Innovations, Patent Races and Endogenous Growth

Joseph Zeira

Hebrew University of Jerusalem, Harvard University and CEPR

April 2003*

Abstract

This paper presents a model of innovations and economic growth, in which patent rates emerge endogenously, as a result of two assumptions: first, R&D is innovation-specific, second, marginal cost of innovation is increasing. The paper then examines the effects of patent races on growth, welfare, and the market structure of R&D, and derives three main results. The first is that patent races reduce significantly the effect of scale on growth. The second result is that R&D is Pareto-inefficient, as too many researchers look for the easy innovations, while too few search for the difficult ones. The third result is that risk aversion leads to concentration of R&D in few firms, to reduce risk of patent race. Interestingly this does not contribute to growth but rather to more duplication.

JEL Classification: O31, O40.
Keywords: Endogenous Growth, Innovations, Patent Races, R&D Sector.

Address:
Kennedy School of Government
Harvard University
69 JFK Street
Cambridge, MA 02138
E-mail: joseph_zeira@ksg.harvard.edu

* In working on this paper I have benefited significantly from discussions with Federico Etro. I am very grateful to him. I am also indebted to Daron Acemoglu, Gadi Barlevy, Dani Givon, Chad Jones, Shaul Lach, and Avi Simhon for valuable comments. Remaining errors are all mine.

Innovations, Patent Races and Endogenous Growth

1. Introduction

This paper presents a model of innovations and economic growth, in the spirit of the R&D-based endogenous growth models, which examines how patent races emerge, what is their market structure, and what is their effect on economic growth. A patent race is defined here as a race between many research teams, which try to find the same specific innovation, and the first to find it gets the patent rights. Hence, such patent races can emerge only if research is directed, or innovation-specific. The paper shows that if research is directed, and if innovations differ by the return to innovator, patent rates emerge, and their size is positively related to the return from the innovation. The model then derives a number of results. First, it shows that patent races reduce significantly the effect of scale on economic growth. Second, competitive R&D is Pareto-inefficient, due to duplication. Furthermore, although there is too much R&D for innovations with high returns, there is too little R&D for innovations with low returns. Finally, the paper shows that concentration of much R&D by few research firms can be a result of risk aversion.

The model presents an economy, which grows through innovations that increase the productivity of workers.[1] Innovations differ by difficulty: the amount of researchers needed to find each specific innovation is increasing within each period. This assumption gives rise to patent races. In order to have sustained growth over time, we assume that the costs of the next innovations in line become lower in the next period, since the economy has accumulated more knowledge and the time that has passed enables this knowledge to

[1] This paper therefore models technologies as 'process innovation,' as in Dasgupta and Stiglitz (1980 a, b). Other modeling devices, such as variety, or quality ladders, lead to similar results.

sink in. In other words, we assume that innovation requires two inputs, innovators and waiting time. Time between innovations is inherent to the process of research, as new innovations build on previous ones. Thus, for example, the internal combustion engine could not be invented at the same time as the steam engine, but only years later, after the necessary experience was accumulated. Sometimes the need for an innovation appears only after some time and experience, like air bags for cars. The rest of the model is similar to the original R&D-based models, where individuals choose between production and innovation, thus equalizing expected utilities across the two sectors.

As the scale of the economy – the size of population – increases, the gains from each invention increase, which induces entry to the R&D sector. Innovators can either begin research on a new innovation, which is more costly, or join the patent race in one of the less costly innovations, which many of them do. This significantly reduces the effect of scale on innovation and economic growth. This result is important, since the empirical evidence shows that scale indeed has a small effect on innovation and growth, as shown by Jones (1995a) and others. The explanation this paper offers to these findings is that increased R&D activity leads mainly to larger patent races and thus to more duplication, with only a small or no increase in innovations.

The second result is that equilibrium is inefficient, since there are too many innovators, who search for the low-cost innovations. The paper shows that this waste due to duplication holds even if there are multiple research strategies for each innovation, so that race participants can follow different methods. Then innovators tend to crowd the more promising research strategies, which have higher probability of success. It is shown that while there is too much R&D in the lucrative strategies, there is too little R&D in the less promising and more difficult lines of research.

Finally, the paper discusses the market structure of patent races under risk aversion. Since participating in a patent race is risky, there is an incentive to cooperate by sharing the gains from successful innovation, which reduces risk. Such cooperation is hard to achieve voluntarily, but it can emerge when a single firm employs all the teams that search for an innovation, thus internalizing the patent race. Such a firm hires many teams, even if one team alone can find the innovation, in order to increase the probability of success and to deter potential competition. It is shown that such concentration leads to more R&D than in competitive patent races but to less growth.

This paper is related to the literature of endogenous growth, which began with Romer (1986) and Lucas (1988), and especially to the R&D-based growth models of Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992). These models have been very successful in using the increasing returns to scale of innovations to explain the high rates of global economic growth over the recent two centuries.[2] These models also faced criticism, since the scale effect they predict cannot be supported by the data, as shown by Jones (1995a), Segerstrom (1998) and others.

Several attempts were made, as a result of this critique, to reduce the scale effect in the endogenous growth models. Jones (1985b), Kortum (1997), and Segerstrom (1998) assume that the difficulty of innovation rises over time. This assumption eliminates the scale effect, but also leads to another problematic result, that economic growth is not sustainable without population growth. Young (1988) and Howitt (1999) assume that the

[2] In the years 1820-1992 world GDP increased 40 fold, and GDP per capita increased 8 fold. US GDP per capita increased 17 fold. See Maddison (1995).

scale of production is bounded within each sector, and increased scale translates only into more sectors. Hence, scale has no effect over growth.

This paper follows a different route, by focusing on the role of patent races. It therefore explores a much more micro-oriented analysis of the innovation process, of how innovators compete with one another, on how they might use similar or different strategies in their research etc. Hence, this paper has wider results than just reduction of scale effects, as it studies patent races and their effects. The potential role of patent races has already been acknowledged by Stokey (1995) and Jones and Williams (1998, 2000), but without explicitly modeling patent races. Other papers, that explicitly describe patent races, like Segerstrom, Anant and Dinopoulos (1990) and Etro (2002), use it to analyze very different issues.

