HARVARD | BUSINESS | SCHOOL

Anticommons and Optimal Patent Policy in a Model of Sequential Innovation

Gastón Llanes
Stefano Trento

Working Paper

09-148

ANTICOMMONS AND OPTIMAL PATENT POLICY IN A MODEL OF SEQUENTIAL INNOVATION

GASTÓN LLANES[†] AND STEFANO TRENTO[‡]

ABSTRACT. We present a model of sequential innovation in which an innovator uses several research inputs to invent a new good. These inputs, in turn, must be invented before they can be used by the final innovator. As a consequence, the degree of patent protection affects the revenues and cost of the innovator, but also determines the incentives to invent the research inputs in the first place. We study the effects of increases in the number of required inputs on innovation activity and optimal patent policy. We find that the probability of introducing the final innovation decreases (increases) as the number of inputs increases when inputs are complements (substitutes). We also find that the optimal strength of patents on research inputs is increasing in the degree of substitution between the inputs, but decreasing in the number of inputs for any degree of substitution.

1. INTRODUCTION

Knowledge builds upon previous knowledge. This is true for most innovations nowadays, especially in hi-tech industries like molecular biology, plant biotechnology, semiconductors and software. In some cases, the innovation consists of an improvement of an older version of the same good. In other cases, the research leading to the discovery of the new good depends on the access to research tools, techniques and inputs which are previous innovations themselves.

In any case, innovation activity will in general depend on the access to previous innovations. Depending on the structure of the patent

Date: June 24, 2009.
[†] Harvard Business School, gllanes@hbs.edu.
[‡] Universitat Autònoma de Barcelona, stefano.trento@uab.es.
We are grateful to Michele Boldrin for his guidance and advice. We thank David Levine and participants of seminars at Universidad Carlos III de Madrid and Washington University in St. Louis 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).

system, some of these inventions will be protected by patents. This means that patents affect not only the revenues of the innovator, but also the cost of performing an innovation.

Recent concern has arisen on the possibility that patents (or other kinds of Intellectual Property) can restrict access to research inputs, hindering innovation as a consequence. The innovator and the owners of patents on previous inventions share the revenues of the innovation. As the number of inputs needed in research increases, the innovator faces a *patent thicket* and is threatened by the possibility of *hold-up*, namely the risk that a useful innovation is not developed because of lack of agreement with the patent holders. This problem has been dubbed the *tragedy of the anticommons* (Heller 1998, Heller and Eisenberg 1998). When too many agents have exclusion rights over the use of a common resource, this resource tends to be underutilized, in clear duality with the tragedy of the commons in which too many agents hold rights of use and the resource tends to be overused.

This problem may be particularly acute in biomedical research, where there is a deep controversy over the patenting of gene fragments and research tools. Take for example the case of the MSP1 antigen (Plasmodium Falciparum Merozoite Specific Protein 1), widely recognized as the most promising candidate for an anti-malarial vaccine. A study of the Commission on Intellectual Property Rights (2002) found more than 39 patent families covering DNA fragments, methods for processing fragments, production systems, vaccine delivery systems, etc. As a consequence, a potential innovator willing to commercialize a vaccine based on MSP1 must get prior permission from the owners of these property rights.

Anticommons can arise in biotechnology as well. A good example is Golden Rice, which required payment of up to 40 licenses, depending on the country of commercialization (Graff, Cullen, Bradford, Zilberman, and Bennett 2003).

As a final example, consider the case of software patents, which cover mathematical algorithms and techniques. Software programs have become so complex that any single program may use thousands of algorithms (Garfinkel, Stallman, and Kapor 1991), possibly infringing a large number of patents. This explains the expected increase in patent litigation in this sector in the next years (think of Microsoft vs. the programmers and users of Linux), and the formation of a Patent Commons by firms involved in the Open Source community (IBM, HP, Novel, Sun, etc.).

We address these issues by constructing a model of sequential innovation in which an innovator uses n patented inputs to develop a new invention. Substitutability between the inputs goes from zero (perfect complements) to infinity (perfect substitutes) and the input sellers compete in prices but do not know the exact value of the innovation for the innovator.

We study how the probability of performing the innovation changes as technologies become more complex (n increases) and find that it decreases when the inputs are market complements and increases when they are market substitutes. Therefore, we prove that the anticommons hypothesis may hold when inputs are essential and not easy to substitute. Then, we analyze the limiting economy when $n \to \infty$, and show that the probability of innovation is always less than socially optimal unless the inputs are perfect substitutes. Moreover, the probability of innovation goes to zero when the elasticity of substitution is below a threshold level which is higher than 1.

