J Econ Growth (2011) 16:135–156
DOI 10.1007/s10887-011-9066-1

Innovations, patent races and endogenous growth

Joseph Zeira

Published online: 12 May 2011

Abstract This paper presents a model of innovations and endogenous economic growth with two main assumptions: first, the cost of searching for innovations differs across innovations, and second, innovations take time to find. The paper shows that given these two assumptions together, competition leads to patent races and to duplication of innovative activity. The paper then shows that duplication significantly reduces the effect of scale on growth. It also shows that competitive R&D creates too much research on easy innovations, and too little research on the difficult ones. Finally, the paper shows that risk sharing might increase duplication and reduce growth.

Keywords Innovations · Patent races · Endogenous growth · Duplication

JEL Classification O31 · O40

1 Introduction

This paper presents a model of innovations and economic growth, in which patent races emerge endogenously. A patent race is defined here as a simultaneous race between many teams, all searching for the same innovation and only one finds it first. Such patent races create duplication of innovative activity, which is Pareto-inefficient. The main result of the paper

J. Zeira (✉)
Department of Economics, The Hebrew University of Jerusalem, Mt. Scopus, 91905, Jerusalem, Israel

J. Zeira
LUISS Guido Carli, Rome, Italy

J. Zeira
CEPR, London, UK

J. Zeira
RCEA, Rimini, Italy

is that such patent races emerge when two conditions are met. First, if the cost of research is not uniform across innovations. Second, if innovations are searched simultaneously. Then, the less costly innovations attract more than one team, as the higher return from winning compensates for the lower probability of success. Thus, patent races emerge. A second result of the paper is that optimality of competitive R&D differs across innovations. Easier innovations are over-researched while the harder are under-researched. This has important policy implications with respect to subsidization of R&D. A third result of the paper is related to the strong scale effect of the early endogenous growth models, which has been criticized on empirical grounds. This paper shows that duplication significantly reduces this theoretical scale effect. The fourth result of the paper is that insurance of patent races increases duplication and reduces the rate of economic growth.

The paper synthesizes two important lines of literature, one is microeconomic on duplication in innovation, and one is macroeconomic on endogenous growth. The first literature examines the market structure of the search for an innovation, which involves a patent race and duplication. This literature began with Loury (1979), Lee and Wilde (1980), and Dasgupta and Stiglitz (1980a,b).[1] While this literature focused on a single innovation, the R&D based endogenous growth models introduced an environment with infinite new potential innovations. In the first generation of such models, by Romer (1990), Segerstrom et al. (1990), Grossman and Helpman (1991b), and Aghion and Howitt (1992), all potential innovations have the same cost. As a result no such patent race emerges, as innovators always prefer to develop a new innovation, rather than work on one that is already searched for by someone else.

The two lines of literature are combined in this paper by focusing on the interaction between two assumptions. One is that the cost of research increases with innovations. This assumption has already been added to the endogenous growth literature by Jones (1995b), Kortum (1997) and Segerstrom (1998), mainly to reduce the theoretical scale effect. The second assumption is that many innovations are searched simultaneously, as they take time to find. None of these two assumptions alone leads to patent races and duplication. Only when the two are combined, so that innovations of high costs are searched at the same time as innovations with low costs, patent races and duplication emerge. This is because innovators can choose between a costly innovation, which no one else works on, and a patent race for a less costly innovation, with a lower probability of success, but with higher profits in case of success.

The main idea of the paper is presented by a simple benchmark model, where agents are risk neutral and the economy grows through innovations of the type of expanding variety. Innovations differ by cost, which is measured by the size of a team needed to find an innovation. This size is increasing over innovations, namely they are ordered by size. Time is discrete, so in each period many innovations are searched simultaneously. As a result patent races emerge.

Patent races affect the macroeconomics of growth in a number of ways. First, they reduce the strong effect of scale on growth, which was predicted by early endogenous growth models. This scale effect has been a source of criticism on this literature, since it did not fit the data, as shown by Jones (1995a) and others. There have been a few attempts to cope with this criticism by introducing changes to the growth models. This paper shows that duplication also contributes to the reduction of the scale effect. As scale increases, gains from innovation

[1] Irwin and Klenow (1996) supply empirical evidence for duplication in the US semiconductor industry.

increase and induce entry to the R&D sector. But only some new entrants search for new innovations, while others join existing patent races.[2]

Another result of the model is that the competitive equilibrium is inefficient. There are too many people searching for the low-cost innovations, while some of the high-cost innovations are under-researched. This means that there is too much R&D in the economy, but too little innovations. The paper shows that this result is quite robust, as it holds even if finding an innovation is uncertain. In such a case the probability of finding the innovation increases if more than one team searches for it, so a patent race might not be pure waste. But the paper shows that even then, patent races in equilibrium are larger than optimal, namely that there is duplication.

Finally, the paper extends the model to deal with risk aversion and with the effect of insurance. Since participation in a patent race is very risky, there is a strong incentive for risk sharing. Such insurance can be achieved either financially through venture capital, or by including many teams that search for an innovation within a single firm and by internalizing the patent race. Such a firm hires many teams, even if one is sufficient to find the innovation, in order to increase the probability of being first and to deter potential competitors. The paper shows that such risk sharing, by venture capital or by large research firms, leads to more R&D but to lower growth rates.[3]

The paper is organized as follows. Section 2 presents the benchmark model and Sect. 3 analyzes the equilibrium. Section 4 examines the effect of scale. Section 5 explores the inefficiency of equilibrium while Sect. 6 examines the results when R&D is uncertain. Section 7 extends the analysis to risk aversion and Sect. 8 concludes. The Appendix contains mathematical proofs.

2 The benchmark model

Consider an economy in a discrete time framework. There is a single final good used for consumption and investment, which is produced by labor and by a continuum of capital goods. The production of the final good in period t is described by the following production function:

$$Y_t = L_t^{\alpha} \int_0^{f_t} X_t(j)^{1-\alpha} dj, \tag{1}$$

where Y_t is output of the final good, $X_t(j)$ is input of capital good j in period t and f_t is the total amount of capital goods invented until period $t-1$ and used in production from period t on. Once invented, a capital good is produced and sold by the inventing team. Each capital good is produced from the final good, where b units of the final good make 1 unit of good j. In order to simplify notation assume from here on that $b = 1$, as it has no effect on the results. The capital goods must be invested one period ahead of time and they fully depreciate after one period of production. Note that the production side of the model is similar to Romer (1990), to make the model easily comparable to many endogenous growth models.

[2] Some endogenous growth models, like Stokey (1995) and Jones and Williams (1998, 2000), acknowledge the possibility of duplication and its effect on growth, calling it the "stepping on toes effect," but they do not model it explicitly as this paper does. Segerstrom et al. (1990) and Etro (2002) have patent races without duplication.

[3] Cooperation in research is also analyzed in Cozzi and Tarola (2006), but with a very different motivation of reducing the effect of information lags. Interestingly, in their paper cooperation also reduces growth.

Inventors operate in teams to search for new innovations, namely new capital goods. Each team searches in a single period for a specific potential innovation, so that research is directed and innovation-specific. The size of a team, which is required to find an innovation with certainty, differs across innovations. Furthermore, the size required to find innovation j in period t, $s_{j,t}$, is assumed to depend both on j and on the recent frontier of technology f_t. This required size is therefore described by a function S:

$$s_{j,t} = S(j, f_t). \tag{2}$$

The function S is assumed to be continuous and increasing in j, as the easier innovations are found earlier. It is also assumed to be decreasing in f_t, since innovators gain from the accumulation of knowledge. This is the famous "spillover assumption" of the R&D based endogenous growth models.