The paper is also related to the microeconomic literature on duplication in innovation, as in Dasgupta and Stiglitz (1980 a, b). This model embeds this literature within the endogenous growth framework. The original endogenous growth models assume that all potential innovations are equally costly, so that innovators can always turn to a new innovation instead of duplicating. This model restores duplication by assuming that the cost of innovation is increasing.

The paper is organized as follows. Section 2 presents the benchmark model and Section 3 analyzes the equilibrium. Section 4 examines the effect of scale and Section 5 discusses the inefficiency of equilibrium. Section 6 extends the analysis to multiple research strategies. Section 7 examines how risk aversion leads to concentration of R&D by large research firms and Section 8 concludes.

2. The Model

Consider an economy in a discrete time framework. There is a single final good in the economy, which is produced by labor. Output of the final good in period t, Y_t, depends on the labor input in the production sector L_t through the following production function:

(1) $$Y_t = a_t L_t .$$

The productivity of labor a_t contains all available technologies at time t. This productivity rises from one period to the next through new innovations, which are therefore called in the literature "process innovations." Innovations are infinitesimal. Each innovation increases productivity of each worker by an amount, which is proportional to last period productivity.[3] Innovation j increases productivity by ba_{t-1}, where $b > 0$. Thus, if I_t innovations are found in period t, the change in productivity over time is described by:

(2) $$a_t = a_{t-1}(1 + I_t b) .$$

We next describe the R&D activity or the search for innovations. Innovations are searched and found by research teams, or by innovation teams. Each team searches for a specific potential innovation. Hence, unlike many endogenous growth models research is directed toward specific innovations. The potential innovations in period t are ordered on the real line $[0, \infty)$. Each innovation j requires a different size of a team, which is denoted by $s(j)$. The function s is increasing, namely potential innovations are ordered by increasing difficulty. Put differently, marginal productivity of innovators is diminishing. The size $s(j)$ can be greater than 1, if more than 1 innovator is needed, or smaller than 1, if an innovator can find j within less than a period of time. The function s is shown in Figure 1.

[Insert Figure 1 here]

A single research team is sufficient to find a potential innovation.[4] It is possible though that the number of teams that search for the innovation is not equal to 1. First, it is possible that a team of size $s(j)$ searches for the innovation part time, q, $q \le 1$. In this case it has probability q of finding the innovation.[5] Second, it is possible that more than one team searches for the same innovation, but only one team finds it first. This first team gets the patent rights and sells the use of the innovation to producers of the final good. Patent rights hold for one period only and from the next period on, innovations become public knowledge. It is assumed that the probability of being first is equal to all, conditional on whether the team searches full or part time. Denote the number of teams that search for innovation j in period t by $n_{j,t}$, which can be any real number due to part time search. Then, the probability of success for each team is:

$$(3) \qquad P_{j,t} = \begin{cases} \dfrac{1}{n_{j,t}} & \text{if team searches full time,} \\ \dfrac{q}{n_{j,t}} & \text{if team searches } q \text{ time and } n_{j,t} \ge 1, \\ q & \text{if team searches } q \text{ time and is the only team.} \end{cases}$$

Finally assume that after some time passes, when more innovations are found and more knowledge is accumulated, potential innovations become easier to find. Formally: after finding innovations $[0, I_t)$ in period t, the remaining innovations shift in period t+1

[3] This proportionality assumption is common to all endogenous growth models. It reflects what is sometimes called the "spillover effect" of innovations.

[4] We therefore implicitly assume that there is only one way to search for each innovation. In Section 6 we replace this assumption with a more realistic one of multiple search strategies.

[5] Alternatively we can assume that an innovation can be found only if the team works full time. In that case the analysis involves only patent races of integer size and not continuous size as in this paper. The results are the same, and a version of the paper with this assumption is available from the author.

to the origin and the labor cost of finding the next innovation decreases from $s(I_t)$ to $s(0)$ and so on. The meaning of this assumption is that finding innovations requires both researchers and waiting time. We can find an innovation either by hiring a larger team now, or by waiting till more knowledge spills over, and then use a smaller team. Furthermore, the production function of innovations by researchers and by waiting time displays constant returns to scale.[6] The main role of this assumption is to enable sustained economic growth over time.

We next describe individuals. The population consists of non-overlapping generations, where each individual lives one period. There is no population growth and the size of each generation is a continuum of size N. As we see below, N is the variable that determines the scale of the economy. Individuals are assumed to be risk neutral.[7] The utility they derive from consumption c is described by:

(4) $$u = c .$$

There is free entry to work either in production or in research. A worker, who produces the final good in period t, sells it in the market to earn income. If she uses a new innovation in production, she has to pay patent fees to the patent holder, who has innovated it in the same period t.[8] Innovators, who belong to a successful team, receive income from patent fees. Unsuccessful innovators have zero income. We assume that individuals decide whether to search for an innovation or work in production before they know whether they succeed or not in being the first to find the innovation.

[6] Doubling the amount of researchers by spreading them equally in two consecutive periods of time doubles the amount of innovations.
[7] This assumption is changed in Section 7, where the effect of risk aversion is analyzed.
[8] A version of the model where innovations are used one period later yields similar results.

3. Equilibrium

3.1 The Markets for Innovations

Consider an innovation *j* found in period *t*, which is sold to production workers in the same period *t*. The inventing team that has got patent rights has monopoly over the innovation. Denote the patent fee paid in period *t* for innovation *j* by $z_{j,t}$. The production workers are willing to purchase the innovation as long as their net income from using it is greater or equal to their income without using the innovation:

(5) $$ba_{t-1} - z_{j,t} \geq 0 .$$

Hence, the demand for the innovation is described by the following step function:

(6) $$Q_{j,t}(z_{j,t}) = \begin{cases} L_t & \text{if } \; z_{j,t} \leq ba_{t-1} \\ 0 & \text{if not.} \end{cases}$$

The monopoly patent holder maximizes profits by setting the fee at the maximum price and all workers purchase the innovation and use it. The patent fee is therefore equal to

(7) $$z_{j,t} = ba_{t-1} .$$

Hence, patent fees are equal for all innovations in period *t*.

