



**Patent policy, patent pools,
and the accumulation of
claims in sequential
innovation**

**Gastón Llanes
Stefano Trento**

Working Paper

10-005

Copyright © 2009 by Gastón Llanes and Stefano Trento

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Patent policy, patent pools, and the accumulation of claims in sequential innovation[☆]

Gastón Llanes

Harvard Business School, Rock Center 222, Soldiers Field Rd., Boston, MA 02138, USA

Email: gllanes@hbs.edu

Stefano Trento

*Universitat Autònoma de Barcelona, Departament d'Economia i Història Econòmica,
Edifici B Office B-174, 08193 Bellaterra, Barcelona, Spain*

Email: stefano.trento@uab.es

July 24, 2009

Abstract

We present a dynamic model where the accumulation of patents generates an increasing number of claims on sequential innovation. We study the equilibrium innovation activity under three regimes: patents, no-patents and patent pools. Patent pools increase the probability of innovation with respect to patents, but we also find that: (1) their outcome can be replicated by a licensing scheme in which innovators sell complete patent rights, and (2) they are dynamically unstable. We find that none of the above regimes can reach the first or second best. Finally, we consider patents of finite duration and determine the optimal patent length.

Key words: Sequential Innovation, Patent Pools, Anticommons

JEL: L13, O31, O34

[☆]We are grateful to Michele Boldrin for his guidance and advice. We thank Antonio Cabrales, Marco Celentani, Belén Jerez and participants of seminars at Harvard Business School, Universidad Carlos III de Madrid, and Universitat Autònoma de Barcelona for useful comments and suggestions. All remaining errors are our responsibility. We gratefully acknowledge financial support from the Ministry of Education of Spain (Llanes, FPU grant AP2003-2204), the Ministry of Science and Technology of Spain (Trento, grant SEJ2006-00538), and the Comunidad Autónoma de Madrid (Trento).

1. Introduction

Patents are intended to enhance private investment in R&D through the monopoly power they grant to the innovator over the commercial exploitation of her invention. Generally, innovations are sequentially linked. For instance, the invention of the radio would have been impossible without the previous discovery of electromagnetic waves. This sequential nature of innovation introduces the issue of how to divide the revenues of the sequence of innovations among the different innovators. Suppose that two innovations may be introduced sequentially (the second innovation cannot be introduced until the first one has been introduced). If a patent is granted to the first innovator, she may obtain a claim over part of the revenues of the second innovator. The policy maker is confronted with an important trade-off: if the patent covering the first innovation is strong, it may imply that the second innovation becomes unprofitable. On the other hand, if the patent is weak, it may provide low incentives to introduce the first innovation.

This problem has been studied in depth by the literature on sequential innovation, pioneered by Scotchmer (1991). Usually, this literature analyzed the optimal division of profits between two sequential innovators. But what happens when each innovation builds on several prior inventions? Going back to the case of radio, it was not only electromagnetic waves that radio was based upon, but also high-frequency alternator, high-frequency transmission arc, magnetic amplifier, the crystal detector, diode and triode valves, directional aerial, etc. In the words of Edwin Armstrong (inventor of FM radio) “it was absolutely impossible to manufacture any kind of workable apparatus without using practically all of the inventions which were then known”.

In this paper, we extend the literature on sequential innovation by analyzing the case in which patents generate cumulative claims on subsequent innovations. If the number of claims is large, innovators may face a patent thicket and may be threatened by the possibility of hold up, namely the possibility that a socially desirable innovation fails to be performed due to the

lack of agreement with previous inventors Shapiro (2001).

Patent thickets are pervasive in hi-tech industries, like software, hardware, biotechnology and electronics. For example, in the 1980s IBM accused Sun Microsystems of infringing some of its 10,000 software patents (<http://www.forbes.com/asap/2002/0624/044.html>); development of golden rice required access to 40 patented products and processes (Graff et al., 2003); and there are 39 patent families “potentially relevant in developing a malaria vaccine from MSP-1” (Commission on Intellectual Property Rights, 2002).

When a patent thicket arises the innovator must pay license fees on many patented previous discoveries, which may lead to low innovation. Heller and Eisenberg (1998) were the first to suggest that a reduction in innovation activity would have stemmed from what Heller (1998) defines the *tragedy of the anticommons*. This phenomenon arises when too many agents have rights of exclusion over a common, scarce resource, and as a consequence the common resource is under-utilized, in clear duality with the tragedy of the commons. In our case, the anticommons could arise if too many patentees have exclusive claims on separate components of the state-of-the-art technology, placing an obstacle to future research. In this paper we build a model to analyze whether anticommons in sequential innovation is a theoretical possibility.

Abstracting from transaction costs, or the possibility that one or more patentees refuse to license their idea therefore blocking innovation, anticommons are similar to complementary monopoly: many monopolistic input providers selling their inputs to a final good producer. The problem of complementary monopoly was first analyzed by Cournot (1838). He modeled a competitive producer of brass who has to use copper and zinc as perfect complement inputs in production. He showed that, when the inputs are sold by two different monopolists, the total cost of producing brass is higher than when the two inputs are sold by the same monopolist. Sonnenschein (1968) showed that complementary monopoly is equivalent to duopoly in quantity

with homogeneous goods, and Bergstrom (1978) generalized this result to a general number n of inputs and any degree of complementarity among them.

Recently Shapiro (2001) and Lerner and Tirole (2004) applied the instruments of complementary monopoly to the analysis of patent pools. Their results reinforced the results on complementary monopoly: patent pools (or equivalently a single monopolist owning all the patented inputs) reduces the cost of innovation when patents are complements, and it increases it when they are substitutes. Boldrin and Levine (2005) and Llanes and Trento (2009) also made use of complementary monopoly to show that, as the number of complementary patents increases, the probability that a future innovation is profitable goes to zero.

All these papers, while making important contributions, present static models. In other words, the first innovation has been already invented, so patents and patent pools only affect the profitability of introducing a second innovation. This introduces an important asymmetry between previous and future innovations. We believe that adding a dynamic dimension is an important step towards a better understanding of the mechanism of anticommons in sequential innovation.

Developing a dynamic model is important for several reasons. First, it will eliminate any bias stemming from the asymmetric treatment of old and new ideas. Second, patent policy will affect not only current but also future innovative activity. Third, it will allow us to analyze the problem of assigning resources to promote innovations with low commercial value (basic research). Fourth, some of the previous findings in the literature will be affected by the introduction of the temporal dimension, and new issues will arise precisely because of this modification.

We present a dynamic model to study the division of profit when each innovation builds on several prior inventions. There is a sequence of innovations $n = 1, 2, \dots$. Innovation n cannot be introduced until innovation $n - 1$ has been introduced. Each innovation has a commercial value (the profit it

generates as a final good), which is random and private information of the innovator, and a deterministic cost of R&D to be developed.

Our model provides a good description of the innovation process in several industries. For example, in the software industry the first programs were written from scratch, and therefore built on little prior knowledge. As more and more programs were developed, they progressively became more dependent on technologies introduced by the first programs. According to Garfinkel et al. (1991), nowadays software programs contain thousands of mathematical algorithms and techniques, which may be patented by the innovators who developed them. Similar examples can be found in other hi-tech industries.