We can now compare our assumption, as described by Eq. 2, with the main R&D growth models. First note that according to most models, where innovation is instantaneous in continuous time, there is no distinction between f_t and j. As a result S depends on f_t only. Hence, in terms of this model, the first-generation endogenous growth models of Romer (1990), Grossman and Helpman (1991b) and Aghion and Howitt (1992) can be characterized by the following version of S:

$$S(j, f_t) = \frac{1}{af_t}. \tag{3}$$

Hence, productivity of each innovator is linear with respect to the frontier of technology and is equal to af_t. To gain a more intuitive understanding of this assumption, note that while $S(j, f_t)$ is the marginal cost of increasing technology by an absolute unit, the size times the level of previous technology, $f_t S(j, f_t)$, is the cost of a proportional increase of technology. Hence, (3) means, that the original endogenous growth models assumed that the cost of a proportional increase of technology is constant. Later, a group of papers by Jones (1995b), Kortum (1997), and Segerstrom (1998), which are now termed 'semi-endogenous growth models,' assumed, following Evenson and Kislev (1976), that the cost of increasing technology proportionately is not constant but increasing, due to growing complexity of innovation. Thus, in the framework of this paper the Jones (1995b) formulation can be described by the following S function:

$$S(j, f_t) = \frac{1}{af_t^{1-\theta}}, \quad \text{where } 0 < \theta < 1. \tag{4}$$

This paper combines the assumption that innovations become harder to search with the assumption that innovations take time to develop, which is captured by the discrete time framework. Thus j and f_t are not identical any longer, as reflected in (2). As a result, several innovations are searched simultaneously, while their costs are not equal. The paper examines the general version of the cost function (2), but focuses mainly on one specification, which leads to a balanced growth path:

$$S(j, f_t) = \frac{1}{af_t} s\left(\frac{j}{f_t}\right). \tag{5}$$

Assume that the function s is increasing and satisfies: $s(1) = 1$. This specification means that the cost of a proportional increase in technology is increasing with the distance of the innovation from the previous technology frontier.

As described above, a team of the right size can find innovation j in period t with certainty.[4] Assume that a smaller team has probability zero of finding the desired innovation. Also assume that an innovation can be searched and found by more than one team, but only one team finds it first. Therefore, only this first team receives the patent rights and becomes the unique producer and seller of capital good j.[5] Also assume that patent rights hold for infinity.[6] Assume that the probability to be first in the patent race is equal for all teams who can find it. Hence, if the number of teams in the race to find j in period t is $n_t(j)$, this probability to be first is equal to:

$$\frac{1}{n_t(j)}. \tag{6}$$

The population in the economy consists a continuum of size N of infinitely lived individuals. Individuals are assumed to be risk neutral.[7] The utility they derive in time 0 from life-time consumption is:

$$u = E_0 \sum_{t=0}^{\infty} \frac{c_t}{(1+\rho)^t}. \tag{7}$$

In every period individuals can choose whether to become workers or innovators. Workers earn a wage w_t in period t. Innovators in a winning team get an equal share from the value of the firm that becomes a monopoly, which produces the intermediate good j. Innovators in a team that did not win get nothing. At the beginning of each period people maximize expected income by choosing between production and research, between innovations to search for, and in each innovation between the various teams in the patent race. As a result a team maximizes expected income of its members.

3 Equilibrium in the benchmark model

3.1 Profit maximization

Producers of the final good maximize profits given wages w_t, prices of the capital goods $p_{t-1}(j)$ for all j and the interest rate r_t, which is set in period $t-1$ and paid in t. The first order condition with respect to labor is:

$$w_t = \frac{\partial Y_t}{\partial L_t} = \frac{\alpha Y_t}{L_t}. \tag{8}$$

The first order condition with respect to the capital good j is:

$$(1+r_t)p_{t-1}(j) = \frac{\partial Y_t}{\partial X_t(j)} = (1-\alpha)L_t^{\alpha} X_t(j)^{-\alpha}. \tag{9}$$

The producer of the capital good is a monopoly, due to her patent rights. Given the demand for the capital good in period $t-1$ by (9), the maximization of profits by the monopoly j in

4 The case of uncertainty in finding an innovation is analyzed in Sect. 6.

5 This type of patent race, where teams compete against each other for the same innovation, follows the Industrial Organization literature, mainly Loury (1979) and Lee and Wilde (1980). In some endogenous growth models the term patent race was used for a race against the previous or next innovation. See for example Aghion and Howitt (1992).

6 An alternative modeling, where patents have finite lifetime, yields similar results.

7 This assumption is changed in Sect. 7, where the effect of risk aversion is analyzed.

$t-1$ leads to the following first order condition:

$$1 = \frac{(1-\alpha)^2}{1+r_t} L_t^{\alpha} X_t(j)^{-\alpha} = (1-\alpha) p_{t-1}(j). \tag{10}$$

Hence, the price of the capital good is constant over j and over time:

$$p_{t-1}(j) = \frac{1}{1-\alpha}. \tag{11}$$

As a result the quantities of capital goods are equal for all industries:

$$X_t(j) = L_t (1-\alpha)^{\frac{2}{\alpha}} (1+r_t)^{-\frac{1}{\alpha}}. \tag{12}$$

Substituting (12) in (1) we get that output of the final good in period t is equal to:

$$Y_t = f_t L_t (1-\alpha)^{\frac{2(1-\alpha)}{\alpha}} (1+r_t)^{\frac{\alpha-1}{\alpha}}. \tag{13}$$

Hence, output per production worker is proportional to the level of technology f_t.

The value of innovation j, invented and patented in period t, $v_t(j)$, is equal to the discounted sum of monopoly profits from that period on. From (11) and (12) we get that maximized profits in t are:

$$p_t(j) X_{t+1}(j) - X_{t+1}(j) = L_{t+1} \alpha (1-\alpha)^{\frac{2-\alpha}{\alpha}} (1+r_{t+1})^{-\frac{1}{\alpha}}.$$

Note that this profit is equal for all innovations. Also, since consumers are risk neutral, the interest rate is equal to their subjective discount rate and is therefore constant over time. Let us denote this constant interest rate by r: $r = \rho$. Hence, the value of an innovation invented and patented in period 0 is equal to:

$$v_0(j) = v_0 = \alpha (1-\alpha)^{\frac{2-\alpha}{\alpha}} (1+r)^{-\frac{1}{\alpha}} \sum_{t=0}^{\infty} \frac{L_{t+1}}{(1+r)^t}. \tag{14}$$

From here on we focus on period 0 without loss of generality.

3.2 Income levels and the size of the patent race

Workers and innovators choose their profession in the beginning of each period according to expected income. Wages are derived from conditions (8) and (13) and are equal in period 0 to:

$$w_0 = f_0 \alpha (1-\alpha)^{\frac{2-2\alpha}{\alpha}} (1+r)^{\frac{\alpha-1}{\alpha}}.$$

The expected income of an inventor in period 0 is equal to the value of innovation v_t times the probability to find the innovation first, divided by the size of the team:

$$\frac{1}{n_0(j)} \frac{v_0}{S(j, f_0)} = \frac{1}{n_0(j) S(j, f_0)} \alpha (1-\alpha)^{\frac{2-\alpha}{\alpha}} (1+r)^{-\frac{1}{\alpha}} \sum_{t=0}^{\infty} \frac{L_{t+1}}{(1+r)^t}.$$

Individuals compare this expected income to the alternative income from production and as long as it is higher, more teams enter the patent race.