3.2. Income and Employment

Due to free entry the expected income of workers must be equal for production workers and for innovators as well, before knowing whether they have succeeded or not in their patent race. Denote the income of production workers in period *t* by w_t. Since their gains from new technologies are equal to what they pay as patent fees we get:

(8) $$w_t = a_{t-1} .$$

The income of a successful innovator, whose team wins the patent race for innovation *j* in period *t*, is equal to:

(9) $$\frac{z_{j,t}L_t}{s(j)} = \frac{ba_{t-1}L_t}{s(j)}.$$

The income of an unsuccessful innovator, whose team does not find the innovation first, is equal to 0. The size of employment in the production sector L_t is related to the size of the R&D sector, R_t, through the equilibrium condition in the labor market:

(10) $$L_t = N - R_t.$$

3.3. The Size and Number of Patent Races

We next turn to determine the amount of innovations in period *t*, I_t, and the size of each patent race $n_{j,t}$. The expected income of an innovator whose team searches for innovation *j* an amount of time *q*, with more teams searching, is:

(11) $$\frac{qz_{j,t}L_t}{s(j)n_{j,t}} = bqa_{t-1}\frac{N-R_t}{n_{j,t}s(j)}.$$

Hence, a research team has an incentive to raise *q* as much as possible at the *j*, which has the lowest available $n_{j,t}s(j)$. Hence, $q = 1$, unless the patent race is satiated. Then innovators compare (11) to the alternative income from production, which is qa_{t-1}. As long as expected income of innovators is higher, more teams join and $n_{j,t}$ increases. Equilibrium is reached when expected income in the two sectors, production and innovation, is equal. Hence, the number of teams working on innovation *j* in period *t* is:

(12) $$n_{j,t} = \frac{b(N-R_t)}{s(j)}.$$

The size of the patent race for innovation j falls as the required size of the team $s(j)$ increases. It therefore positively depends on the returns from successful innovation. The more lucrative innovations attract more research teams.

We next turn to determine how many innovations are found in each period. As innovations become more difficult, the size of the patent rate declines, until it reaches 1. There cannot be a smaller race, since then $b(N - R_t) < s(j)$, and as a result the expected income of innovators in such a part time team, derived from equations (3), (9), and (10), satisfies: $bqa_{t-1}(N - R_t) / s(j) < qa_{t-1}$. Hence, no one joins this race. As a result, research stops at exactly where the patent race consists of one team only, and that determines the amount of innovations in period t:

(13) $$s(I_t) = b(N - R_t)\ .$$

The amount of innovations found in period t depends positively on the size of the production sector $N - R_t$. This is the scale effect.

3.4. Equilibrium R&D

As shown above, the size of each patent race depends on the overall size of the R&D sector R_t. At the same time the size of the R&D sector itself depends on the individual patent races, since:

(14) $$R_t = \int_0^{I_t} n_{j,t} s(j) dj\ .$$

Using (12) and (13) we can rewrite the size of the patent race as a function of the amount of innovations I_t only:

(15) $$n_{j,t} = \frac{s(I_t)}{s(j)} .$$

Substitute (15) in (14) and get:

(16) $$R_t = I_t s(I_t) .$$

This function, which describes how the size of the R&D sector depends on the amount of innovations, is described by the curve *RD* in Figure 2.

[Insert Figure 2 here]

The second curve in Figure 2, *IN*, shows how the amount of innovations depends on the size of the R&D sector, as described by equation (13). Note, that as the size of the research sector increases the production sector decreases and with it the incentive for innovation. When the R&D sector reaches $\overline{R} = N - s(0)/b$, the production sector becomes so small that no innovation is profitable and $I = 0$. The intersection of the two curves, *RD* and *IN*, determines a unique equilibrium, which determines the rate of innovation and the size of the R&D sector. The equilibrium is a steady state, and from here on we delete time subscripts. The equilibrium amount of innovation can be described by the following equilibrium condition, which combines (13) and (16):

(17) $$s(I)(1 + bI) = bN .$$

The amount of innovations determines the equilibrium rate of growth, g:

(18) $$g = \frac{a_t - a_{t-1}}{a_{t-1}} = bI_t = bI .$$

We can now relate the size of the R&D sector to the rate of growth of the economy. From (13), (16) and (18) we get:

$$R = Ib(N - R) = g(N - R) .$$

Hence:

(19) $$\frac{R}{N} = \frac{g}{1+g}.$$

Note that the rate of growth g changes only with the relative size of the R&D sector and not with its absolute size. This result fits the empirical findings of Jones (1995a).

In order to better understand the equilibrium in this model and how it is reached, it is useful to focus on the equilibrating mechanism in comparison with other endogenous growth models. Consider a situation where the expected income of innovators, equation (11), is high and there is high incentive to become innovators. In the original endogenous growth models the equilibrium is reached by reducing the production sector, as more people become researchers, until the scale of production becomes small enough. In Jones (1995b) and Segerstrom (1998) equilibrium is reached by reduction of expected returns through searching for more costly innovations, namely increasing the denominator in (11). This paper presents an additional mechanism, which further increases the denominator in (11), namely increasing the size of the patent race, thus reducing the probability of success.

3.5. Patent Races

As shown above the equilibrium determines not only the size of the research sector and the amount of new innovations per period, but also the maximum size of patent races n, which is equal to:

(20) $$n = \frac{b(N-R)}{s(0)} = \frac{s(I)}{s(0)}.$$

Hence, patent races emerge if the difficulty of innovation is increasing. Intuitively, as some innovations create higher income than others, they attract more research teams, so that patent races emerge. Note, that patent races can become large only if difficulty of innovation increases enough, so that $s(I)$ is much larger than $s(0)$.

From here on we assume that the difficulty of innovation is not only increasing but it is unbounded, so that

(21) $$s(j) \xrightarrow[j\to\infty]{} \infty \, .$$

This assumption can be justified as follows. The size of team s is the cost of finding an innovation, so $1/s$ is the marginal productivity of innovators. Hence, (21) means that marginal productivity of innovators is not only diminishing, but it satisfies the Inada condition as well. If s is unbounded, patent races become larger and larger with the rate of innovation and their size is unbounded as well.

4. The Effect of Scale

In this section we study the effect of changes in scale, namely in size of population N, on the equilibrium. We use this analysis in order to contrast the model both with the original R&D based endogenous growth models and with the data. An analysis of the effect of scale is best done by use of Figure 2. As N increases, the *IN* curve shifts parallel to the right. That increases both the size of the R&D sector R and the amount of innovations I. The size of patent races also increases with scale, as $s(I)$ increases.