Formally, our model is a multi-stage game in discrete time with uncertain end. Interestingly, the probability of reaching the next period is determined endogenously. Our theoretical contribution is to present a simple dynamic model that can obtain closed form solutions for the sequence of probabilities of innovation. The equilibrium concept we use is Subgame Perfect Equilibrium with Markovian Strategies (Markov Perfect Equilibrium).

We are interested in determining equilibrium dynamics under three scenarios: patents, no patents and patent pools. The novel aspect of our model is that patents generate cumulative claims on the sequence of innovations. Patents affect the innovator in two ways: on one hand, the innovator has to pay licensing fees to all previous inventors, but on the other hand, she will collect licensing revenues from all subsequent innovators, in case they decide to innovate. Therefore, it is not clear what is the net effect of patents on innovation as the sequence of innovations progresses.

We find that with patents, innovation becomes harder and harder with more complex innovations. The probability of innovation goes to 0 as $n \rightarrow \infty$. The probability of innovation is higher than in the static case, but not enough to stop the tragedy of the anticommons from happening.

In the no patents case, on the other hand, the probability of innovation is constant and depends on the degree of appropriability of the commercial

value of the innovation in the final goods sector. The no patents case will provide higher innovation than the patents case unless the innovator can appropriate a very small fraction of the value of the innovation.

When ideas are protected by patents, the formation of a patent pool increases the probability of innovation for all innovations. Interestingly, the probability of innovation with a pool is constant and higher than what it would be in the static case. This result strengthens the findings of Shapiro (2001), Lerner and Tirole (2004), and Llanes and Trento (2009) for static models. The comparison between the pool and the no patents case depends once again on the degree of appropriability of the value of the innovation in the latter case.

We present two additional findings which are new to the literature of patent pools. First, the patent pool outcome can be replicated by a scheme in which each innovator buys all patent rights from the preceding innovator, instead of paying only for the permission to use the idea (innovator 1 sells all the rights over innovation 1 to innovator 2, who sells all the rights over innovations 1 and 2 to innovator 3 and so on). This means that the complete sale of patent rights will generate higher innovation than licensing. However, this scheme may be difficult to implement when the nature of innovations is difficult to describe ex-ante, in which case patent pools would be easier to enforce. Second, we find that pools are dynamically unstable: the temptation to remain outside the pool increases as the sequence of innovations advances. This means that early innovators have more incentives to enter the pool than later innovators.

We find the optimal innovation policy that maximizes the expected welfare of the sequence of innovations. We find that innovation is suboptimal in the three policy regimes. In the no patents regime, there is a dynamic externality: innovators do not take into account the impact of their decision on the technological possibilities of future innovators. In the two other policy regimes, the inefficiency stems from asymmetric information and market

power: patent holders do not know the exact value of the innovation, but they know its probability distribution. This generates a downward sloping expected demand for the use of their ideas, and market power implies a price for old ideas above marginal cost.

It is interesting to analyze the second best innovation policy, implemented through transfers between innovators when the patent office does not know the value of the innovations. In this case, we find that patents are larger than these transfers, and therefore the patent regime cannot even attain the second best. Finally, we extend the basic model to analyze what happens when the sequence of innovations does not stop after an innovation fails, and to analyze the optimal patent policy when patents have finite length.

Our paper is related to Hopenhayn et al. (2006), which also presents a model of cumulative innovation with asymmetric information. However, the focus of that paper is different. In their paper innovations are substitutes: the introduction of a new product automatically implies the disappearance of old versions in the market. Patents block the introduction of subsequent innovations for the duration of the patent. The question they study is how to allocate monopoly power in the final goods market to successive innovations. A trade-off arises because the promise of property rights to the first innovator limits what can be offered to the second innovator. In our paper, innovations are complementary: all prior inventions are necessary to introduce a new idea. We study what is the effect of intellectual property rights on the pricing of old ideas. The problem is that granting too many rights on sequential innovations implies an increase in licensing fees, potentially hindering innovation as a consequence. Therefore, the two papers offer complementary analysis of the process of sequential innovation when the value of innovations is private information.

2. Innovation with patents

There is a sequence of innovations $n = 1, 2, \dots$. Innovation n cannot be introduced until innovation $n - 1$ has been introduced. Formally, the model is a multi-stage game with uncertain end. At each stage, an innovator may introduce an innovation. If the innovation is performed, the game continues and further innovations may be introduced. If the innovation fails to be performed, the game ends and no other innovation can be introduced (we will relax this assumption in Section 10). We will see that the probability that the game continues is determined endogenously.

The innovation process is deterministic. At stage n , the innovator may introduce the new idea by incurring in an R&D cost of ε . The new idea is based on $n - 1$ previous ideas. These previous ideas are protected by patents, which means that the innovator has to pay licensing fees to the $n - 1$ previous innovators (patent holders), in case she wants to introduce the innovation. The cost of innovation is the sum of the cost of R&D and the licensing fees paid to patent holders.

The innovation process reflects the fact that usually earlier innovations do not have a solid background to build upon, while as the market comes to maturity further innovations are more and more indebted to previous ones.

Each idea has a commercial value v_n , which represents the revenues obtained by selling the new product in the final goods market. In order to concentrate on the effects of patents on innovation activity, we will assume that the innovator is a perfect price discriminator in the final goods market, which means that the commercial value of the innovation is equal to the social surplus generated by the new product.

The value of the innovation is private information of the innovator. Patentees only know that v_n is drawn from a uniform distribution between 0 and 1, with cumulative distribution function $F(v_n) = v_n$.

The new idea will be protected by a patent of infinite length, which means that the innovator can request licensing fees from all subsequent innovators

(we will relax this assumption in Section 11). The total revenues of the innovation equal the commercial value of the innovation plus the future licensing revenues.

The timing of the game within each stage is the following: (i) the $n - 1$ patent holders set licensing fees p_n^i , (ii) Nature extracts a value for v_n from distribution $F(v_n)$, (iii) the innovator decides whether to innovate ($I_n = 1$) or not ($I_n = 0$).

At each stage, patent holders only care about maximizing the expected future licensing revenues. Let J_n^i denote the expected future licensing revenues of innovator i at stage n , given that stage n has been reached. Then,

$$\begin{aligned} J_n^i &= Pr_n p_n^i + Pr_n Pr_{n+1} p_{n+1}^i + \dots \\ &= \sum_{m=n}^{\infty} p_m^i \prod_{k=n}^m Pr_k, \end{aligned} \tag{1}$$

where Pr_n is the probability that the n^{th} innovation is introduced, given that all prior innovations have been introduced. Notice that the probabilities Pr_n work as intertemporal discount factors, which arise endogenously from the specification of the model.

J_n^i can also be expressed in a recursive way:

$$J_n^i = Pr_n (p_n^i + J_{n+1}^i), \tag{2}$$

This means that with probability Pr_n the innovation is performed, and the patent holder gets the licensing fee from the innovator plus the continuation value of her expected licensing revenues.