In the basic model, the research inputs have been already invented, and are protected by strong patents. The government could reduce the patent thicket by granting less patents, or by reducing patent breadth. However, weaker patents imply a reduced incentive to discover research inputs in the first place. In Section 4, we turn to the analysis of optimal patent policyr. Patent policy affects the division of profits between the input innovators and the final innovator. We obtain the patent policy that maximizes the joint probability of inventing all the inputs and the final good. We find that this optimal policy is decreasing in complexity: when the number of inputs required in the R&D process increases, the optimal response is to reduce patent strength.

There is an extended literature on Sequential Innovation, which is mainly concerned with the optimal division of profits between sequential innovators. The main contribution of this paper is to extend the analysis of patent policy and optimal division of profit when innovation activity requires access to several prior inventions. As such, the two key insights of the model are: (i) patents will be more harmful when research inputs are complementary, and (ii) optimal patent policy should decrease when innovations become more complex.

As we discuss in the following section, our paper is also related to the literatures of complementary monopoly and patent pools. A patent pool is a cooperative agreement among patent holders, through which they agree on the licensing terms of a subset of their patents. In Section 5, we analyze the effects of the formation of patent pools on innovation activity within our model, and study similarities and differences of our findings with previous papers in this literature.

1.1. **Related literature.** There is an extended literature on Sequential Innovation (Scotchmer 1991, Green and Scotchmer 1995, Chang 1995, Scotchmer 1996, Hopenhayn, Llobet, and Mitchell 2006), which is mainly concerned with the optimal division of profits between sequential innovators. Generally, in these models, there are two innovations that have to be introduced sequentially (the second innovation cannot be introduced until the first one has been introduced), and the objective is to find the patent policy that maximizes the incentives to invest in both innovations. In this paper we generalize these models by assuming that any innovation is based on n previous innovations. We find that the effect of increasing complexity on innovation activity depends on the degree of substitutability/complementarity between these previous innovations.

Our contribution is to extend the analysis of optimal patent policy and optimal division of profits to the case in which the innovator requires access to *several* prior inventions. In doing so, we link the literature of sequential innovation to the literatures of complementary monopoly and patent pools.

Cournot (1838) was the first to analyze the complementary monopoly problem. He modeled a competitive producer of brass who must buy zinc and copper from two separate monopolists (zinc and copper are perfect complements), and showed that (i) the price of the inputs is higher than the price that a single provider would set, (ii) the total cost of the inputs is increasing in the number of inputs, (iii) in the limit, as $n \to \infty$ the cost of the inputs is such that the demand for the final good is zero.

This analysis has been later extended in various directions. Sonnenschein (1968) showed that the Cournot's theory of complementary monopoly is the dual of its theory of duopoly with quantity competition and homogeneous goods, when the marginal cost of production is zero. Bergstrom (1978) built upon this intuition to extend the duality result to imperfect substitutes. Recently, Boldrin and Levine (2005a) applied the structure of complementary monopoly to a model of sequential innovation: they set up a model of one innovation building upon a number of (perfect complementary) previous innovations, and show that as this number increases the cost of innovating also increases, eventually hitting the upper bound for which the innovation is no longer profitable. We extend the analysis to a framework of oligopoly with any degree of substitutability. This generalization is relevant because it allows us to study the joint effect of the number of oligopolists and the degree of substitutability on the probability of innovation. In particular we find that, when the research inputs are market complements (i.e. not only when they are perfect complements), the probability of innovation decreases in the complexity of innovation.

Cournot's theory has also been used by the literature on patent pools. Shapiro (2001) was the first to suggest that patent pools may be anticompetitive when they are formed by substitute patents, and procompetitive when formed by complementary patents. Lerner and Tirole (2004) built a model to generalize Shapiro's results for imperfect substitutability. Given the similarity between our model and Lerner and Tirole's, in Section 5 we present an analysis of the effect of patent pools on innovation. We find that Lerner and Tirole's results generalize to our case also. However, given that in our model input decisions are continuous, we can be more precise in our definition of complementarity, which complements previous analysis.

However, it is important to remark that, even though our model is related with Lerner and Tirole's, our focus is completely different. We are interested in analyzing the effect of increases in n on the optimal patent policy, while they are interested in the effects of price collusion on innovation and welfare for n fixed.