Note, that innovation and growth begin only when the economy passes some threshold scale or size. When the economy is small and $N < N_0 = s(0)/b$, the *IN* curve in Figure 2 is zero everywhere and there are no innovations and no economic growth. The reason is that the scale is so small that the incentives for innovation are too low. Only when the size of population exceeds the threshold N_0, innovations and growth begin.

It is important to note that the positive effect of scale on innovations is diminishing with scale. One reason is that additional innovations require more researchers, since *s* is increasing. The other reason is that as scale increases, some of the additional innovators do not search for new innovations but join existing patent races. They increase these races and do not add to innovation and growth. To see it formally use equation (17), which describes the relationship between the scale of the economy and the amount of innovation, to show that the effect of scale satisfies:

(22) $$\frac{\partial I}{\partial N} = \frac{1}{s'(I)(I + b^{-1}) + s(I)} < \frac{1}{s(I)} \xrightarrow[N \to \infty]{} 0 .$$

Hence, the effect of scale on innovations is positive but diminishing, and even goes to zero when scale becomes very large. The effect of scale on the rate of economic growth is similar, since $g = bI$. Figure 3 shows how the equilibrium amount of innovations I depends on the scale N.

[Insert Figure 3 here]

A comparison of these results with the original endogenous growth models reveals many similarities. First, an economy with fixed population can have sustained economic growth through innovations. Another similar result is that scale is critical to the take-off of economic growth, as there is a threshold size necessary for innovation and growth. In both models scale increases the size of the R&D sector. But here the models depart significantly. While in the original endogenous growth models all new researchers contribute to finding innovation, in this model only some find new innovations, while others join existing patent races. Hence, the scale effect on economic growth, which is so strong in the original endogenous growth models, is diminishing in this model and even approaches zero.

The reduction of the scale effect is important, since it is found to be in strong contrast with the empirical evidence. While scale of production increased significantly in the world in the recent two centuries, TFP growth has been fairly stable, at a rate of around 1% annually.[9] We can therefore further specify the function *s* in order to better fit the main stylized facts of global growth, namely increasing scale and bounded growth rates. From here on we assume that the function *s* is not only unbounded, but it rises to infinity at some finite amount of innovations. Formally, we assume that there exists a level *I** such that:

(23) $$s(j) \xrightarrow[j \to I^*]{} \infty .$$

This additional restriction on *s* yields the following results. First, since the curve *RD* in Figure 2 is now bounded by *I**, the equilibrium amount of innovations is bounded by *I** as well. Hence, the rate of growth is bounded too: $g \le bI^*$. This specification therefore fits the empirical findings of Jones (1995a) and others. Furthermore, as the marginal effect of scale on growth diminishes, its effect on the size of patent races increases.

5. Patent Races and Pareto Inefficiency

If many innovation teams search for the same innovation, while one team alone can find it, there is inefficiency due to misallocation of resources. To see this formally, consider a central planner, who reallocates individuals between production and R&D. This planner can assign only one innovation team for each innovation, and assign all others to work in production. This reduces the size of the R&D sector from $\int_0^I n_j s(j)dj$ to $\int_0^I s(j)dj$, as $n_j \ge 1$ for all *j*, and increases the size of the production sector. The rate of growth

[9] See Maddison (1995), Table 2.6.

remains the same, while the level of output in each period increases. That means the equilibrium is not Pareto-Optimal.

This inefficiency is due to duplication of innovation activity in patent races. Such duplication inefficiency has been already observed, mostly in microeconomic studies of innovation, such as Dasgupta and Stiglitz (1980 a, b). Most endogenous growth models do not have duplication, since R&D is not direct. Stokey (1995) and Jones and Williams (1998, 2000) acknowledge the possibility of duplication and its effect on endogenous growth, which Jones and Williams call the "stepping on toes effect," but do not model it explicitly. This paper fills this gap by modeling endogenous patent races.[10]

The inefficiency of equilibrium raises the question whether there is too much or too little R&D, which has been recently discussed in Stokey (1995), Jones and Williams (1998, 2000), and Li (2001). This model gives a different answer to this question. There is both too much and too little R&D. Due to duplication there is too much R&D on innovations, which attract a patent race of more than 1 team. We next show that there is not enough R&D on the innovations beyond I, which are delayed to next period. To show it, consider maximization of the discounted sum of present and future incomes, since individuals are risk-neutral. The subjective discount rate is an intergenerational rate ρ. Increasing innovation only in period 0 increases present and future discounted output by:

$$aL\frac{1+\rho}{\rho-g}b.$$

Present production falls by $as(I)$. Since according to equation (13), the equilibrium rate of innovation satisfies $s(I)=bL$, the benefit from increasing innovation is greater than the

[10] Segerstrom, Anant and Dinopoulos (1990) have patent races without duplication. Etro (2002) has patent races and duplication, but focuses on different issues than this paper.

cost. Hence, equilibrium innovation is below the social optimum. Note, that this result depends on some specific assumptions of the model, such as finite patent time, etc. But the result is clear: R&D should not be supported for innovations that attract large patent races, but only for innovations that are marginal, at *I* and beyond.

6. Equilibrium with Multiple Research Strategies

As shown in Section 5, the equilibrium in the benchmark model leads to Pareto-inefficiency, since too many innovators are trying to find the same innovation. Note, that this result seems to depend on an implicit assumption, that all innovation teams search for the innovation in the same way, except that only one of them finds it first. Hence, the inefficiency might disappear if innovators could search for the same innovation in many different ways. Then they can diversify the search for the innovation and increase the probability of finding it. In this section we examine this possibility of multiple research strategies for each innovation. Interestingly, having multiple research strategies does not remove the inefficiency and patent races are still overcrowded. The research strategies, which are mostly crowded, happen to be those with the highest success probabilities.