The innovator's payoff is $I_n(v_n + J_{n+1}^i - c_n - \varepsilon)$, where $c_n = \sum_{i=1}^{n-1} p_n^i$ is the sum of licensing fees paid to previous innovators.

We will focus on Markov strategies. A strategy for player i specifies an action conditioned on the state, where actions are prices and the state is simply the number of previous innovations introduced. The equilibrium

concept is Markov Perfect Equilibrium.

Perfectness implies that future prices will be determined in following subgames, as the result of a Nash equilibrium. Thus, players understand that no action taken today can influence future prices and probabilities. Current actions only affect the probability that the following stage is reached, through the effect of current prices on the probability of innovation. We have just proved the following lemma:

Lemma 1. J_m^i for $m > n$ does not depend on any action taken at stage n .

The game is solved recursively. The solution to the innovator's problem is straightforward. Given v_n and c_n , the innovator forecasts J_{n+1}^n , and decides to innovate ($I_n = 1$) if the revenues from the innovation exceed the cost of innovation:

$$I_n = \begin{cases} 1 & \text{if } v_n + J_{n+1}^n \geq c_n + \varepsilon, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

which implies that the probability of innovation is $Pr_n = 1 + J_{n+1}^n - c_n - \varepsilon$.

At stage n , patent holders want to maximize their expected licensing revenues from stage n onwards. They know their decisions do not affect J_{n+1}^i (they can only affect the probability that stage $n + 1$ is reached), and decide a licensing fee p_n^i , taking the decision of the other patent holders as given. The patent holder's problem is:

$$\max_{p_n^i} J_n^i = Pr_n (p_n^i + J_{n+1}^i), \quad (4)$$

which leads to a price equal to $p_n^i = (1 - \varepsilon)/n$.

Imposing symmetry, $p_n^i = p_n$ and $J_n^i = J_n$ for all i . Replacing prices and probability in J_n , we get $J_n = \left(\frac{1-\varepsilon}{n} + J_{n+1}\right)^2$. Rearranging this equation, $J_{n+1} = \sqrt{J_n} - \frac{1-\varepsilon}{n}$, which is a decreasing sequence, converging to 0 as $n \rightarrow \infty$.

The sequence of probabilities of innovation is:

$$Pr_{n+1} = \left(Pr_n - \frac{1 - \varepsilon}{n}\right)^{1/2}, \quad (5)$$

which is also a decreasing sequence converging to 0 as $n \rightarrow \infty$. This means that innovation gets harder and harder with more complex innovations (those that are based on more previous innovations).

The probability of innovation decreases with complexity because patent holders do not take into account cross-price effects: patent holder i set the price of her patent by equating the marginal revenue and the marginal cost of increasing her license fee. The marginal revenue is simply the additional revenue in case the new innovation is performed. The marginal cost is the reduction in expected demand, and depends on the fact that - since all patents are essential for the new innovation - increasing the price of patent i decreases the probability of innovation. As a matter of fact, increasing the price of patent i reduces the expected demand for all other inputs, but this effect does not enter the marginal cost for patent holder i . This generates the anticommons effect that closely resembles the tragedy of the commons: patent holders ignore this cross-price effects and set a price that is higher than the price they would set if they coordinated (see section 4).

In practice the anticommons can be interpreted as a combination of coordination failure and market power. Each patent produces a claim over part of the revenues generated by subsequent innovations. Since each patent is essential, and patent holders do not take into account cross-price effects, they set a license fee that is too high. As the number of claims increases, the coordination problem gets worse, until eventually the new innovator is left with negative expected profits with probability one.

3. Innovation without patents

Suppose that a policy reform completely removes patents. This change has two effects on innovation. First, the revenues of the innovator in the final goods sector will decrease as a result of imitation. Specifically, assume that the innovator can only appropriate a fraction $\phi \in [0, 1]$ of the consumer surplus generated by the innovation. Second, innovators will not pay licensing

fees to previous innovators, nor will they charge for the use of their ideas in subsequent innovations. Therefore, $c_n = 0$ and $J_n = 0$ in the previous model.

At each stage: (i) nature extracts a value of the innovation v_n , and (ii) the innovator decides to innovate or not. The innovator will innovate if $\phi v_n \geq \varepsilon$ and will not innovate otherwise. Thus, the probability of innovation is constant and equal to $1 - \varepsilon/\phi$ if $\phi > \varepsilon$. If $\phi \leq \varepsilon$, then the probability of innovation is zero.

4. Patent pools

In this section we analyze what happens when licensing fees are set cooperatively by a collective institution like a patent pool. At each stage, the pool maximizes the future expected revenues of *current* patent holders. The pool will set a symmetric price for all current patent holders. Once an innovation is performed, the innovator becomes a member of the pool in all subsequent stages. In the first stage there is no pool because no innovation has been introduced (the pool plays from stage 2 onwards).

The probability of innovation is $Pr_n = 1 + J_{n+1} - (n - 1)p_n - \varepsilon$, and the pool's problem is

$$\max_{p_n} J_n = Pr_n (p_n + J_{n+1}). \quad (6)$$

The difference with respect to the non-cooperative case is that the pool recognizes cross-price effects, and therefore is encouraged to set lower prices than in the no-pool case.

A higher J_{n+1} fosters innovation in two ways. First, it increases the future revenues of the innovator. Second, it encourages the pool to set a lower price, because it increases the loss of current patent holders if the sequence of innovations is stopped.

The equilibrium price is $p_n = \frac{1-\varepsilon}{2(n-1)} - \frac{n-2}{2(n-1)}J_{n+1}$, which is equal to the price a pool would set in a static model (see section 8.1) *minus* an additional term arising from the pool's concern of keeping future revenues.

The probability of innovation becomes $Pr_n = \frac{1-\varepsilon}{2} + \frac{n}{2}J_{n+1}$. Introducing price and probability in J_n , we get

$$J_n = \frac{1}{n-1} \left(\frac{1-\varepsilon}{2} + \frac{n}{2}J_{n+1} \right)^2. \quad (7)$$

This is a first order non-linear difference equation. J_n is decreasing in n and converges to 0 as $n \rightarrow \infty$.

The sequence in terms of probabilities is

$$Pr_n = \frac{1-\varepsilon}{2} + \frac{1}{2}Pr_{n+1}^2. \quad (8)$$

which is a constant sequence such that $Pr_n = 1 - \sqrt{\varepsilon}$ for $n \geq 2$. To determine Pr_1 we need J_2 , which is equal to $(1 - \sqrt{\varepsilon})^2$. Then, $Pr_1 = \min\{1, 2(1 - \sqrt{\varepsilon})\}$.

5. Comparison

Figure 1 shows the evolution of the probability of innovation in the three cases studied above: infinitely lived patents, no patents and patent pool. The cost of R&D is $\varepsilon = 0.2$ and we consider $\phi = 1$ (full appropriation) and $\phi = 0.3$ (the innovator appropriates 30% of the social surplus generated by the new product) for the no patents case.

Comparing the patent and no-patent cases, we can see that patents increase the probability of the first innovations but decrease the probability of further innovations. The number of innovations for which patents increase the probability depends on ϕ . For example, when $\phi = 1$, patents only increase the probability of the first innovation. Nevertheless, even when $\phi = 0.3$ the probability increases only for the first two innovations. For patents to increase the probability of several innovations, it is necessary that ϕ is very small and close to ε (i.e. when there is very little appropriability without patents).