Hence, entry in period 0 continues as long as:

$$n_0(j) S(j, f_0) \leq \sum_{t=0}^{\infty} \frac{L_{t+1}}{(1+r)^t} \frac{1-\alpha}{f_0(1+r)} = \frac{1-\alpha}{f_0} \sum_{t=1}^{\infty} \frac{L_t}{(1+r)^t}.$$

Note that the number of teams in a patent race is an integer. Hence, the equilibrium size of the patent race for innovation j in time 0 is:

$$n_0(j) = \text{int}\left\{\frac{1-\alpha}{S(j, f_0)}\frac{1}{f_0}\sum_{t=1}^{\infty}\frac{L_t}{(1+r)^t}\right\}. \tag{15}$$

Equation 15 implies that the size of the patent race is negatively related to the required size of a team, since less costly innovations are more profitable. Hence, the size of patent races is diminishing with j. It also depends positively on the future scales of production.

Denote by $R_t(j)$ the number of innovators who try to find innovation j in period t. It is equal to the number of teams, which participate in the patent race, times the size of each of these teams. Hence the number of innovators searching for j in period 0 is:

$$R_0(j) = n_0(j)S(j, f_0).$$

The function $R_0(j)$ has a see-saw shape. It rises with j up to $\frac{1-\alpha}{f_0}\sum_{t=1}^{\infty}\frac{L_t}{(1+r)^t}$, where $n_0(j)$ falls by 1 to a smaller integer and then rises again. This continues until the size of the patent race is 1.

3.3 The equilibrium amount of new innovations

We next determine how many innovations are searched and found in each period. As research teams become larger, the size of the patent race declines, until it reaches 1. This is the last innovation researched in this period, since an innovation of a higher j, with a larger required size, generates income that is lower than the wage and will not be searched for. Clearly, all previous innovations with smaller required size are found with certainty. Hence, the amount of innovations in period 0 and the new technology frontier are determined by:

$$S(f_1, f_0) = \frac{1-\alpha}{f_0}\sum_{t=1}^{\infty}\frac{L_t}{(1+r)^t}. \tag{16}$$

Equation 16 presents the scale effect in this economy, through the positive effect of future scales of production on the current amount of innovation. Note that if these future scales of production are not sufficiently large, there might be a situation of no innovation at all, if:

$$S(f_0, f_0) \geq \frac{1-\alpha}{f_0}\sum_{t=1}^{\infty}\frac{L_t}{(1+r)^t}.$$

Finally, note that Eq. 16 enables us to rewrite the size of each patent race slightly differently, by substituting in Eq. 15:

$$n_0(j) = \text{int}\left\{\frac{S(f_1, f_0)}{S(j, f_0)}\right\}. \tag{17}$$

This equation implies that patent races appear endogenously in this economy and their size can be larger than 2, if the size function S rises sufficiently within a period of time. This is clearly the main result of the paper. Equation 17 also indicates that this result reflects two main assumptions of the paper. The first is that costs of innovation are increasing, namely $S_j > 0$, and the second is that innovation takes time, so that all innovations from f_0 to f_1 are searched in parallel, where $f_1 > f_0$.

3.4 Equilibrium R&D

Knowing the size of each patent race and the equilibrium amount of innovations enables us to calculate the overall size of the R&D sector in period 0, R_0:

$$R_0 = \int_{f_0}^{f_1} n_0(j) S(j, f_0) dj = \int_{f_0}^{f_1} S(j, f_0) \text{int} \left\{ \frac{S(f_1, f_0)}{S(j, f_0)} \right\} dj. \tag{18}$$

Hence the size of the R&D sector is a function of the previous technology frontier and of the future scales of production, as these two variables determine f_1 through (16). Finally we add the labor market equilibrium condition, that the sum of workers in the production and the R&D sectors equals the total number of workers in the economy:

$$L_0 + R_0 = N. \tag{19}$$

We can now fully characterize the general equilibrium in the model. Since f_1 depends on the previous technology frontier and on the future scales of production, so do R_0 and hence L_0. We therefore get the following dynamic condition:

$$L_0 = \phi \left[f_0, \sum_{t=1}^{\infty} \frac{L_t}{(1+r)^t} \right].$$

In general, the solution of this forward looking dynamic equation is quite complicated and depends on the size function S. In the next subsection we show that for the size function (5) this solution is a balanced growth path.

3.5 A balanced growth path

For tractability we restrict the analysis from here on to the specific size function (5).

Proposition 1 *If the size function is described by* (5)*, the equilibrium is characterized by constant sizes of the production and R&D sectors and by a constant rate of growth g, which is determined by the following equation:*

$$s(1+g) \frac{r}{(1-\alpha)a} + R(g) = N, \tag{20}$$

where the function R is defined by:

$$R(g) = a^{-1} \int_0^g s(1+i) int \left\{ \frac{s(1+g)}{s(1+i)} \right\} di.$$

Proof See Appendix.

Note that R, which describes the size of the R&D sector, is an increasing and continuous function. Since s is also increasing and continuous, Eq. 20 has a unique positive solution if:

$$N \geq \frac{r}{(1-\alpha)a} s(1) = \frac{r}{(1-\alpha)a}.$$

Below this threshold it does not pay to innovate and the rate of growth is 0. Hence, growth is driven by innovation activity in the R&D sector, as in the original endogenous growth models, but innovation is accompanied by patent races, which emerge endogenously. This emergence of patent races is the main result of this paper. The effects of these patent races are examined in the following sections.

3.6 Discussion

As explained above, patent races emerge as a result of two main assumptions of the model, that innovations differ in costs and that many innovations are researched simultaneously. It should be emphasized that it is not the mere targeting of R&D at specific innovations, as in Aghion et al. (2008), that drives this result. To highlight it, note for example that in Grossman and Helpman (1991a) R&D is also targeted, but since innovations do not differ by cost, no patent race emerges.

Also, this paper demonstrates the emergence of patent races in an expanding-variety type of model. But the same results can be obtained in other types of endogenous growth models, like quality ladders, etc. To see this note that if quality increases in period 0 by s steps, from q_0 to $q_0(1+\gamma)^s$, and each of these steps requires a larger team than the previous step, patent races emerge and the results are similar.

Finally, it is easy to show that the specification (5) of the size function S is not arbitrary and is even required in order to obtain a balanced growth path. Along such a path the technology frontier must grow at a constant rate g while the size of the labor force is fixed at L. Hence it follows from Eq. 16 that the function S should be homogenous of degree -1. Hence, the size function must be the same as in (5).

4 A reduction of the scale effect

After proving the existence of equilibrium with patent races, the paper turns to examine the macroeconomic effects of such patent races, focusing first on the scale effect. In the first-generation endogenous growth models the rate of economic growth depends strongly on scale, which is incompatible with the empirical evidence, as shown by Jones (1995a) and others. This critique led to what is called second generation endogenous growth models, which try to eliminate the scale effect from the theory. This literature consists of two main lines. The first is called semi-endogenous growth models and it shows that if the cost of innovation rises with the frontier and if population is growing, the scale effect can disappear.[8] This theory is described by the size function (4) in this paper. The second theory, called Schumpeterian growth models, assumes that costs of innovations are not rising per innovation but are spread over more and more sectors, which can also reduce and even eliminate the scale effect.[9] This section shows that patent races also reduce the scale effect even for first-generation growth models and can do it significantly.

Examine first how patent races affect the scale effect in the benchmark model. Equation 20 implies that the rate of growth g depends positively on the scale of the economy N. But the next proposition shows that this effect is much smaller under patent races than in the first-generation endogenous growth models.