For the sake of simplicity consider the case, where the size of research teams is fixed and equal to 1, namely one innovator in each team, for innovations $0 \le j \le I^*$, while it is infinite for innovations beyond I^*. Assume that for each innovation $0 \le j \le I^*$ there are infinitely many different ways to search for it. We call them 'research strategies' and number them by m, $1 \le m < \infty$. The probability of finding the innovation by using strategy m full time is p_m. If an innovator works only q time, $q \le 1$, the probability of finding the innovation is qp_m. The strategies are ordered by decreasing probability of success: $p_1 > p_2 > p_3 > \dots$. The strategies are independent of one another, so that the probability of finding the innovation is $\sum_{m=1}^{M} p_m$ if strategies 1, …, M are followed. If all strategies are followed the probability of success is 1. In other words, one and only one of the research strategies can enable innovators to find the innovation.

The probabilities of the research strategies can in principle vary over innovations and over time, but for tractability, assume that they are the same for all innovations and for all times. Hence, we must further specify these probabilities to be: $p_m = p(1-p)^{m-1}$ for all *m*, where $0 < p < 1$. Under this specification, if an innovation is not found in period *t*, after using methods 1,…, *M*, innovators can use the remaining untried research strategies from next period on, and by renumbering the strategies, *M*+1 becoming 1 etc., the conditional probabilities of success become the original probabilities. Hence, the probabilities of the various research strategies are the same, whether research on the innovation has just begun, or whether it has been going on for some time.

We assume, as in the benchmark model, that a research strategy can be used by a number of teams. Conditional on the success of this strategy, if it leads to finding the innovation, each of the teams who follow it has equal probability of reaching the innovation first, if working full time. Hence, if the number of teams who try to find innovation *j* using research strategy *m* in time *t*, is $n_{m,j,t}$, the success probability of each is:

(24) $$P_{m,j,t} = \begin{cases} \dfrac{p_m}{n_{m,j,t}} & \text{if the team works full time,} \\ \dfrac{qp_m}{n_{m,j,t}} & \text{if the team works } q \text{ time and } n_{m,j,t} \geq 1, \\ qp_m & \text{if the team works } q \text{ time and it is the only team.} \end{cases}$$

Note that most of the analysis of the benchmark case carries through to this case of multiple research strategies. The market for each innovation looks the same, patent fees are the same and so are the wage rate and the employment levels. The expected returns for an innovator, who uses research strategy *m* to find innovation *j* and works only *q* time, are:

(25) $$\frac{qp_m ba_{t-1}(N-R_t)}{n_{m,j,t}}.$$

First, innovators look for the strategy with the highest $p_m / n_{m,j,t}$ and then use it for the maximum amount of time. They set $q=1$, if possible. If the patent race of the *m* strategy is too large they work part time on *m* and part time on *m*+1. This equates the expected returns from all strategies. Since expected income from innovation and from production is equal we get that the size of the patent race for innovation *j* of those who use strategy *m* should be:

(26) $$n_{m,j,t} = p_m b(N-R_t).$$

Hence, patent races emerge in this case as well. The strategies with the highest success probability have the largest races. As *m* increases and the success probability diminishes, so does the size of the patent race. It can be shown that the research strategies that are followed are $m=1,...,M$, where *M* is the highest number that satisfies: [11]

(27) $$p_M \geq \frac{1}{b(N-R)}.$$

This condition determines the strategies that are followed and also the probability *P* of finding the innovation in each period of time: $P = p_1 + p_2 + ... + p_M$.

[11] We suppress the time index from here on, as it is clear that the equilibrium is a steady state.

In order to close the equilibrium we calculate the size of the R&D sector:

(28) $$R = I^* \sum_{m=1}^{M} n_m = I^* b(N-R) \sum_{m=1}^{M} p_m = I^* b(N-R)P.$$

A simple calculation yields:

(29) $$\frac{R}{N} = \frac{I^* bP}{1 + I^* bP}.$$

Hence, the share of the R&D sector in population increases with the probability of success *P*. Note that *I*bP* is the rate of economic growth *g*. Equation (29), which reflects the amount of researchers needed to achieve this success level in equilibrium, is presented by the curve *RN* in Figure 4.

The probability of success depends on the size of the R&D sector also through the scale effect, since:

(30) $$P = \sum_{m=1}^{M} p_m = \sum \left\{ p_m \middle| p_m \geq \frac{1}{b(N-R)} \right\} = \sum \left\{ p_m \middle| \left(1 - \frac{R}{N}\right) p_m \geq \frac{1}{bN} \right\}.$$

This step function, which describes how the probability of success depends on scale, is presented by the curve *PR* in Figure 4. The intersection of the two curves determines the equilibrium in the economy.

[Insert Figure 4 here]

The equilibrium determines the number of research methods adopted each period *M*, the probability of finding innovations *P*, the expected rate of growth of productivity $g = I^* bP$, and the relative size of the R&D sector $g/(1+g)$. The effect of scale on innovation and growth is similar to the benchmark model. As *N* increases, innovations become more profitable, the *PR* curve shifts to the right, more research strategies are followed, and more innovations are found. The positive effect of scale on growth is diminishing though, as patent races increase as well.

The main result of this version of the model is, that despite the many research strategies for each innovation, we still have duplication of innovative activity through patent races. Actually, there are too many innovators using the more promising strategies, with low m, while the economy can benefit from putting innovators to work on the marginal strategies, like $M+1$, as that increases the chance of finding the innovation. Hence, the equilibrium is not Pareto-efficient. A Pareto-improving policy should aim at reducing the number of innovators working on the more promising research strategies 1, …, M, while promoting the more risky research beyond M. The problem is that such a policy is very hard to implement, due to moral hazard and other practical difficulties.

7. Risk Aversion and Concentration of R&D

Our paper shows that if the economy is large enough, there will be patent races and many innovators will participate in these races. But in reality we observe many cases in which innovations are searched by a small number of large R&D firms and in many cases R&D is even concentrated within a single monopoly. Interestingly, our model can account for this phenomenon as well, by attributing it to risk aversion. To see this let us replace our original assumption of risk neutrality by risk aversion. Note that although the gains from innovation are very high, the probability of success is low, if the patent race is large. This creates a strong incentive for innovators to form a coalition, where they will share the gains from innovation if one of them finds it first. This way they have the same expected income, but less risk. Such arrangements are hard to create cooperatively, due to problems of free riding and contract enforcement. These problems can be solved by large research firms, which hire many research teams to search for an innovation.