When ideas are protected by patents, the formation of a patent pool in-

creases the probability of innovation. Figure 1 shows that the probability of innovation with patent pools is always larger than the patents case. Moreover, with a pool the probability of innovation does not go to zero as $n \rightarrow \infty$. The comparison with the no patents case depends on ε and ϕ . If ϕ is low, a patent pool increases the probability of all innovations. When ϕ is high, the pool increases the probability of the first innovation, and decreases the probability of all posterior innovations.

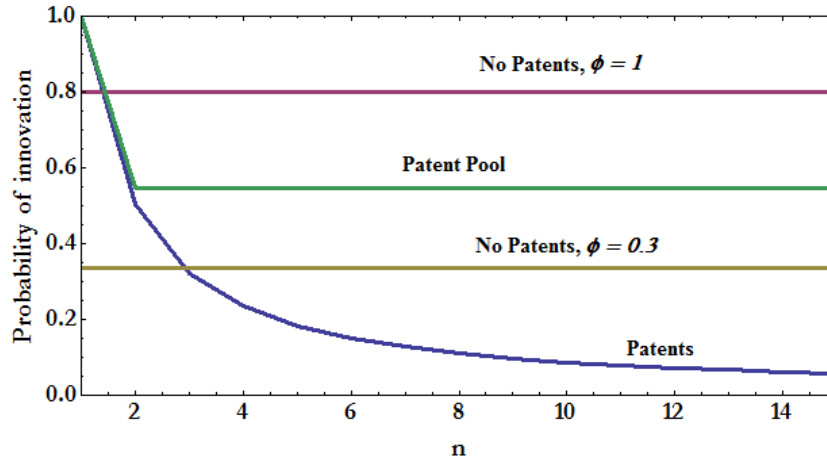


Figure 1: Comparison of equilibria.

6. Complete sale of patent rights

The tragedy of the anticommons stems from fragmented ownership of complementary patents. In this case, the probability of innovation decreases as more innovations are introduced, converging to 0 as $n \rightarrow \infty$. The formation of a patent pool would alleviate this problem by concentrating all pricing decisions on one entity. In this section, we discuss a possible alternative solution, which is to enforce the sale of complete patent rights, instead of allowing the sale of individual access rights through licensing fees. These

patent rights, in turn, can be resold to other innovators. In this case, innovator 1 would sell the complete patent rights over innovation 1 to innovator 2 for a price \wp_1 . Innovator 2 then would sell the patent rights on innovations 1 and 2 to innovator 3 for a price \wp_2 , and so on. We will show that this mechanism eliminates the coordination failure at the basis of the anticommons, and that it replicates the innovation outcome under a patent pool.

The cost of innovation n becomes $\varepsilon + \wp_{n-1}$, its expected revenues $v_n + Pr_{n+1}\wp_n$, and the probability that innovation n is performed $Pr_n = 1 - \varepsilon - \wp_{n-1} + Pr_{n+1}\wp_n$. At stage n , innovator $n-1$ solves the following maximization problem:

$$\max_{\wp_{n-1}} J_n^s = Pr_n \wp_{n-1} \quad (9)$$

where J_n^s represent expected revenues of selling the $n-1$ patent rights to innovator n .

Solving the maximization problem yields a price for patent rights $\wp_{n-1} = \frac{1-\varepsilon}{2} + \frac{1}{2}J_{n+1}^s$, and a sequence of probabilities of innovation $Pr_n = \frac{1-\varepsilon}{2} + \frac{1}{2}Pr_{n+1}^2$ exactly as in the patent pool case. As before, the sequence is $Pr_1 = \min\{1, 2(1 - \sqrt{\varepsilon})\}$, and $Pr_n = 1 - \sqrt{\varepsilon}$ for $n \geq 2$.

An alternative policy arrangement leading to the same result would be the following: restoring the possibility of licensing access rights, but at the same time allowing subsequent competition between the licensee and the original licensor. In this case, if innovator n licenses the use of innovation n to innovator $n+1$, then innovator $n+2$ can license the use of innovation n from innovators n and $n+1$. Under this policy arrangement, innovator n will only get positive revenues from the licensing of her innovation to innovator $n+1$, because at stage $n+1$ she is a monopolist. After that, she will face competition from other innovators, and Bertrand competition will imply a licensing fee equal to zero.

These schemes may be difficult to implement when the nature of innovations is difficult to describe ex-ante. For example, when selling the rights over innovation n to innovator $n+1$, it is difficult to describe what innova-

tion $n + 2$ may be. In this case, complete contracts may be difficult to write, making patent pools easier to enforce.

7. Endogenous patent pool formation

In section 4 we have assumed that all innovators, after innovating, automatically join the patent pool. Let us now endogenize this choice, by analyzing the incentives for innovator n to join the pool. In particular let us compare the expected revenues from joining the pool (J_n from section 4) with the expected revenues from setting the price of her patent non-cooperatively (J_n^o).

We start with the non-cooperative choice. For expositional clarity let us refer to the patent pool members as *insiders* and to the non-cooperative member as the *outsider*. The pool maximizes the expected revenues of the insiders:

$$\max_{p_n^i} J_n^i = Pr_n (p_n^i + J_{n+1}^i) \quad (10)$$

where p_n^i is the cooperative price of any insider's patent, and $Pr_n = 1 + J_{n+1}^i - (n - 2)p_n^i - p_n^o - \varepsilon$, with p_n^o denoting the price of the outsider's patent.

On the other hand, the outsider maximizes:

$$\max_{p_n^o} J_n^o = Pr_n (p_n^o + J_{n+1}^o). \quad (11)$$

From first order conditions we know that $J_n^o = (n - 2)J_n^i$, meaning that if there is one outsider it is convenient to be her. Now let us compare the expected revenues from not joining the pool (J_n^o) to the expected revenues of joining the pool given that everybody else is in the pool (J_n from section 4). In equilibrium, deviating from the pool produces an expected revenue of $J_n^o = (3 - \sqrt{1 + 8\varepsilon})/4$, which is constant and only depends on the R&D cost ε . If, on the other hand, innovator n decides to join the pool together with the $n - 1$ previous innovators, her expected revenue is $J_n = (1 - \varepsilon)^2/(n - 1)$ which is decreasing in n . This is because the patent pool maximizes the

joint profits, thus keeping total cost of innovating constant. This constant amount must be divided among an increasing number of insiders, therefore the expected revenue of an insider is decreasing in n . Figure 2 shows the gains from deviating from the pool, that is the difference between J_n^o and J_n , as a function of n . For $n \geq 4$ it is convenient for the innovator to set her price non-cooperatively.

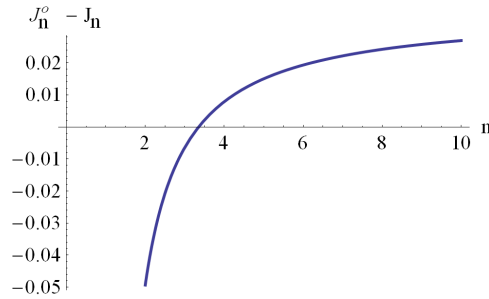


Figure 2: Gains from not joining the patent pool.