Proposition 2 *The scale effect in this model is described by the following derivative:*

$$\frac{dg}{dN} = \frac{a}{s(1+g) + s'(1+g)\frac{r}{1-\alpha} + \int_0^g s(1+i)\frac{d}{dg} int\left\{\frac{s(1+g)}{s(1+i)}\right\} di}. \tag{21}$$

[8] See Jones (1995b), Kortum (1997) and Segerstrom (1998).

[9] See Young (1998), Peretto (1998), Aghion and Howitt (1998), Howitt (1999), and Madsen (2008). A third attempt to eliminate the scale effect in endogenous growth models is made by Cozzi and Spinesi (2004), through information and transmission costs.

This effect is smaller than the scale effect in the first-generation models, where the size function is (3). *Furthermore, if the function s is convex, the scale effect diminishes with scale and if* $s(1+g)\underset{g\to G}{\longrightarrow}\infty$ *the scale effect is even diminishing to zero.*[10]

Proof See Appendix.

The reduction of the scale effect is due to two mechanisms. The first is increasing cost of innovation, as expressed by the second element in the denominator of (21). A higher scale induces more innovations, but it also increases the number of inventors required to find the new innovations, and this mitigates the increase in innovations. The second mechanism is through patent races, as reflected by the third element in the denominator of (21). When scale increases, and innovation is more profitable, more teams enter the race for each innovation, which also mitigates the scale effect. The reason is that as scale increases, many more inventors crowd the existing patent races and fewer inventors search for new innovations, as they become more difficult and costly. Of these two mechanisms, the first one is the strongest, as implied by Eq. 21. But remember that both mechanisms are due to the same assumption, namely that the function s is increasing, that innovations differ by cost.

Interestingly, this model of patent races also sheds some light on the semi-endogenous growth models. To show it assume that population is growing at a constant rate n, and that the size function is equal to:

$$S(j,f)=\frac{1}{af^{1-\theta}}s\left(\frac{j}{f}\right). \tag{22}$$

This function combines together our specification (5) with the Jones specification (4). It assumes, as in Jones (1995b), that the cost of a proportional increase in technology is increasing with the frontier, where θ measures this increasing difficulty. As shown in Appendix, in the analysis of Eq. 22, the rate of growth g along a balanced growth path is given by:

$$1+g=(1+n)^{\frac{1}{\theta}}.$$

Hence, the rate of growth depends only on n and on θ, but not on any parameter related to duplication or to patent races. Only the level of technology is negatively affected by patent races, as shown in the same analysis in the Appendix. Interestingly, Jones (1995b) does not model patent races and duplication explicitly, but represents them by diminishing marginal productivity of innovators and finds that it affects the rate of growth negatively. In this model duplication is explicitly modeled and it has no effect on the rate of growth.

5 Patent races, duplication and optimality

In the economy described in this paper many teams search for the same innovation, while one team alone can find it with certainty. Such duplication clearly constitutes Pareto-Inefficiency. A central planner can assign only one team for each potential innovation, and can assign all other teams to production. This reduces the R&D sector from the equilibrium size $R_0=\int_{f_0}^{f_1}S(j,f_0)n_0(j)dj$ to the required smaller size $R_0^*=\int_{f_0}^{f_1}S(j,f_0)dj$, without changing the rate of economic growth. Hence, the market equilibrium is not Pareto-Optimal. This inefficiency due to duplication of innovation in patent races has already been explored by Loury (1979), Lee and Wilde (1980) and Dasgupta and Stiglitz (1980a,b) in models of a

[10] A similar bound on the rate of innovation is assumed in Cozzi and Spinesi (2004).

single innovation. The model in this paper determines not only the sizes of patent races, but the equilibrium amount of innovations as well. Therefore, it enables us to ask an additional question: is there too much or too little innovation? This question has been studied by Stokey (1995), Jones and Williams (1998, 2000), Li (2001) and others, but not in a framework of patent races and duplication. This paper shows that there is too little innovation as a result of duplication. The intuition is clear: duplication increases the R&D sector on expense of the production sector. As a result the incentive to innovate is reduced. Hence, in this model there can be too much R&D for some innovations, where the patent race is large, but there might be too little R&D for other innovations, which are delayed to future periods.[11]

To show it formally, consider a central planner, who maximizes the discounted sum of present and future levels of aggregate consumption, as individuals are risk-neutral, using the individual intergenerational subjective discount rate r. Assume that the size requirement for innovation is given by (5). Hence, the central planner maximizes:

$$\sum_{t=0}^{\infty} \frac{L_t^{\alpha} \int_0^{f_t} X_t(j)^{1-\alpha} dj - \int_0^{f_{t+1}} X_{t+1}(j) dj}{(1+r)^t}, \tag{23}$$

given the labor constraint in the case of no duplication, namely that in each period t:

$$L_t = N - \frac{1}{a} \int_0^{\frac{f_{t+1}}{f_t}-1} s(1+i) di = 0. \tag{24}$$

Proposition 3 *The optimal growth rate of the economy consists of a constant growth rate g^*, which is determined uniquely by the following condition:*

$$\frac{rs(1+g^*)}{a} - \frac{g^* s(1+g^*)}{a} + \frac{1}{a} \int_0^{g^*} s(1+i) di = N. \tag{25}$$

The optimal growth rate exceeds the equilibrium rate: $g^* > g$.

Proof See Appendix.

There are three reasons why the optimal rate of growth exceeds the equilibrium rate. First, the R&D sector economizes by avoiding duplication. This is reflected in the third term in the LHS of (25) relative to the function R. Second, the optimal allocation takes into consideration the effect of current R&D on reducing the cost of future R&D, which is not taken into consideration in the competitive equilibrium. This is reflected in the second negative element in the LHS of (25), which does not appear in (20). Third, inventors are monopolists and they can earn higher profits than production costs. This is reflected in the element $rs(1+g)/(1-\alpha)/a$ in (20), which is larger than $rs(1+g)/a$ in Eq. 25.

We can therefore summarize that the optimal rate of growth exceeds the competitive rate of growth for three main reasons: it avoids duplication, which is caused by patent races, it takes into consideration the spillover effect, and it avoids monopoly rents. Note that this does not mean that the R&D sector is too small in this economy. It means that there is too much R&D on innovations close to the previous technology frontier f_0, and too little R&D (actually 0) on innovations beyond f_1, which should be researched in the present, if the

[11] Recently Saint-Paul (2003) has also made a distinction between innovations, but not according to cost, but according to identity of innovators: philanthropists vs. profit seekers.

optimal allocation is followed. This conclusion sheds some light on recent recommendations to subsidize commercial R&D. According to this model the lucrative innovations should not be subsidized, but rather taxed, to reduce the sizes of the patent races, which are inefficient. The only innovations that are worth subsidization are those that seem to be too difficult to find and too costly. They might be so today, but it is actually sub-optimal to delay them to the future.

6 Uncertain outcomes

The benchmark model assumes that finding an innovation by a team of the required size is certain. In this section we relax this assumption to examine how robust are the results of the paper, mainly the duplication result. If finding an innovation is uncertain, adding teams to the race might increase the probability of finding it. This section shows that even if there are some benefits to patent races, their equilibrium size is greater than optimal, so there is still duplication and inefficiency, as in the benchmark model. This section also examines another extension of the benchmark model, which is the possibility of serendipity, namely, the possibility that some innovations might be found by pure chance, without working hard to find them.