Interestingly, such firms do not eliminate the inefficiency created by patent races, despite their incentive to reduce the number of innovators, in order to increase the returns from successful innovation per innovator. The reason is that such firms have an opposite incentive to increase the number of innovators in order to increase the chance of getting first, and also to deter potential entrants. Hence, there are still too many researchers. Furthermore, it is shown below that under such cooperation there is more duplication than under competition and less innovation and growth.

To formalize the analysis, we introduce a small change to the benchmark model. Assume that consumers are risk averse and that utility is described by:

(31) $$u = \log c \, .$$

Let us also assume that innovators do not work full time in innovation, but work x of their time as production workers, $x < 1$.[12] Hence, even in case of failure in innovation they have some positive income: $xw_t = xa_{t-1}$. We further assume for simplicity that each innovator can work at one innovation at most, namely that $s(0) > 1$.[13] Let us denote the size of fees per innovator in a winning team by f:

(32) $$f = \frac{b(N-R)}{s(j)} \, .$$

For the sake of simplicity assume, as in the benchmark model, that there is only one research strategy for each innovation.

---

[12] Otherwise no individual enters a patent race with risk of zero income, and the model becomes trivial.
[13] The ability to develop more than one innovation already supplies some insurance and reduces the need to create mechanisms for risk sharing.

We first solve the model under perfect competition, when research monopolies are not allowed, or when the cost of forming them is very high. A team enters the patent race if its expected utility is greater or equal to that of working in production. Hence, if the number of the other teams is *n*, an additional team enters until:

$$\frac{1}{n+1}\log(f+x)+\left(1-\frac{1}{n+1}\right)\log x=0\,. \tag{33}$$

This condition describes how the sizes of patent races depend on the fee per innovator *f*, which depends on the size of the research sector *R*. It determines the equilibrium together with a condition that relates the size of the R&D sector *R* to the sizes of the patent races.

We next turn to the case of concentrated research within firms, which enable an increase of expected utility through risk sharing. For the sake of simplicity we overlook the problems of contract enforcement, and view such firms as groups of researchers, who divide equally the returns from a successful innovation. We further assume that firms participate in a Cournot competition. Let the number of teams in such a firm be *k*. This number is chosen to maximize the expected utility of each innovator, taking the number of teams in the other firms, *n*, as given. Hence, *k* is determined by maximization of:

$$\left(\frac{k}{n+k}\right)\log\left(\frac{f}{k}+x\right)+\left(1-\frac{k}{n+k}\right)\log x=\log x+\frac{k}{n+k}\log\left(\frac{f}{kx}+1\right). \tag{34}$$

Note, that this maximization presents the opposing interests of the firm. On the one hand it has an incentive to reduce *k* in order to increase the return to innovator in case of success, *f*/*k*. On the other hand it has incentive to increase *k* in order to increase the probability of success *k*/(*n*+*k*).

The overall amount of research activity for this innovation is determined by the entry condition. Firms enter the market until the expected income of innovators equals that in the production sector, namely until (34) is equal to zero:

(35) $$\log x + \frac{k}{n+k}\log\left(\frac{f}{kx}+1\right) = 0 .$$

The maximization of expected utility and the entry condition together determine the equilibrium *k* and *n*. The number of teams in each firm *k*, and the overall number of teams working on the innovation *n* + *k*, determine the number of research firms (*n* + *k*)/*k*.

We can now outline the first result of this section. Since (34) is larger than (33), due to optimality, the number of teams in this case, *n* + *k*, is larger. In other words, allowing firms to internalize the patent race and to share risk increases the number of teams searching for the same innovation. The intuition for this result is straightforward: the ability to share risk induces more people to become innovators and to join the patent race. Hence, allowing firms to concentrate research for innovation leads to more R&D, and to more duplication.

A second result is obtained when *f* declines, as *s* increases and innovations become more difficult. It can be shown that as *f* declines, the number of firms decreases, and the total number of teams falls as well. Thus, as the size of innovation teams increases, the patent race becomes more concentrated. Hence, for the most difficult innovations we have a research monopoly, which completely internalizes the patent race. Such a monopoly completely eliminates the risk of participation in the patent race and hence innovators enter as long as expected income from innovation exceeds that from production. Hence, the number of teams in such a monopoly *k* is determined by:

(36) $$\frac{f}{k} = 1 - x \,.$$

As f further declines innovation stops when $k = 1$. In this case the research monopoly consists of one big team only, when $f = 1 - x$. This marginal innovation is the same as in the case of competition. But under cooperation $N - R$ is lower. The reason is that the size of the R&D sector is larger, due to larger number of teams for lower j. Hence, in order for f to be equal to $1 - x$ we need to have lower $s(I)$ and hence lower I. Thus, allowing concentration increases the R&D sector, but reduces the rate of innovation and the rate of growth in the economy.

8. Summary and Conclusions

This paper analyzes the role of patent races in an endogenous growth model. It introduces patent races by adding two assumptions to the model. One is that research is directed, namely innovation-specific. The other assumption is that innovations require not only the labor of innovators, but also waiting time to ripen up. As a result of these two assumptions, patent races emerge endogenously in the model. The paper then shows that patent races significantly reduce the positive effect of scale on growth. A larger economy increases the incentive to innovate, but most new innovators crowd existing patent races, while only few search for additional innovations.

But the paper's main contribution lies beyond the effect of scale on growth, in its analysis of patent races. It examines when these races emerge, when they are run between and when within innovating firms, which research strategies are used more than others, etc. The paper also contributes to the welfare analysis of R&D. It directs our focus on where there is too much R&D and where there is too little. The paper shows that researchers tend to crowd the more promising research strategies, which have higher probabilities of success, while there is too little R&D at the low probabilities of success.

The paper stops short of offering policy recommendations, as the practical problems of formulating an optimal policy are huge. An ideal policy should be to support only those innovators who travel the less frequented ways, namely those who try strategies with lower probabilities of success. Hence, this model suggests that research incentives should be given only to those who deviate from the crowd and who are doing less standard and more risky research. It is hard to find practical policies that identify such researchers, but it is definitely worth searching for.

References

Aghion, Philippe, and Howitt, Peter. "A Model of Growth Through Creative Destruction," *Econometrica*, Vol. 60 (2), March 1992, p. 323-351.