2. The model

There are n research inputs $(x_1, ..., x_n)$ and a potential innovator who may use the n inputs in R&D in order to invent a new good. The n inputs have already been invented and are ready to be produced. We make this assumption in order to concentrate on the effects of the pricing of old ideas on innovation activity. In Section 4 we will extend the model to see what happens when the n inputs have to be invented as well.

The structure of Intellectual Property Rights is such that each input is protected by a patent, granting its owner a monopoly over it. Each patent is owned by a different patentee and thus each of the n inputs is supplied by a different producer. Given that the inputs are imperfect substitutes of each other, the factor market is a differentiated goods oligopoly. The input sellers compete in prices and the value of the innovation is private information of the innovator.

2.1. Technology.

The innovator can perform R&D to invent a new final good according to the following CES technology:

$$y = A \left(\sum_{i=1}^{n} {x_i}^{\rho} \right)^{\frac{1}{\rho}}, \tag{1}$$

where y is a measure of the R&D effort, A is a scale parameter, x_i is the amount of input i used, n is the number of inputs and $\rho \in (-\infty, 1]$ is a technological parameter related to the substitutability between the inputs.

We will find easier to work with the elasticity of substitution $\sigma = 1/(1-\rho)$. As is well known, $\sigma = 0$ represents perfect complements, $0 < \sigma < \infty$ represents imperfect substitutes, and $\sigma \to \infty$ represents perfect substitutes.

The innovator faces an indivisibility problem, meaning that a minimum amount of R&D effort is required to invent a new good. When the R&D effort is below that threshold level there is no innovation. Without loss of generality we can set the threshold level at 1, so that the indicator function for the innovation is:

$$I = \begin{cases} 1 & \text{if } y \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

We set the scale parameter A in (1) equal to $n^{(\rho-1)/\rho}$ in order to eliminate any returns from specialization or division of labor. Usually CES production functions exhibit a property called increasing returns to specialization (or love for variety in the case of utility functions). Following an argument similar to Romer (1987), suppose that the production function is $y = \left(\sum_{i=1}^{n} x_i{}^{\rho}\right)^{1/\rho}$, and let X be the total quantity of inputs used in production. Because of symmetry, all inputs will be used in the same quantity in the equilibrium, so $x_i = X/n$ for all i, and output will be equal to $y = n^{(1-\rho)/\rho} X$. There are positive returns to specialization because an increase in n holding X constant causes output to increase. We eliminate this effect by introducing $A = n^{(\rho-1)/\rho}$ in the production function.

The complexity of the innovation is measured by n. More complex technologies use a larger number of components or require more research tools in order to be developed. Each input is produced with a constant marginal cost of $\varepsilon > 0$. We assume that the resources used to produce inputs are sold in a competitive market, so that the private and the social cost of producing inputs coincide. The assumption of no returns to specialization guarantees that the social cost of performing the innovation does not change as technologies become more complex. In other words, there is no technological advantage or disadvantage from increases in n.

In section 2.4 we will provide the intuition behind our innovation technology. Before we do it, let us continue with the description of the model.

2.2. **Value of the innovation and structure of the information.** The social value of the innovation, v, is the total surplus generated by the new product. To focus on the factor market, we will assume that the innovator is a perfect price discriminator in the final goods market. This means that the revenue of the innovator coincides with the social value of the innovation.

The value of the innovation is private information of the innovator. This may be because the innovator has better information about the characteristics of the new product or about the valuation of the consumers. The sellers of inputs only know that v has a cumulative distribution $F(v)$. Therefore $F(v)$ is the probability that the innovation has a return less or equal to v. In Section 6 we show that the assumptions of perfect price discrimination and asymmetric information can be relaxed without altering the results.

The hazard function is defined as $h(v) = f(v)/(1-F(v))$, where $f(v)$ is the density function corresponding to $F(v)$. In order to guarantee the quasi-concavity of the maximization problem of the input producers, the following assumption will hold throughout the paper:

Assumption 1 (Nondecreasing hazard function)**.** *$h(v) > 0$ and $h'(v) \geq 0$ on a support $[\underline{v}, \bar{v}]$, and $h(v) = 0$ outside of this support.*

This assumption on the hazard function is very general, and holds for most continuous distribution functions. Notice that we are not restricting $\underline{v}$ nor $\bar{v}$ to be of finite value.

