., Thomas, R.P.** (1973). *The Rise of the Western World: A New Economic History*. Cambridge
39 University Press, Cambridge. 31
- 40 **Parente, S.L., Prescott, E.C.** (1994). "Barriers to technology adoption and development". *Journal of Political*
41 *Economy* 102, 298–321. 32
- 42 **Porter, M.E., Stern, S.** (2000). "Measuring the "ideas" production function: Evidence from international patent
43 output". NBER Working Paper 7891. 35
- 44 **Prescott, E.C.** (1998). "Needed: A theory of total factor productivity". *International Economic Review* 39,
45 525–553. 37
- 46 **Ramey, G., Ramey, V.** (1995). "Cross-country evidence of the link between volatility and growth". *American*
47 *Economic Review* 85, 1138–1151. 39
- 48 **Rebelo, S.** (1991). "Long-run policy analysis and long-run growth". *Journal of Political Economy* 99 (3),
49 500–521. 40
- 50 **Rivera-Batiz, L., Romer, P.** (1991a). "Economic integration and endogenous growth". *Quarterly Journal of*
51 *Economics* 106, 531–555. 42

- 1 Rivera-Batiz, L., Romer, P. (1991b). "International trade with endogenous technological change". *European Economic Review* 35, 971–1004. 1
- 2 2
- 3 Rodriguez-Clare, A. (1996). "The division of labor and economic development". *Journal of Development Economics* 49, 3–32. 3
- 4 4
- 5 Romer, P. (1986). "Increasing returns and long-run growth". *Journal of Political Economy* 94, 1002–1037. 5
- 6 Romer, P. (1987). "Growth based on increasing returns due to specialization". *American Economic Review* 77, 56–62. 6
- 7 Romer, P. (1990). "Endogenous technological change". *Journal of Political Economy* 98, 71–102. 7
- 8 Romer, P. (1994). "New goods, old theory and the welfare costs of trade restrictions". *Journal of Development Economics* 43, 5–77. 8
- 9 Rosenberg, N. (1976). *Perspectives on Technology*. Cambridge University Press, Cambridge. 9
- 10 Samuelson, P. (1965). "A theory of induced innovations along Kennedy–Weisacker lines". *Review of Economics and Statistics*, 444–464. 10
- 11 11
- 12 Schmookler, J. (1966). *Invention and Economic Growth*. Harvard University Press. 12
- 13 Schumpeter, J.A. (1942). *Capitalism, Socialism and Democracy*. Harper, New York. 13
- 14 Segerstrom, P.S., Anant, T.C.A., Dinopoulos, E. (1990). "A Schumpeterian model of the product life cycle". *American Economic Review* 80, 1077–1091. 14
- 15 Shell, K. (1973). "Inventive activity, industrial organization and economic growth". In: Mirrlees, J.A., Stern, N.H. (Eds.), *Models of Economic Growth*. Wiley, New York. 15
- 16 16
- 17 Shleifer, A. (1986). "Implementation cycles". *Journal of Political Economy* 94, 1163–1190. 17
- 18 Solow, R.M. (1956). "A contribution to the theory of economic growth". *Quarterly Journal of Economics* 70, 65–94. 18
- 19 19
- 20 Spence, M. (1976). "Product selection, fixed costs, and monopolistic competition". *Review of Economic Studies* 43, 217–235. 20
- 21 Stiglitz, J.E. (1970). "Factor price equalization in a dynamic economy". *Journal of Political Economy* 78, 456–488. 21
- 22 22
- 23 Swan, T.W. (1956). "Economic growth and capital accumulation". *Economic Record* 32, 334–361. 23
- 24 Thoenig, M., Verdier, T. (2003). "Trade induced technical bias and wage inequalities: A theory of defensive innovations". *American Economic Review* 93, 709–728. 24
- 25 Ventura, J. (1997). "Growth and interdependence". *Quarterly Journal of Economics* 112, 57–84. 25
- 26 Vernon, R. (1966). "International investment and international trade in product-cycle". *Quarterly Journal of Economics* 80, 190–207. 26
- 27 27
- 28 Wood, A. (1994). *North–South Trade, Employment and Inequality: Changing Fortunes in a Skill Driven World*. Clarendon Press, Oxford. 28
- 29 29
- 30 Yang, G., Maskus, K.E. (2001). "Intellectual property rights, licensing and innovation in an endogenous product-cycle model". *Journal of International Economics* 53, 169–187. 30
- 31 Young, A. (1928). "Increasing returns and economic progress". *Economic Journal* 38 (152), 527–542. 31
- 32 Young, A. (1991). "Learning by doing and the dynamic effects of international trade". *Quarterly Journal of Economics* 106, 369–406. 32
- 33 33
- 34 Young, A. (1993). "Substitution and complementarity in endogenous innovation". *Quarterly Journal of Economics* 108, 775–807. 34
- 35 Young, A. (1995). "The tyranny of numbers: Confronting the statistical realities of the East Asian growth experience". *Quarterly Journal of Economics* 110, 641–680. 35
- 36 36
- 37 Xie, D. (1998). "An endogenous growth model with expanding ranges of consumer goods and of producer durables". *International Economic Review* 39, 439–460. 37
- 38 38
- 39 Zilibotti, F. (1994). "Endogenous growth and intermediation in an archipelago economy". *Economic Journal* 104, 462–473. 39
- 40 40
- 41 Zilibotti, F. (1995). "A Rostovian model of endogenous growth and underdevelopment traps". *European Economic Review* 39, 1569–1602. 41
- 42 42
- 43 43

Proof of Raw Subject Index

1	_____	complementarity	1
2	Page: 50	multiple steady states	2
3	Romer regime	expectational indeterminacy	3
4	_____	_____	4
5	Page: 52	Page: 53	5
6	learning and sunspots	expectational stability	6
7			7
8			8
9			9
10			10
11			11
12			12
13			13
14			14
15			15
16			16
17			17
18			18
19			19
20			20
21			21
22			22
23			23
24			24
25			25
26			26
27			27
28			28
29			29
30			30
31			31
32			32
33			33
34			34
35			35
36			36
37			37
38			38
39			39
40			40
41			41
42			42
43			43