Aghion, Philippe, and Howitt, Peter. *Endogenous Growth Theory*. Cambridge, MA: MIT Press, 1998.

Barro, Robert J., and Sala-i-Martin, Xavier. *Economic Growth*. New York: McGraw Hill, 1995.

Dasgupta, Partha and Stiglitz, Joseph. "Uncertainty, Industrial Structure, and the Speed of R&D," *Bell Journal of Economics*, Vol. 11 (1), Spring 1980 (a), p. 1-28.

Dasgupta, Partha and Stiglitz, Joseph. "Industrial Structure and the Nature of Innovative Activity," *Economic Journal*, Vol. 90 (358), June 1980 (b), p. 266-293.

Etro, Federico. "Patent Races and Schumpeterian Growth," mimeo, February 2002.

Grossman, Gene M. and Helpman, Elhanan. *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press, 1991.

Howitt, Peter. "Steady Endogenous Growth with Population and R & D Inputs Growing," *Journal of Political Economy*, Vol. 107 (4), August 1999, p. 715-730.

Jones, Charles I. "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, Vol. 110 (2), May 1995 (a), p. 495-525.

Jones, Charles I. "R & D-Based Models of Economic Growth," *Journal of Political Economy*, Vol. 103, August 1995 (b), p. 759-784.

Jones, Charles I. And Williams, John C. "Measuring the Social Return to R&D," *Quarterly Journal of Economics*, Vol. 113, November 1998, p. 1119-1135.

Jones, Charles I. And Williams, John C. "Too Much of a Good Thing? The Economics of Investment in R&D," *Journal of Economic Growth*, Vol. 5, March 2000, p. 65-85.

Kortum, Samuel. "Research, Patenting, and Technological Change," *Econometrica*, Vol. 65 (6), November 1997, p. 1389-1419.

Li, Chol-Won. "On the Policy Implications of Endogenous Technological Progress," *The Economic Journal*, Vol. 111, May 2001, p. C164-C179.

Lucas, R.E. Jr. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, Vol. 22, July 1988, p. 3-42.

Maddison, Angus. *Monitoring the World Economy 1820-1992*. Paris, France: Organization for Economic Co-operation and Development, 1995.

Romer, Paul M. “Increasing Returns and Long-Run Growth,” *Journal of Political Economy*, Vol. 94, 1986, p. 1002-1037.

Romer, Paul M. “Endogenous Technical Change,” *Journal of Political Economy*, Vol. 98 (5, pt.2), October 1990, p. S71-S102.

Segerstrom, Paul S. “Endogenous Growth without Scale Effects,” *The American Economic Review*, Vol. 88 (5), December 1998, p. 1290-1310.

Segestrom, Paul S., Anant T.C.A. and Dinopoulos, Elias. “A Schumpeterian Model of the Product Life Cycle,” *The American Economic Review*, Vol. 80 (5), December 1990, p. 1077-1091.

Stokey, Nancy L. “R&D and Economic Growth,” *Review of Economic Studies*, Vol. 62 (3), July 1995, p. 469-489.

Young, Alwyn. “Growth without Scale Effects,” *Journal of Political Economy*, Vol. 106 (2), February 1998, p. 41-63.