An important assumption is that the distribution of values of the innovation does not change with n. This assumption, together with the absence of returns to specialization in the R&D technology imply the following lemma:

Lemma 1. *The probability that an invention is socially optimal does not depend on its complexity.*

Proof. The probability that an innovation is socially optimal is the probability that its social value is larger or equal than its social cost. The social cost of an innovation coincides with the resources used to produce it. Therefore, the probability that an innovation is socially optimal is $Prob\,(v - \sum_{i=1}^{n} \varepsilon\, x_i \geq 0)$. Because of the symmetry in the innovation technology, $x_i = 1/n$, so this probability becomes $1 - F(\varepsilon)$, which depends on the distribution of social values of the innovation and the marginal cost of the inputs but *not* on the number of inputs used in R&D. ∎

In this paper, we are interested in studying the effects of increasing technological complexity on the probability of innovation. Lemma 1 assures that a change in n affects this probability only through a change in the number of inputs used in research, but not through a change in the social value or cost of the innovation. In other words, we want to compare innovations with different n but the same net social value. In Section 6 we relax these assumptions by letting the value of the innovation be a function of n and allowing returns to specialization in the R&D technology. We find that the main results of the paper are not significantly affected by a change in these assumptions.

2.3. **Market interaction.** The players of the game are the n input sellers and the innovator. A strategy for input seller i is a choice of price for her input. A strategy for the innovator is a function $g : \mathbb{R}^n_+ \times v \rightarrow \mathbb{R}^n_+$, namely a demand x_i for each input, as a function of the price of all the inputs and the realization of the value of the innovation.

The timing of the game is as follows: (i) the input producers simultaneously set the price of their inputs, (ii) Nature extracts a value v from the distribution $F(v)$, and (iii) given prices, the innovator calculates the input mix that minimizes the cost of innovation and decides whether to innovate or not.

The equilibrium concept we use is Symmetric Subgame Perfect Equilibrium (SSPE). A set of strategies $\{p_i\}_{i=1}^n, g$ is a SSPE if it is a Nash equilibrium of every subgame of the original game, and $p_i = p$ for all i.

The payoff for input producer i is $x_i\,(p_i - \varepsilon)$ and the payoff of the innovator is $I\,v - \sum_{i=1}^n p_i\,x_i$.

2.3.1. *Innovator's Problem.* Given input prices $\{p_i\}_{i=1}^n$, the innovator solves the following Cost Minimization Problem (CMP):

$$\begin{aligned} c &= \min \sum_{i=1}^{n} p_i x_i \\ s.t. \quad & n^{-\frac{1-\rho}{\rho}} \left(\sum_{i=1}^{n} {x_i}^{\rho} \right)^{\frac{1}{\rho}} \geq 1. \end{aligned}$$

The solution to this problem is the set of conditional factor demands x_i and the minimum cost of innovation c. Given c, the innovator will perform the innovation $(I = 1)$ if $v \geq c$.

2.3.2. *Input Seller's Problem.* When setting the price the sellers of inputs do not know the realization of v. They only know that given $\{p_i\}_{i=1}^n$ the probability that $v \geq c$ (the probability of innovation) is $1 - F(c)$. Therefore, the expected demand of input firm i is $E(x_i) =$

$(1-F(c))\, x_i$, and its Profit Maximization Problem (PMP) is:

$$\max_{p_i} \quad \Pi_i = (1-F(c))\, x_i (p_i - \varepsilon),$$

where both c and x_i come from the CMP of the innovator.

2.4. **Interpretation of the innovation technology.** The CES specification of the R&D technology is a simple and general way to introduce substitutability and complementarity between the inputs used in research. In our model, ideas have economic value because they are embodied in physical objects (Romer 1990, Boldrin and Levine 2002, 2005b). The innovator uses these physical objects to innovate, not the abstract ideas. Accordingly, the input decision is not discrete (to use the idea or not), but rather continuous (the research inputs can be used in variable amounts).

Our description of the R&D process is a good description of many innovation technologies. First, we can think that the inputs are used in R&D in variable amounts, and once the innovation is performed they are no longer needed. This interpretation fits well for sectors that use a large number of research tools, like biomedical research, where the use of clones and cloning tools, laboratory equipment and machines, reagents, computer software and many other research instruments is required in R&D, and can be used in variable amounts.

The second interpretation is that inputs are actually components of the final innovation, and are used to produce each copy of it. This interpretation resembles more the case of already patented code lines used in new software, hardware components for computers, and a variety of cases in electronics, semiconductors and other similar industries. In the early radio industry, for example, according to 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", like the high-frequency alternator, high-frequency transmission arc, magnetic amplifier, the crystal detector, diode and triode valves, directional aerial, etc.