Patent pools can improve innovation activity, but are dynamically unstable. Early innovators have more incentives to enter the pool than subsequent innovators. This might explain why governments in some cases have to enforce the creation of patent pools, as the US government did in the radio and aircraft cases for example.

8. Socially optimal innovation

The relevant measure of welfare is the expected social value generated by the sequence of innovations. The social value of an innovation is equal to the increase in consumer surplus minus the cost of the resources spent in R&D. Therefore, at stage n , the social value generated is $v_n - \varepsilon$ if an innovation is performed, and 0 otherwise. Let W be the expected social value. Then,

$$W = \sum_{n=1}^{\infty} E(v_n - \varepsilon / I_n = 1) Pr(I_n = 1) + E(0 / I_n = 0) Pr(I_n = 0)$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} E(v_n - \varepsilon / I_n = 1) \prod_{m=1}^n Pr_m \\
&= \sum_{n=1}^{\infty} \int_{w_n}^1 \frac{v_n - \varepsilon}{1 - w_n} dv_n \prod_{m=1}^n (1 - w_m) \\
&= \sum_{n=1}^{\infty} \left(\frac{1 + w_n - 2\varepsilon}{2} \right) \prod_{m=1}^n (1 - w_m), \tag{12}
\end{aligned}$$

where w_n is the smallest v_n such that the innovation is performed. In the cases studied above, $w_n = \varepsilon/\phi$ when there are no patents and $w_n = c_n + \varepsilon - J_{n+1}$ with patents or patent pools.

Suppose now that the decision of whether to innovate or not is taken by a centralized agency or social planner. The social planner has to determine $\{w_n\}_{n=1}^{\infty}$, namely what is the minimum value of v_n she would require to perform the innovation at stage n . The planner may decide to perform an innovation even when the realization of v_n is less than ε , if the expected gain from future innovations exceeds the current loss in terms of welfare.

Proposition 1 (Socially optimal innovation). *In order to maximize expected social welfare, the innovation should be performed at stage n if and only if $v_n \geq w_n^*$, where*

$$w_n^* = \begin{cases} 0 & \text{if } \varepsilon \leq E(v_n) = 1/2, \\ \sqrt{2\varepsilon - 1} & \text{if } \varepsilon > E(v_n) = 1/2. \end{cases} \tag{13}$$

Proof. Because previous decisions are irrelevant once a stage is reached, the social planner's problem at stage n is exactly the same as the problem at stage $n + 1$, which means that $w_n = w$ for all n . The social planner wants to maximize

$$W = \sum_{n=1}^{\infty} \left(\frac{1 + w - 2\varepsilon}{2} \right) (1 - w)^n \tag{14}$$

$$= \left(\frac{1 + w - 2\varepsilon}{2} \right) \frac{1 - w}{w}$$

The first order condition is $-(1 + w^2 - 2\varepsilon)/(2w^2) \leq 0$, with equality if $w \geq 0$. The value that equates the first order condition, $w^* = \sqrt{2\varepsilon - 1}$, makes sense only when $\varepsilon \geq 1/2$. On the other hand, when $w \rightarrow 0$, the first order condition converges to $\text{sign}(2\varepsilon - 1)\infty$, which means that $w^* = 0$ only if $\varepsilon < 1/2$. ■

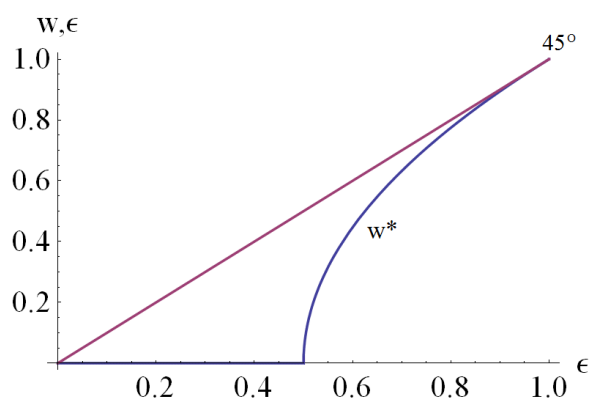


Figure 3: Socially optimal innovation.

Proposition 1 implies that innovation will be suboptimal in the three cases studied above, unless $\varepsilon = 0$. There are three reasons why this is so: dynamic externalities, market power and asymmetric information.

The dynamic externality is best described by analyzing the no patents case. Without patents, the innovator will perform the innovation when $v_n \geq \varepsilon/\phi$. Given that $w_n^* \leq \varepsilon$, the innovator may decide not to perform an innovation when it is socially optimal to do so, even if $\phi = 1$. There is a dynamic externality: the innovator ignores the effect of her decision on the technological possibilities of future innovators. This effect is well known in the literature of sequential innovation (Scotchmer, 1991; Hopenhayn et al., 2006), and is similar to the one found in the literature of moral hazard in

teams (see for example Holmstrom, 1982), where each agent internalizes only his reward from the effort exerted.

With respect to the patents and patent pool cases, the inefficiency arises from a different source: market power and asymmetric information. Because patentees care about the stream of future licensing revenues, they internalize the effect of today’s decision on future innovation. However, asymmetric information implies a downward sloping expected demand for old innovations, and market power implies inefficient pricing of patents, which leads to suboptimal innovation. As the number of holders of rights on innovation increases, the inefficiency due to market power increases.

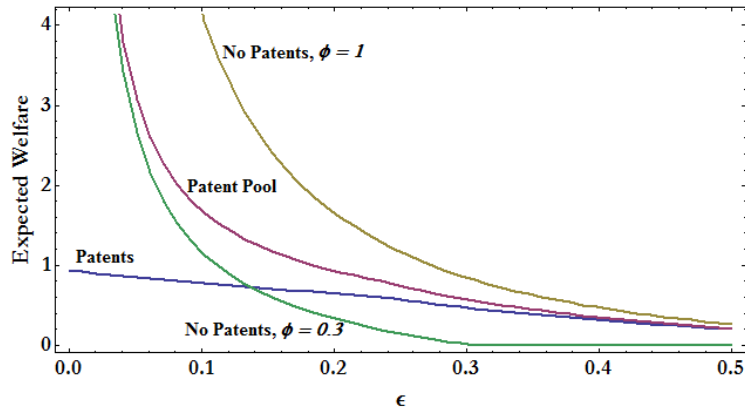


Figure 4: Comparison of expected welfare.

8.1. Static versus dynamic incentives

Previous models of complementary monopoly, sequential innovation and patent pools were static (Shapiro, 2001; Lerner and Tirole, 2004; Boldrin and Levine, 2005; Llanes and Trento, 2009). It is interesting to ask what changes when we add the dynamic dimension.

To see what happens in the static case, assume only one innovation is being considered. The innovation uses $n - 1$ old ideas, which have already

been invented. If the innovation is performed, the innovator obtains a value v from a uniform distribution between 0 and 1, and incurs in a cost ε of R&D. The probability of innovation is $Pr = 1 - \varepsilon - c_n$, with patents or patent pool and $Pr = 1 - \varepsilon/\phi$ without patents.