Figures

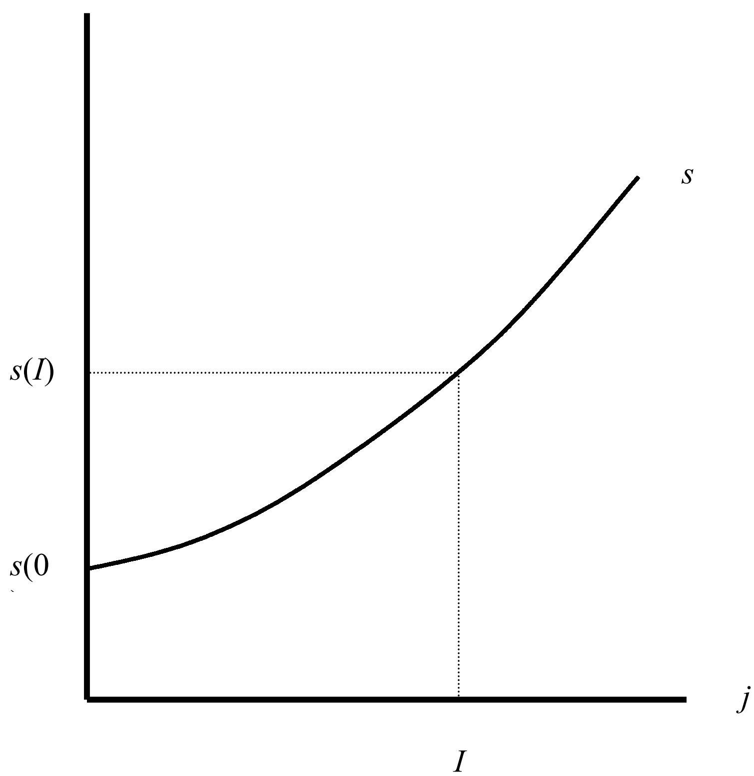

Figure 1

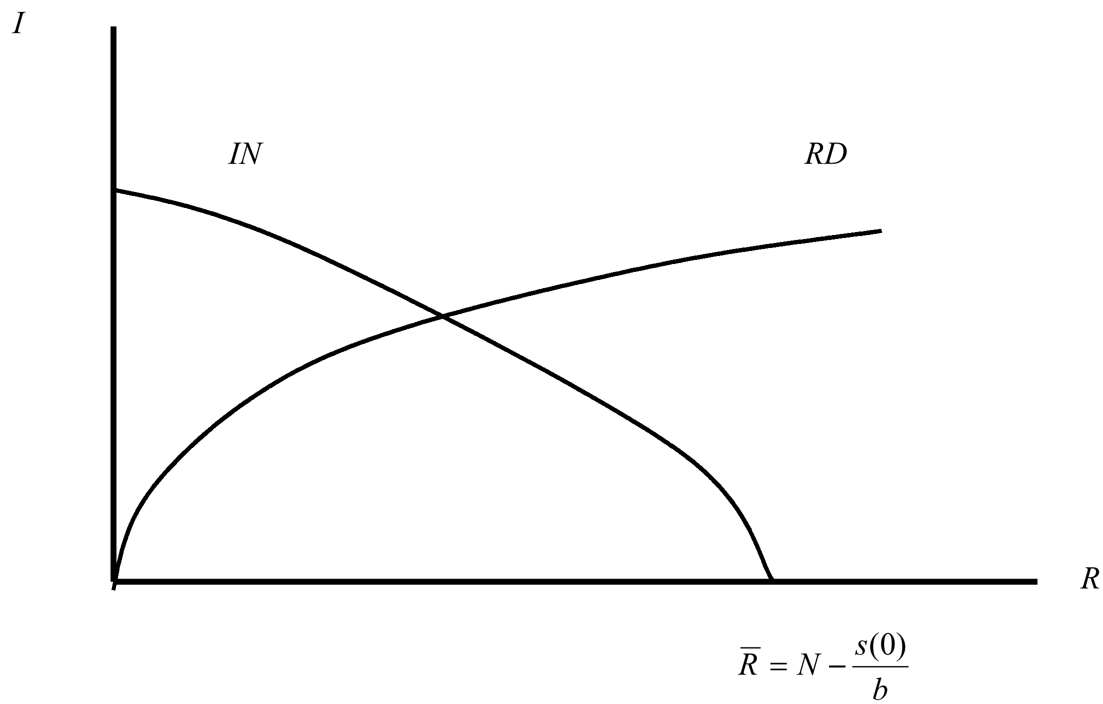

Figure 2

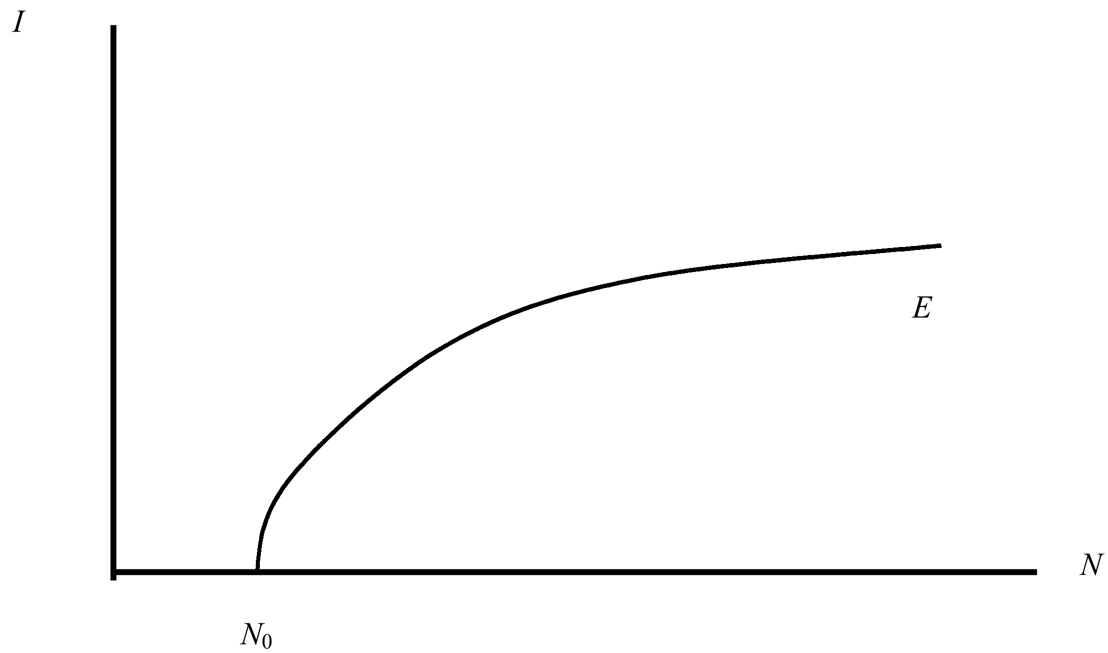

Figure 3

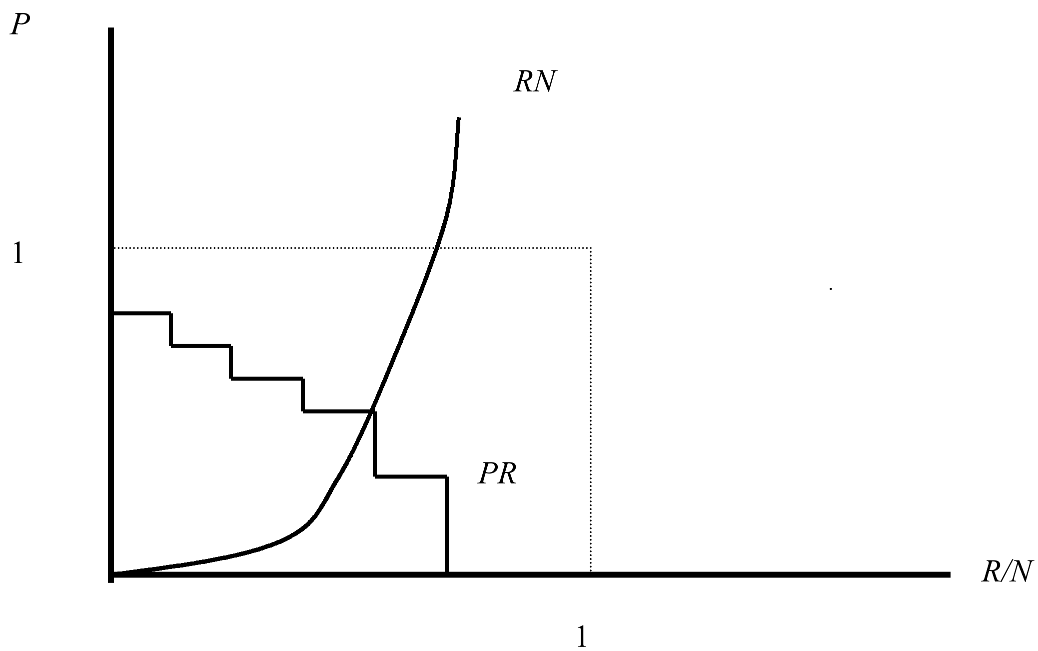

Figure 4