In order to analyze the case of R&D uncertainty, consider the benchmark model with the size function (5), with one addition, that success in finding innovation j in one period by a team with the required size is no longer certain and has a probability $Q > 0$. In the case of uncertainty we have to be explicit in defining the frontier and the distance from the frontier, as not all searched innovations are found. Assume that in period t an amount J_t of innovations are searched but only I_t are found. The innovations that are found are shifted to the left and added to formerly found innovations while keeping the initial ordering, so that the new frontier becomes: $f_t = f_{t-1} + I_t$. Innovations that have not been found, of which there are $J_t - I_t$, shift to the right of f_t, keeping their initial order.

Let n teams search for innovation j. As before, we assume that if more than one team finds j by the end of the period, only one of these teams finds it first, with equal probabilities. Denote the probability of a team to find the innovation first by $q(n)$. Since this probability is equal for all teams, and since these events are disjoint, the probability that any team finds the innovation first is $nq(n)$. Since this probability is also equal to $1 - (1 - Q)^n$, the probability that a team finds the innovation first is:

$$q(n) = \frac{1 - (1 - Q)^n}{n}. \tag{26}$$

This probability is equal to Q when there is only one team, $n = 1$, and it can be shown that this probability is diminishing with n, with the number of teams in the patent race, and it even converges to zero as the number of teams tends to infinity. The following proposition describes the equilibrium in the economy with uncertain innovations.

Proposition 4 *Equilibrium is a balanced growth path with a constant rate of growth g and a constant rate of innovation h, $h > g$, so that in each period t : $I_t = f_t g$ and $J_t = f_t h$. The size of a patent race for innovation $j = f_t(1 + i)$, which is denoted $n(i)$, is the largest integer n that satisfies:*

$$q(n) = \frac{1 - (1 - Q)^n}{n} \geq Q \frac{s(1 + i)}{s(1 + h)}. \tag{27}$$

The size of the patent race at the marginal innovation, $n(h)$, is equal to 1. The equilibrium rate of innovation h is determined by:

$$N = \frac{rs(1+h)}{aQ(1-\alpha)} + R_Q(h), \tag{28}$$

where the function R_Q describes the size of the R&D sector: $R_Q(h) = a^{-1}\int_0^h n(i)s(1+i)di$. The equilibrium rate of growth g is given by:

$$g = \int_0^h [1-(1-Q)^{n(i)}]di. \tag{29}$$

Proof See Appendix.

Proposition 4 shows that the main result of the paper, namely that patent races emerge when the cost of innovation is increasing and when many innovations are searched simultaneously, holds in this case as well, when invention is uncertain. To see it, note that the race to the first innovation is the largest and its size is $n(0)$. This size satisfies: $s(1+h) \geq Q/q[n(0)]$. Clearly, if the rate of innovation h is sufficiently high and if the size function s is sufficiently increasing, this patent race can be quite large.

We next show that there is duplication in this model and that the equilibrium described in Proposition 4 is not Pareto-Efficient, even if increasing the size of the patent race can increase the probability of finding the innovation. To see this we ask the following question: how can the economy maximize the rate of growth with a given size of the R&D sector, namely without affecting current output. In other words, we ask whether there can be a Pareto improvement. Formally, we address the following maximization problem, where k denotes the desired rate of innovation and $m(i)$ denotes the desired size of the patent race:

$$\max \int_0^k [1-(1-Q)^{m(i)}]di \text{ s.t.} \left[\frac{1}{a}\int_0^k m(i)s(1+i)di = R\right]. \tag{30}$$

Proposition 5 *The optimal allocation (30) is different than the equilibrium allocation described in Proposition 4. It enables a greater rate of innovation: $k > h$, which means that on average it requires a smaller size of patent races.*

Proof See Appendix.

Proposition 5 therefore implies that even in this case of uncertain outcomes, the equilibrium is Pareto-inefficient. Patent races are too big, so even if it increases the probability of finding innovations, too few innovations are searched, so the overall rate of growth is lower. Hence, the results of Sect. 5 apply here as well.

Finally, we can approach the issue of serendipity, namely the possibility that some innovations are found accidentally, without effort and without directed research. Assume that each innovation has a probability P in each period of being found by chance. Assume that alternatively the innovation can be searched for by teams, as in the benchmark model. If the innovation is found by chance, the teams fail. Hence, their probability of success is $Q = 1 - P$. Note though, that this case is simpler than the case analyzed above in this section, since the probability of success does not depend on the number of teams. Hence, the probability to be first is simply Q/n, where n is the number of teams. It is easy to see that the main results of the model hold in this case as well, as long as P is not equal to 1, which is not realistic.

7 Risk aversion and concentration of R&D

This section presents an extension to the benchmark model, where risk neutrality is replaced by risk aversion. This is very relevant for such a model, since participating in a patent race is risky, as one team wins big, while the others get zero income. Clearly such risk creates a demand for insurance. This insurance can be supplied through a number of mechanisms. One possible mechanism is venture capital. The financing of R&D is undertaken by large financiers who can average risks over many research teams, including over teams that compete with one another in the same patent race. Another mechanism is cooperation between competing teams, usually through firms, that hire the various teams together and internalize the patent race, or part of it, within the firm. There is much evidence on such R&D cooperation.[12]

As shown below, such cooperation and concentration of the patent race within a firm, or insurance by venture capitalists, does not reduce the number of teams in a patent race and does not eliminate duplication, but rather increases it. The reason is that risk reduction increases the incentive for new teams to enter the race. This section shows that insurance increases duplication and as a result it reduces the overall rate of growth.

To formalize the analysis, the benchmark model is slightly changed with respect to individuals, but the production and innovation parts remain unchanged. Assume that the population consists of two types of people. People of the first type live one period each, in non-overlapping generations and the size of each generation is N. Each individual derives the following utility from consumption:

$$u = \ln c.$$

People of this type can work either in production, in which case each supplies 1 unit of labor, or in R&D, in which case each still works a share e of her time as a production worker, where $e < 1$. This assumption is made to avoid zero income if they fail in the patent race.[13] The second type of people is a group of size zN of risk neutral infinitely lived individuals, with intertemporal utility equal to (7). They operate as financiers.

We describe first the equilibrium in this economy without cooperation between teams and then introduce cooperation within firms. Note that the main result of such cooperation is insurance, so the analysis is similar in the case of venture capital. A new team, which enters the race in period 0 to find innovation j, compares the expected utility of its members to the expected utility of production workers. If the number of teams already in the race is $\bar{n}_0(j)$ the new team enters the race if an only if:

$$\frac{1}{1+\bar{n}_0(j)} \ln\left(\frac{v_0}{S(j, f_0)} + ew_0\right) + \left(1 - \frac{1}{1+\bar{n}_0(j)}\right) \ln(ew_0) \geq \ln(w_0). \tag{31}$$

Subtracting the logarithm of wages from both sides yields the following entry condition:

$$\frac{1}{1+\bar{n}_0(j)} \ln\left(\frac{v_0}{w_0 S(j, f_0)} + e\right) + \left(1 - \frac{1}{1+\bar{n}_0(j)}\right) \ln e \geq 0. \tag{32}$$

We can therefore conclude that the overall size of the patent race $n_0(j)$ is equal to $1+\bar{n}_0(j)$, where $\bar{n}_0(j)$ is the largest integer that satisfies (32).