With patents, the patent holder's problem is to maximize $Pr p^i$, the equilibrium price is $\frac{1-\varepsilon}{n}$ and the probability of innovation is $\frac{1-\varepsilon}{n}$. We have shown that in the dynamic model, the probability is $\frac{1-\varepsilon}{n} + J_{n+1}$, with $J_{n+1} > 0$. This extra term arises because the innovator gets licensing revenues from future innovators. Dynamic incentives imply a higher probability of innovation, but the increase is not enough to prevent the probability from converging to 0 as $n \rightarrow \infty$.

A patent pool would consider cross-price effects, which would lead to a price of $\frac{1-\varepsilon}{2(n-1)}$ and a probability of innovation of $\frac{1-\varepsilon}{2}$. The probability of the corresponding dynamic model is $\frac{1-\varepsilon}{2} + (n-1)J_n$, with $J_{n+1} > 0$. In this case, the extra term arises not only due to the future licensing revenues of the innovator, but also because the pool is concerned with keeping the future licensing revenues of current patent holders.

With respect to the no patents case, the profit-maximizing decision is the same as in the dynamic case. This means that innovators will perform the innovation if $\phi v_n \geq \varepsilon$, which leads to a probability of $Pr = 1 - \varepsilon/\phi$. However, in the dynamic case innovation is suboptimal even when $\phi = 1$, which contrast with the static case, where innovation is socially optimal because there is no intertemporal link between innovations and therefore there is no externality.

9. Dynamic externalities and optimal transfers

In the previous section we have shown that sequential innovation is sub-optimal because of the presence of dynamic externalities and asymmetric information. Without patents, current innovators do not take into account the effect of their decisions on the innovation possibilities of future inno-

vators. The solution to this problem would require intertemporal transfers between innovators. Patents provide a way to transfer resources from future innovators to current innovators, but we have shown that with patents, market power leads to high licensing fees for old innovations, and therefore to low innovation. In this section we ask how close can the government get to the social optimum when it does not have information on the value of innovations (second best analysis).

To do this, we will use a simplified 2-period version of the general model. In the first period, innovator 1 has the option of introducing an innovation with value v_1 and cost ε . If innovator 1 decides to perform the innovation, in the second period, innovator 2 can introduce an innovation with value v_2 and cost ε . Innovator 1 does not know v_2 .

To determine the social optimum, we have to assume the social planner knows v_1 at stage 1, and v_2 at stage 2. It is likely to think that the government would have reduced information on v_n , but assuming the social planner does not know v_n would imply that innovation decisions without patents give higher welfare than the social optimum, which does not make sense. Later we will analyze government policy, and we will assume that the government does not know v_n .

9.1. First best

Let us begin by finding the optimal innovation policy in this 2-stage model. At stage 2 the value and cost of the first innovation are sunk. Therefore, the second innovation should be performed if $v_2 \geq \varepsilon$, and should not be performed otherwise. Consider now the first innovation. The social planner will introduce this innovation if

$$\begin{aligned} v_1 + Pr(v_2 \geq \varepsilon) E(v_2 - \varepsilon | v_2 \geq \varepsilon) &\geq \varepsilon, \\ v_1 + (1 - \varepsilon) \left(\frac{1 - \varepsilon}{2} \right) &\geq \varepsilon. \end{aligned} \tag{15}$$

which leads to a probability of innovation $Pr_1^* = \min\{1, \frac{(1-\varepsilon)(3-\varepsilon)}{2}\}$.

Without patents, the probability of introducing the second innovation is $Pr_2 = 1 - \varepsilon$, which is optimal. However, the probability of introducing the first innovation is also $Pr_1 = 1 - \varepsilon$, which is less than optimal. The reason is the same as in Section 8: the first innovator does not take into account the effect of her decision on the innovation possibilities of the second innovator.

With patents, innovator 1 sets a licensing fee p_1 to try to extract part of the surplus of innovator 2 (in this 2-period model, the patent and patent pool cases are the same). The probability of innovation of innovator 2 is $Pr_2 = 1 - \varepsilon - p_1$. Innovator 1 maximizes:

$$\max_{p_1} v_1 - \varepsilon + (1 - \varepsilon - p_1) p_1, \quad (16)$$

which leads to a price $p_1 = (1 - \varepsilon)/2$. The probabilities of innovation are $Pr_1 = \min\{1, \frac{(1-\varepsilon)(5-\varepsilon)}{4}\}$ and $Pr_2 = (1 - \varepsilon)/2$, so $Pr_1 \leq Pr_1^*$ and $Pr_2 < Pr_2^*$. This is due to the combined effects of asymmetric information and market power.

Therefore, the 2-period model presents a simplified version of the general model but still allows to capture the externality and asymmetric information problems.

9.2. *Second best: optimal transfers*

One way to correct the dynamic externality would be to allow for transfers from future innovators to current innovators. We have seen that patents fail to convey appropriate incentives because of asymmetric information. In this subsection we analyze what is the optimal transfer a government should set to maximize expected welfare when it does not have information on the value of innovations, and we compare it with that of the patents case.

Assume that the government does not know v_1 nor v_2 . In this case, the government cannot make the transfer depend on the realization of v_2 , and it will be impossible to reach the optimum. The best the government can do is

to set a transfer equal to t if innovator 2 innovates, and 0 otherwise.

In this case, innovator 1 will innovate if $v_1 + Pr_2 t \geq \varepsilon$, and innovator 2 will innovate if $v_2 \geq \varepsilon + t$. The government wants to maximize expected welfare:

$$\begin{aligned} W &= Pr_1 \left(E(v_1 - \varepsilon/v_1 + Pr_2 t \geq \varepsilon) + Pr_2 E(v_2 - \varepsilon/v_2 \geq \varepsilon + t) \right) \quad (17) \\ &= \frac{1}{2}(1 - \varepsilon)(2 - \varepsilon - t)(1 - \varepsilon + t(1 - t - \varepsilon)). \end{aligned}$$

Solving this problem we get that the second best transfer with asymmetric information is:

$$t^* = \frac{3 - 2\varepsilon - \sqrt{6 - 6\varepsilon + \varepsilon^2}}{3}, \quad (18)$$

where $t^* < p_1$. Therefore, even if the government does not know v_2 , it will set a lower transfer than the licensing fee of innovator 1 with patents. This is due to the combined effects of asymmetric information and market power with patents.

10. Ongoing innovation

In this section we analyze what happens if the sequence of innovations does not stop after one innovation fails to be performed. There is a sequence of innovations $n = 1, 2, \dots$, just as before, but now there can be many trials until an innovation is successful.

Innovator n, j is the j^{th} innovator trying to introduce innovation n ($j - 1$ previous innovators tried to introduce innovation n without success). The innovator has to pay licensing fees to $n - 1$ patentees (the $n - 1$ previous successful innovators), and obtains a draw v_{nj} from the same distribution as before. If the revenues from the innovation are higher than the cost, innovator n, j will introduce the innovation, and in the next stage, innovator $n + 1, 1$ will try to introduce innovation $n + 1$. If revenues are lower than cost, innovator n, j fails to introduce innovation n , which will then be tried

by innovator $n, j + 1$ in the following stage. This innovator will face the same $n - 1$ patent holders and will have a new draw of the value of innovation $v_{n,j+1}$.

For this model, we need to be more specific about the time dimension. Specifically, assume that stages correspond to time periods. At each period only one trial for one innovation is performed. The discount factor of innovator and patent holders is β . At stage $n + j$ the game is summarized by a state $\{n, j\}$.

Let J_{nj}^i be the expected future licensing revenues of patentee i at trial j of innovation n , given that stage $n + j$ has been reached under state $\{n, j\}$. Expressed in a recursive way:

$$J_{nj}^i = Pr_{nj} (p_{nj}^i + \beta J_{n+1,1}^i) + (1 - Pr_{nj}) \beta J_{n,j+1}^i, \quad (19)$$

where Pr_{nj} is the probability that innovation n is introduced in trial j . With probability Pr_{nj} , the patent holder gets the price p_{nj}^i plus the continuation value of J of the first trial of the next innovation, $J_{n+1,1}^i$, appropriately discounted by β . With probability $1 - Pr_{nj}$, the innovation will not be introduced and the patent holder gets the continuation value of J corresponding to the next trial of the current innovation.

The profit of the innovator is $I_{nj} (v_{nj} + \beta J_{n+1,1} - c_{nj} - \varepsilon)$.

Just as before, subgame perfection implies that the patent holders take $J_{n+1,1}^i$ and $J_{n,j+1}^i$ as given when deciding p_{nj}^i . The profit maximizing price is $p_{nj}^i = (1 - \varepsilon + \beta J_{n,j+1}^i)/n$. In a symmetric equilibrium, $p_{nj}^i = p_n$ and $J_{nj}^i = J_n$ for all i, j .

Replacing in the probability of innovation, we get $Pr_n = \frac{1-\varepsilon}{n} + \beta J_{n+1} - \frac{n-1}{n} \beta J_n$, and introducing in the expression for J_{nj}^i :

$$J_n = \frac{1}{1-\beta} \left(\frac{1-\varepsilon}{n} + \beta J_{n+1} - \frac{n-1}{n} \beta J_n \right)^2. \quad (20)$$

Rearranging this expression:

$$J_{n+1} = \frac{1}{\beta} \left(\sqrt{(1-\beta)J_n} - \frac{1-\varepsilon}{n} \right) + \frac{n-1}{n} J_n, \quad (21)$$

which is a decreasing sequence converging to 0 as $n \rightarrow \infty$.

The sequence in terms of probabilities is:

$$Pr_{n+1} = \frac{1-\beta}{\beta} \left(Pr_n + \frac{n-1}{n} \frac{\beta}{1-\beta} Pr_n^2 - \frac{1-\varepsilon}{n} \right)^{1/2}, \quad (22)$$

which is also a decreasing sequence converging to 0 as $n \rightarrow \infty$. Therefore, the main conclusions of the basic model still hold under when innovation does not stop when a single innovation fails.

11. Finite patents

In this section we analyze what happens if patents have finite length. Each stage corresponds to one period and only one innovation is attempted at each period. If the innovator decides to introduce the innovation, she obtains a patent for L periods. This means that the innovator has to pay patents for L previous innovations, but also charges licenses to L future innovators.

The main difficulty of the present analysis is that now the identity of the patent holders matters. The price and future expected licensing revenues will be different for different patent holders, depending on how long will it take for her patent to expire.

The innovator will introduce the innovation if the revenues from innovation are larger than the cost:

$$v_n + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k \geq \sum_{i=n-L}^{n-1} p_n^i + \varepsilon, \quad (23)$$

which means that the probability of innovation is

$$Pr_n = 1 + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k - \sum_{i=n-L}^{n-1} p_n^i - \varepsilon. \quad (24)$$

The L current patent holders differ in their objective functions. Let J_n^i be the future expected revenues of patent holder i at stage n , given that stage n has been reached. Then,

$$J_n^i = Pr_n(p_n^i + J_{n+1}^i). \quad (25)$$

The patent holder charging a license for the last time is patent holder $n - L$, so $J_{n+1}^{n-L} = 0$. The patent of $n - L + 1$, on the other hand, will last for one more period, so $J_{n+1}^{n-L+1} = Pr_{n+1} p_{n+1}^{n-L+1}$. In this way, we can construct the future expected revenues of the L patent holders.

The profit maximization problem is

$$\max_{p_n^i} J_n^i = Pr_n(p_n^i + J_{n+1}^i). \quad (26)$$

The first order condition is $-p_n^i - J_{n+1}^i + Pr_n = 0$, so $p_n^i + J_{n+1}^i = Pr_n$ and $J_n^i = Pr_n^2$ for all i . This also implies that $p_n^{n-L} = Pr_n$.

We are interested in stationary equilibria, which means that $Pr_n = Pr$ for all n . This, together with the first order condition, implies that $p_n^i = Pr(1 - Pr)$ for $i \geq n - L$. Replacing in the probability of innovation, we get:

$$\begin{aligned} Pr &= 1 + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k - \sum_{i=n-L}^{n-1} p_n^i - \varepsilon \\ &= 1 + \sum_{m=1}^{L-1} Pr(1 - Pr)Pr^m + PrPr^L - (L - 1)Pr(1 - Pr) - Pr - \varepsilon \end{aligned} \quad (27)$$

Solving for Pr , we get:

$$Pr = \frac{L + 1 - \sqrt{(L - 1)^2 + 4L\varepsilon}}{2L} \quad (28)$$

which is the stationary equilibrium probability of innovation.

Figure 5 shows the probability of innovation as a function of the patent length for $\varepsilon = 0.2$. We can see that the probability of innovation decreases with L , which means that patents hurt more than benefit the innovator. This is because the innovator has to pay licenses that are certain to the patent holders, but the future licensing revenues are uncertain, as they depend on future innovations being performed.

It is also interesting to see that $Pr \rightarrow 0$ when $L \rightarrow \infty$ and $Pr \rightarrow 1 - \varepsilon$ when $L \rightarrow 0$, which correspond to the previously analyzed patents and no patents cases (with $\phi = 1$).

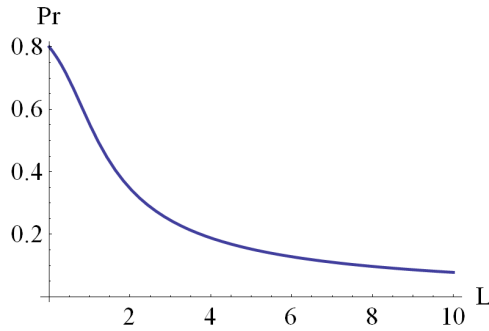


Figure 5: Probability of innovation and patent length.