Note that as the required size of the innovation team S becomes larger, the size of the patent race declines. Hence, as j increases, patent races become smaller, until they reach the

12 See a survey of empirical studies of such cooperation in Cozzi and Tarola (2006).

13 Reducing the life horizon from infinity to one period already increases the risk of patent races.

size 1. Then there is only one team that participates in the race, there is no risk involved and the entry condition (32) becomes:

$$\frac{v_0}{w_0 S(j, f_0)} \geq 1 - e.$$

Hence, the marginal innovation that is searched in period 0, which is the new technology frontier f_1, is described by the following condition:

$$S(f_1, f_0) = \frac{v_0}{(1-e)w_0} = \frac{1-\alpha}{1-e}\frac{1}{f_0}\sum_{t=1}^{\infty}\frac{L_t}{(1+r)^t}. \tag{33}$$

Given (32) and (33), we can solve the rational expectation equilibrium, in a similar way to the solution of the benchmark model. Similarly, if the required size follows (5), the equilibrium is a balanced growth path.

We next turn to examine concentration of research teams within firms, which creates insurance, since the returns from innovation are divided by all teams in the firm, if one of the teams wins the race. In that case, if m is the number of teams working in the firm on innovation j, and if $\bar{n}_0(j)$ is the number of the other teams working on innovation j, the expected utility of each researcher in the firm is described by:

$$\frac{m}{m+\bar{n}_0(j)}\ln\left(\frac{v_0}{mS(j, f_0)} + ew_0\right) + \left(1 - \frac{m}{m+\bar{n}_0(j)}\right)\ln(ew_0). \tag{34}$$

The firm maximizes this expected utility with respect to m over all integers greater or equal than 1. The firm enters the patent race if this maximum expected utility is greater than the utility of a producer, $\ln(w_0)$.

Clearly the expected utility of a researcher in a firm, described by (34), is larger than in a single team, described by (31), since the firm is not restricted to set m to be 1, as in (31), but it can increase expected utility by increasing m, which increases insurance. Hence, as (34) exceeds (31), we get that $\bar{n}_0(j)$ for firms is larger than for single teams. As a result, the size of the patent race tends to be larger when there is concentration of research teams within firms. The intuition for that is simple. Collaboration creates insurance, so research teams are willing to take more risk and therefore more research teams crowd each race.

We next examine how much innovation is performed under collaboration. Note that entry continues until we have a firm of one team only and no risk at all, so that the entry condition is similar to the case of no collaboration and is given by the same condition (33). But under collaboration the sizes of the future production sectors $\{L_t\}$ are smaller, since patent races are larger and the R&D sector is larger as well. Thus, the same condition (33) implies a smaller rate of innovation, namely f_1 is smaller under collaboration, or insurance, than without it. Hence, collaboration increases patent races and the R&D sector, but reduces the amount of innovations and thus the rate of growth. Hence, collaboration increases duplication and as a result it reduces economic growth.[14]

[14] Note that this result holds with respect to patent race risk. If the innovation is risky, as in Sect. 6, this result becomes ambiguous. The reason is that in the marginal innovation there is no patent race risk, but it is still risky. Insurance reduces this risk on the one hand, but it reduces future production sectors and reduces the incentive to innovate on the other hand. Hence, its full effect is ambiguous.

8 Summary and conclusions

This paper examines the conditions that can lead to patent races and R&D duplication in a growing economy, which has infinite potential innovations. The paper shows that such a phenomenon can emerge when two conditions are met and interact with each other. The first is that potential innovations differ in costs, and the second is that finding innovations takes time. When these two conditions hold, innovations of different costs are searched simultaneously, so that patent races and duplication emerge. The paper further shows that duplication reduces the scale effect and that it leads to Pareto-inefficiency.

But the paper's results raise a number of interesting issues for future research and for policy. The first issue is the difficulty of finding empirical support to duplication in patent races. As shown in Sect. 7 many innovations are searched by large firms instead of competitive patent races, due to risk aversion. It means that many patent races are not observed empirically, since they are internalized within large firms. Despite this, duplication in the search for these innovations is even larger than in competition, as shown in Sect. 7. Thus, empirical research on duplication seems to be an intriguing area for future research.

A second issue raised by the paper is related to the welfare analysis of R&D. It is shown that researchers tend to crowd the more promising research strategies, with lower costs, or with higher probabilities of success, while there is too little R&D at high cost projects, or projects with low probabilities of success. This result should inspire us to think more seriously on the policy of R&D subsidization, which is quite common. The paper shows that subsidization should be targeted at the innovators that seem to be less promising, who travel the less frequented ways, who try strategies with lower chances of success, who deviate from the crowd and do less standard and more risky research. The big practical question is of course how to identify such researchers?

Acknowledgments This paper is dedicated to my uncle, the inventor Jack Drillick (1925–1991). In working on this paper I have benefited significantly from discussions with Federico Etro and I am very grateful to him. I am also indebted to the editor of the journal and to three anonymous referees, to Daron Acemoglu, Gadi Barlevy, Jacques Crémer, Guido Cozzi, Paolo Giordani, Dani Givon, Chad Jones, Shaul Lach, Albert Marcet, Motty Perry, Luca Zamparelli and especially to Avi Simhon for very valuable comments. I also thank the Hebrew University Research Fund and the Aaron and Michael Chilewich Chair for financial assistance. Remaining errors are all mine.

Appendix

Proof of Proposition 1 If the size of the required research team is given by (5), the equilibrium rate of technical change is derived from (16) and is given by:

$$s\left(\frac{f_1}{f_0}\right) = (1-\alpha)a\sum_{t=1}^{\infty}\frac{L_t}{(1+r)^t}. \tag{A.1}$$

The equilibrium size of the R&D sector is therefore derived from (18):

$$R_0 = \int_{f_0}^{f_1}\frac{1}{af_0}s\left(\frac{j}{f_0}\right)\operatorname{int}\left\{\frac{s(f_1/f_0)}{s(j,f_0)}\right\}dj. \tag{A.2}$$

Let us use the notation: $i = j/f_0 - 1$, and $g = f_1/f_0 - 1$. Clearly g is the rate of growth of output in period 1. Using this notation (A.1) becomes:

$$s\,(1+g) = (1-\alpha)a \sum_{t=1}^{\infty} \frac{L_t}{(1+r)^t}, \tag{A.3}$$

and (A.2) boils down to:

$$R_0 = \frac{1}{a} \int_0^g s(1+i) int \left\{ \frac{s(1+g)}{s(1+i)} \right\} di. \tag{A.4}$$

The right hand side of (A.4) is a function of g, and it is a continuous and increasing function, as can be shown easily. We therefore denote this function by R:

$$R(g) = a^{-1} \int_0^g s(1+i)\text{int} \left\{ \frac{s(1+g)}{s(1+i)} \right\} di.$$

Hence, the size of the R&D sector in period 0 is increasing with growth in period 1.

Utilizing this result and the labor market equilibrium condition (19) yields:

$$L_0 = N - \frac{1}{a} R \left\{ s^{-1} \left[(1-\alpha)a \sum_{t=1}^{\infty} \frac{L_t}{(1+r)^t} \right] - 1 \right\}. \tag{A.5}$$

Hence, in this case the equilibrium size of the production sector depends only of future sizes of this sector and not on the level of technology. We can therefore write the dynamic condition of the economy (A.5) as:

$$L_0 = \phi\left(\sum_{t=1}^{\infty} \frac{L_t}{(1+r)^t} \right).$$

The function ϕ is decreasing.