11.1. Revenues depend on patent length

We have assumed that the revenues from selling the new product in the final goods market are independent of patent length. In this subsection we analyze what happens when we relax this assumption. Assume the revenues of the innovator are $\phi(L) v_n$, with $\phi'(L) \geq 0$, $\phi''(L) \leq 0$, $\lim_{L \rightarrow 0} \phi(L) = \underline{\phi}$

and $\lim_{L \rightarrow \infty} \phi(L) = 1$. Here, ϕ is the fraction of social surplus the innovator would appropriate without any patent protection, due to trade secrets or first mover advantages.

In this case, the innovator will innovate if

$$\phi(L) v_n + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k \geq \varepsilon + \sum_{i=n-L}^{n-1} p_n^i. \quad (29)$$

Applying a similar procedure as that of the previous case, we obtain the probability of innovation in the stationary equilibrium:

$$Pr = \frac{L + 1 - \sqrt{(L - 1)^2 + 4L\varepsilon/\phi(L)}}{2L}. \quad (30)$$

The effect of patent length on the probability of innovation depends on the functional form of $\phi(L)$. Let $\phi(L) = 1 - \frac{1-\phi}{(L+1)^\gamma}$, where γ measures the speed at which revenues grow when L increases. Figure 6a shows that when ϕ is more concave ($\gamma = 1$), the probability of innovation first increases and then decreases with patent length. The optimal length is positive and finite (in this case $L = 1$). Figure 6b shows that for a lower degree of concavity of $\phi(L)$ it is optimal to completely remove patents. Therefore, the results do not change significantly when the revenues in the final goods sector depend on patent length.

12. Conclusion

In this paper we build a dynamic model where accumulation of patents generates an increasing number of claims on cumulative innovation. The model is intended to reproduce the central feature of innovation activity in hi-tech industries: new products are more complex than old products, because they build on a larger stock of previously accumulated knowledge.

We study the policy that maximizes expected social welfare and compare it with the outcome of three patent policy regimes: patents, patent

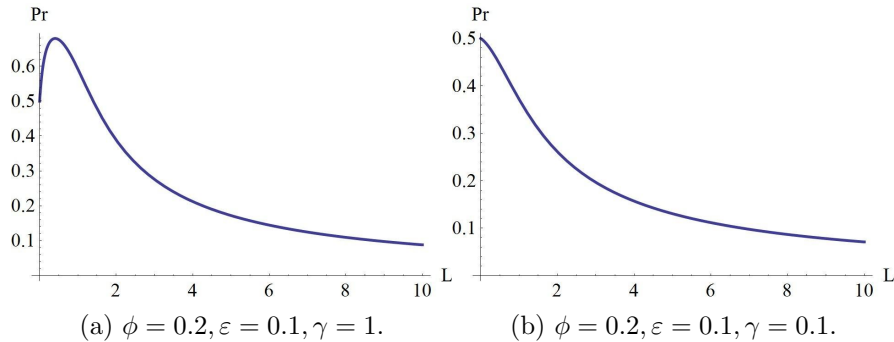


Figure 6: Probability of innovation as a function of patent length.

pools and no patents. We find that, even abstracting from the monopolistic inefficiencies of patents, none of these policies attains the optimum.

With patents, the innovator has to pay an increasing number of license fees to previous innovator. Asymmetric information on the value of the innovation and uncoordinated market power of licensor create an anticommons effect that reduces the incentives to innovate as innovation becomes more complex. The anticommons effect is weaker than in the static case, but it is still strong enough to drive the probability of innovation to zero as the number of licenses grows large. Enforcing a patent pool solves the lack of coordination but not the asymmetric information problem. As a result the outcome of patent pools is more desirable but still it does not achieve the first best. Eliminating patent protection solves the two problems but introduce a non-internalized externality: previous innovations set the foundations for future innovations. Therefore it might be the case that the social cost of one innovation is higher than its instantaneous social value (the social value the innovation creates *per se*), and yet the innovation is socially desirable because it allows the development of further innovations. This is the typical problem faced by basic research.

Then we study alternative solutions to the anticommons: (i) the complete sale of patent rights of each innovator to the next one, and (ii) the

possibility that the licensee compete with the original licensor. Both alternatives exactly replicate the sequence of innovations under the patent pool regime. Another interesting result of the paper is that patent pools are dynamically unstable, as the incentives to remain outside the pool increase as the sequence of innovations progresses.

We also find that the outcome of these three policy arrangements does not even attain the second best. The second best is achievable by means of government transfers, assuming that the government does not know the value of the innovation.

This paper shows that patent protection may be the wrong way to provide incentives to innovation in complex industries like electronics, software and hardware. Enforcing patent pools or eliminating patent protection would improve welfare, but still would not reach the social optimum. We hope this paper will contribute to future research on the design of an optimal innovation policy.

References

- Bergstrom, T. C., 1978. Cournot equilibrium in factor markets. Tech. rep., University of Michigan working paper.
- Boldrin, M., Levine, D. K., 2005. The economics of ideas and intellectual property. *Proceedings of the National Academy of Sciences* 102 (4), 1252–1256.
- Commision on Intellectual Property Rights, 2002. Integrating intellectual property rights and development policy: Report of the commision on intellectual property rights. London.
- Cournot, A., 1838. *Researches Into the Mathematical Principles of the Theory of Wealth*. Irwin (1963).

- Garfinkel, S., Stallman, R., Kapor, M., 1991. Why patents are bad for software. *Issues in Science and Technology* 8 (1), 50–55.
- Graff, G. D., Cullen, S. E., Bradford, K. J., Zilberman, D., Bennett, A. B., September 2003. The public-private structure of intellectual property ownership in agricultural biotechnology. *Nature Biotechnology* 21 (9), 989–995.
- Heller, M. A., Jan. 1998. The tragedy of the anticommons: Property in the transition from marx to markets. *Harvard Law Review* 111 (3), 621–688.
- Heller, M. A., Eisenberg, R. S., May 1998. Can patents deter innovation? the anticommons in biomedical research. *Science* 280 (5364), 698–701.
- Holmstrom, B., 1982. Moral hazard in teams. *The Bell Journal of Economics* 13 (2), 324–340.
URL <http://www.jstor.org/stable/3003457>
- Hopenhayn, H., Llobet, G., Mitchell, M., 2006. Rewarding sequential innovators: Prizes, patents, and buyouts. *Journal of Political Economy* 114 (6), 1041–1068.
URL <http://www.journals.uchicago.edu/doi/abs/10.1086/510562>
- Lerner, J., Tirole, J., June 2004. Efficient patent pools. *American Economic Review* 94 (3), 691–711.
- Llanes, G., Trento, S., Jun. 2009. Anticommons and optimal patent policy in a model of sequential innovation. Working Paper 09-148, Harvard Business School.
- Scotchmer, S., Winter 1991. Standing on the shoulders of giants: Cumulative research and the patent law. *Journal of Economic Perspectives* 5 (1), 29–41.
- Shapiro, C., 2001. Navigating the patent thicket: Cross licenses, patent pools, and standard setting. In: Jaffe, A. B., Lerner, J., Stern, S. (Eds.), *Innovation policy and the economy*. Vol. 1. MIT Press for the NBER, pp. 119–150.

Sonnenschein, H., 1968. The dual of duopoly is complementary monopoly: or, two of cournot's theories are one. *The Journal of Political Economy* 76 (2), 316–318.