We next show that the only rational expectations equilibrium which solves (A.5) is a constant size of the production sector L over time. Define the following value function:

$$V_1 = \sum_{t=1}^{\infty} \frac{L_t}{(1+r)^t}.$$

Then, according to (A.5), $L_0 = \phi(V_1)$. It follows that:

$$V_0 = \frac{V_1 + L_0}{1+r} = \frac{V_1 + \phi(V_1)}{1+r}. \tag{A.6}$$

It is clear from this equation and from ϕ being decreasing, that the effect of the future value on the present value satisfies:

$$\frac{\partial V_0}{\partial V_1} < \frac{1}{1+r}.$$

Hence the only rational expectations equilibrium to (A.6) is a constant level of V. If V is constant over time, so is L. If L is constant over time Eq. A.3 implies that the rate of growth is constant over time as well and is given by:

$$s\,(1+g) = (1-\alpha)aL\frac{1}{r}.$$

From this equation we can calculate the equilibrium size of the production sector, and by adding the size of the R&D sector we get:

$$s(1+g)\frac{r}{(1-\alpha)a}+R(g)=L_0+R_0=N.$$

This is Eq. 20. □

Proof of Proposition 2 A straight derivation of Eq. 20 yields Eq. 21. To see the reduction of the scale effect, note first, by comparing (3) and (5), that in the classical R&D growth models the function s is constant and equal to 1. As a result, in these models the derivative of g with respect to N should be a. Note that in the denominator of (21) we have: $s(1+g)>s(1)=1, s'(1+g)>0$, and also:

$$\int_0^g s(1+i)\frac{d}{dg}\text{int}\left\{\frac{s(1+g)}{s(1+i)}\right\}di>0.$$

Hence, the denominator of (21) is larger than 1, even at $g=0$, and hence dg/dN is smaller than a. As a result, the scale effect is smaller in a model with patent races.

If s is convex, $s'(1+g)$ rises with scale, as g increases with scale, hence it follows from (21) that the scale effect is diminishing. If the rate of growth is bounded, namely if $s(1+g)\underset{g\to G}{\longrightarrow}\infty$, then the scale effect even converges to zero as scale increases. □

Analysis of Eq. 22

Assume that the economy is on a balanced growth path, where technology grows at a rate g and producers are a constant share of population, so they grow at the rate n as well. The equilibrium condition (16) becomes:

$$S(f_1,f_0)=\frac{1}{af_0^{1-\theta}}s\,(1+g)=\frac{(1-\alpha)L_0}{f_0}\frac{1+n}{r-n}.$$

Hence:

$$\frac{L_0}{f_0^\theta}=\frac{s(1+g)(r-n)}{(1-\alpha)(1+n)a}. \tag{A.7}$$

Since the RHS of (A.7) is constant over time, so is the LHS and the rate of growth of innovation is:

$$1+g=(1+n)^{\frac{1}{\theta}}.$$

Hence, the steady state rate of growth in the economy does not depend on the s function, namely on the amount of duplication.

Calculation of the size of the R&D sector that follows the proof of Proposition 1 yields:

$$R_0=f_0^\theta R(g). \tag{A.8}$$

Equations A.7 and A.8 together yield the following labor market equilibrium condition:

$$N_0=L_0+R_0=f_0^\theta\left[R(g)+\frac{s(1+g)(r-n)}{a(1-\alpha)(1+n)}\right].$$

This means that although duplication has no effect on the steady state rate of growth, it has an effect on the level of technology relative to the size of population along the balanced growth path. It is clear from the above condition that as the function s increases f_0 becomes smaller, since duplication increases the R&D sector.

Proof of Proposition 3 Maximization of (23) subject to the resource constraints (24) leads to the following Lagrangian:

$$\sum_{t=0}^{\infty} \frac{L_t^{\alpha} \int_0^{f_t} X_t(j)^{1-\alpha} dj - \int_0^{f_{t+1}} X_{t+1}(j) dj + q_t \left[N - a^{-1} \int_0^{\frac{f_{t+1}}{f_t}-1} s(1+i) di - L_t \right]}{(1+r)^t}$$

The first order conditions are:

$$-1 + \frac{(1-\alpha) L_{t+1}^{\alpha} X_{t+1}(j)^{-\alpha}}{1+r} = 0, \tag{A.9}$$

for the capital goods at $0 \leq j \leq f_{t+1}$,

$$q_t = \alpha L_t^{\alpha-1} \int_0^{f_t} X_t(j)^{1-\alpha} dj, \tag{A.10}$$

for labor and:

$$-X_{t+1}(f_{t+1}) + \frac{L_{t+1}^{\alpha} X_{t+1}(f_{t+1})^{1-\alpha}}{1+r} - \frac{q_t}{a} s\left(\frac{f_{t+1}}{f_t}\right) \frac{1}{f_t} + \frac{q_{t+1}}{(1+r)a} s\left(\frac{f_{t+2}}{f_{t+1}}\right) \frac{f_{t+2}}{f_{t+1}^2} = 0, \tag{A.11}$$

for the technology frontier f_{t+1}.

From (A.9) we get:

$$X_t(j) = L_t \left[\frac{1-\alpha}{1+r} \right]^{\frac{1}{\alpha}}.$$

Substituting in (A.10) we get:

$$q_t = \alpha f_t \left[\frac{1-\alpha}{1+r} \right]^{\frac{1-\alpha}{\alpha}}.$$

Substituting this Lagrange multiplier and the amount of X_t in (A.11) we get:

$$\frac{a}{1+r} L_{t+1} - s\left(\frac{f_{t+1}}{f_t}\right) + \frac{1}{1+r} s\left(\frac{f_{t+2}}{f_{t+1}}\right) \frac{f_{t+2}}{f_{t+1}} = 0.$$

Substituting in this condition the constraint (24) and denoting $f_{t+1}/f_t = 1 + g_{t+1}$ we get:

$$s(1+g_{t+1}) = \frac{1+g_{t+2}}{1+r} s(1+g_{t+2}) + \frac{aN}{1+r} - \frac{1}{1+r} \int_0^{g_{t+2}} s(1+i) di. \tag{A.12}$$

Note that if in the steady state $g < r$, as shown below, then the derivative of g_{t+1} with respect to g_{t+2} is smaller than 1. Due to the transversality condition the dynamic path of g cannot explode and thus (A.12) implies that the optimal path is a constant rate of growth g^*, which is given by:

$$(r-g^*) s(1+g^*)/a + a^{-1} \int_0^{g^*} s(1+i) di = N.$$

This is Eq. 25. Note that r exceeds g^*, since otherwise the R&D sector exceeds overall population and there is no production, which is sub-optimal. Since $g^* < r$, the left hand side of (25) is increasing in g^*, and hence the optimal rate of growth g^* is unique.

Next compare the equilibrium rate of growth g with the optimal rate of growth g^*, by comparing (20) and (25) as functions of the rate of growth. Note first, that the size of the R&D sector is smaller in optimum due to lack of duplication:

$$a^{-1}\int_0^g s(1+i)di < R(g).$$

Also: $\frac{r}{(1-\alpha)}s(1+g) > (r-g)s(1+g)$.

Hence, the LHS in (20) is larger than the LHS in (25) for any g. Therefore, the optimal growth rate g^* exceeds the equilibrium growth rate g. □

Proof of Proposition 4 Profits of producers and of inventors and value of innovations are the same as in the benchmark model. The expected income of an innovator in period 0, who works on innovation j, for which n teams are searching, is:

$$\frac{q(n)}{S(j, f_0)}v_0 = q(n)\frac{af_0}{s(1+i)}v_0,$$

where $i = j/f_0 - 1$. Hence, teams are entering as long as this income exceeds wages in the production sector, namely as long as:

$$q(n) \geq s(1+i)\left[a(1-\alpha)\sum\nolimits_{t=1}^{\infty}\frac{L_t}{(1+r)^t}\right]^{-1}. \tag{A.13}$$

Define $n(i)$ to be the largest integer that satisfies (A.13). It is easy to see that $n(i)$ is decreasing in i. It falls until it reaches 1. Then the probability of finding the innovation is Q. Hence, similar to the benchmark model, the rate of innovation h_0 is given by:

$$Q = s(1+h_0)\left[a(1-\alpha)\sum\nolimits_{t=1}^{\infty}\frac{L_t}{(1+r)^t}\right]^{-1}.$$

Substituting this condition in (A.13) we get that $n(i)$ is the largest integer n that satisfies:

$$q(n) \geq Q\frac{s(1+i)}{s(1+h_0)}.$$

It easily follows that the size of the R&D sector in period 0 is:

$$R_0 = a^{-1}\int_0^{h_0} n(i)s(1+i)di = R_Q(h_0).$$

In a similar way to the proof of Proposition 1 it can be shown that the economy follows a balanced growth path with fixed labor L and fixed h and that:

$$L = \frac{rs(1+h)}{Qa(1-\alpha)}.$$

Thus the equilibrium in the labor market, $N = L + R$, leads to the equilibrium condition (28). The calculation of the equilibrium rate of growth (29) is straightforward. □

Proof of Proposition 5 The Lagrangian of (30) is:

$$\int_{i=0}^{k} \left[1-(1-Q)^{m(i)}\right] di + \lambda \left[R - a^{-1} \int_{i=0}^{k} m(i)s(1+i)di\right].$$

The solution of the optimization therefore looks, for each i, for the integer m that yields the highest value of:

$$1-(1-Q)^m - \frac{\lambda}{a} ms(1+i), \tag{A.14}$$

and at the first order condition with respect to k, which is a continuous variable:

$$1-(1-Q)^{m(k)} - \frac{\lambda}{a} m(k)s(1+k) = 0.$$

From this FOC we get:

$$\frac{\lambda}{a} = \frac{1-(1-Q)^{m(k)}}{m(k)s(1+k)}. \tag{A.15}$$

Substituting in (A.14) for k we get that $m(k)$ is the m that maximizes:

$$1-(1-Q)^m - m\frac{1-(1-Q)^{m(k)}}{m(k)}.$$

It can be shown that the solution to this maximization is: $m(k) = 1$. Namely, the marginal patent race has a size 1, just as in equilibrium. Substituting in (A.15) and back in (A.14) we get that $m(i)$ is the integer m that yields the highest value for:

$$1-(1-Q)^m - Qm\frac{s(1+i)}{s(1+k)}. \tag{A.16}$$

Note that if m as a continuous variable, (A.16) describes a function that is concave, it is equal to 0 for $m = 0$ and its derivative at 0 is: $-\ln(1-Q) - Qs(1+i)/s(1+k)$, which is positive. Hence, this function is first rising and is positive at $m > 0$, and then at some point it declines and reaches 0 at some finite m.

Next, we prove the proposition by negation, assuming that $k \leq h$. According to Proposition 4, $n(i)$ is the largest integer n that satisfies:

$$1-(1-Q)^n - Qn\frac{s(1+i)}{s(1+h)} \geq 0.$$

If $k \leq h$, then $n(i)$ is larger than the largest integer n that satisfies:

$$1-(1-Q)^n - Qn\frac{s(1+i)}{s(1+k)} \geq 0.$$

Note that this is the same function as in (A.16). Since this function is concave and positive after 0, it reaches a maximum before it returns to zero. It therefore follows that: $m(i) \leq n(i)$ for all i and the inequality is strict for some i. Substituting this result at the constraint of the size of the R&D sector in (30) we get that $k > h$. This is a contradiction. □

References

Aghion, P., Dewatripont, M., & Stein, J. C. (2008). Academic freedom, private-sector focus, and the process of innovation. *RAND Journal of Economics, 39*(3), 617–635.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica, 60*(2), 323–351.

Aghion, P., & Howitt, P. (1998). *Endogenous growth theory*. Cambridge, MA: MIT Press.

Cozzi, G., & Spinesi, L. (2004). Information transmission and the bounds to growth. *Topics in Economic Analysis and Policy, BEPress, 4*(1), Article 6.

Cozzi, G., & Tarola, O. (2006). R&D cooperation, innovation and growth. *Journal of Theoretical and Institutional Economics, 162*(4), 683–701.

Dasgupta, P., & Stiglitz, J. (1980a). Uncertainty, industrial structure, and the speed of R&D. *Bell Journal of Economics, 11*(1), 1–28.

Dasgupta, P., & Stiglitz, J. (1980b). Industrial structure and the nature of innovative activity. *Economic Journal, 90*(358), 266–293.

Etro, F. (2002). *Patent races and Schumpeterian growth*, Mimeo.

Evenson, R. E., & Kislev, Y. (1976). A stochastic model of applied research. *Journal of Political Economy, 84*(2), 265–282.

Grossman, G. M., & Helpman, E. (1991a). Quality ladders in the theory of growth. *Review of Economic Studies, 58*(1), 43–61.

Grossman, G. M., & Helpman, E. (1991b). *Innovation and growth in the global economy*. Cambridge, MA: MIT Press.

Howitt, P. (1999). Steady endogenous growth with population and R&D Inputs growing. *Journal of Political Economy, 107*(4), 715–730.

Irwin, D. A., & Klenow, P. J. (1996). High-tech R&D subsidies: Estimating the effects of Sematech. *Journal of International Economics, 40*(3–4), 323–344.

Jones, C. I. (1995a). Time series tests of endogenous growth models. *Quarterly Journal of Economics, 110*(2), 495–525.

Jones, C. I. (1995b). R&D-based models of economic growth. *Journal of Political Economy, 103*(4), 759–784.

Jones, C. I., & Williams, J. C. (1998). Measuring the social return to R&D. *Quarterly Journal of Economics, 113*(4), 1119–1135.

Jones, C. I., & Williams, J. C. (2000). Too much of a good thing? The economics of investment in R&D. *Journal of Economic Growth, 5*(1), 65–85.

Kortum, S. (1997). Research, patenting, and technological change. *Econometrica, 65*(6), 1389–1419.

Lee, T., & Wilde, L. L. (1980). Market structure and innovations: A reformulation. *Quarterly Journal of Economics, 94*(2), 429–436.

Li, C. W. (2001). On the policy implications of endogenous technological progress. *The Economic Journal, 111*(471), 164–179.

Loury, G. C. (1979). Market structure and innovation. *Quarterly Journal of Economics, 93*(3), 395–410.

Madsen, J. B. (2008). Semi-endogenous versus Schumpeterian growth models: testing the knowledge production function using international data. *Journal of Economic Growth, 13*(1), 1–26.

Peretto, P. (1998). Technological change and population growth. *Journal of Economic Growth, 3*(4), 283–311.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5, pt. 2), S71–S102.

Saint-Paul, G. (2003). Growth effects of non proprietary innovation. *Journal of the European Economic Association, 1*(2–3), 429–439.

Segerstrom, P. S. (1998). Endogenous growth without scale effects. *The American Economic Review, 88*(5), 1290–1310.

Segerstrom, P. S., Anant, T. C. A., & Dinopoulos, E. (1990). A Schumpeterian model of the product life cycle. *The American Economic Review, 80*(5), 1077–1091.

Stokey, N. L. (1995). R&D and economic growth. *Review of Economic Studies, 62*(3), 469–489.

Young, A. (1998). Growth without scale effects. *Journal of Political Economy, 106*(2), 41–63.