Under this second interpretation there is a continuum of perfectly competitive innovators. The production function of output is $y = \left(\sum_{i=1}^{n} n^{\rho-1}\, {x_i}^{\rho}\right)^{1/\rho}$ and perfect competition assures that the price of the final output is equal to its marginal cost $c = \sum_{i=1}^{n} p_i\, x_i$. Then, the demand of the final good is $y = 1 - F(c)$, where y and c are the quantity and price of the final good.

The two interpretations lead to exactly the same results, except that in the second one there is a welfare loss from the anti-competitive pricing in the inputs market (the welfare loss would be approximately equal to $(c-\varepsilon)(F(c)-F(\varepsilon))/2$. For expositional purposes, we will stick to the first interpretation.

Obviously, our research technology is not a good description of all possible innovation processes. In particular, it does not fit the case in which the innovator pays only for the permission to use an idea. An elegant alternative formulation is to consider discrete input choices (1 if the input is used and 0 otherwise). This is the approach followed by Lerner and Tirole (2004). In this case, the innovator uses either embodied or abstract ideas, and pays input innovators to have access to these ideas.

In Section 5 we show that our model can be interpreted as a "smooth" version of Lerner and Tirole's. Plainly, the advantage of our approach is tractability. Continuity allows us to use differential calculus, which greatly simplifies our subsequent analysis of the effect of an increase in the number of inputs. The best argument to show that our results would still hold if we were to assume discrete input choices is to notice the similarity of the conclusions of both models on the effect of patent pools when n is fixed.

3. Equilibrium

In this section we solve recursively for the SSPE. Therefore, we begin by solving the Innovator's Problem (second stage of the game). The demands are those of a typical CES production function. The conditional demand of input i and the cost of innovation are:

$$\begin{aligned} x_i &= I\, n^{-\frac{1}{1-\sigma}}\, p_i^{-\sigma} \left(\sum_{i=1}^{n} {p_i}^{1-\sigma}\right)^{\frac{\sigma}{1-\sigma}}, \\ c &= n^{-\frac{1}{1-\sigma}} \left(\sum_{i=1}^{n} {p_i}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \end{aligned}$$

where $\sigma = 1/(1-\rho)$ is the elasticity of substitution between the inputs. The innovator will introduce the new good (I=1) if $v \geq c$.

The restrictions on ρ imply that the elasticity of substitution σ goes from 0 (perfect complements) to ∞ (perfect substitutes).

Given x_i and c, the symmetric equilibrium price p solves

$$p = \underset{p_i \geq \varepsilon}{\operatorname{argmax}}\ (1-F(c))\ x_i\,(p_i-\varepsilon)\,,$$

where $c = n^{-\frac{1}{1-\sigma}}\left(p_i^{1-\sigma}+(n-1)p^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ and $x_i = n^{-1}\, p_i^{-\sigma}\, c^{\sigma}$. It is useful to notice that in the symmetric equilibrium ($p_i = p$ for all i), $c = p$ and $x_i = 1/n$ for all i. Also, $p \geq \varepsilon$ in equilibrium because otherwise firms would be making negative profits and would find it profitable to deviate by setting a higher price.

Because of the nature of Nash equilibria, for any value of n, ε, and $\sigma < \infty$ there exists equilibria where p is so high that the probability of innovation is zero (i.e. profits are zero for all input sellers) but any deviation by a *single* input seller is not enough to make it positive. However, these are trivial equilibria coming from the definition of Nash equilibria without any intrinsic economic value. We are interested in the existence of equilibria with a positive probability of innovation ($p < \bar{v}$).

The following proposition characterizes the solution of the first stage of the game (the Input Seller's Problem).

Proposition 1. *A SSPE with positive probability of innovation ($p < \bar{v}$) exists and is unique. The equilibrium price solves*

$$(p - \varepsilon)\, h(p) = n - \sigma\,(n-1)\,(p-\varepsilon)/p.$$

The conditional input demand is $x = 1/n$, the cost of innovation is $c = p$ and the probability of innovation is $1 - F(p)$.

Proof. The firm wants to maximize $(1-F(c))\, x_i\, (p_i - \varepsilon)$. The derivative with respect to price is:

$$D(p_i) = -f(c)\,\frac{\partial c}{\partial p_i}\, x_i\, (p_i - \varepsilon) + (1 - F(c))\left(\frac{\partial x_i}{\partial p_i}\,(p_i - \varepsilon) + x_i\right).$$

By Shepard's Lemma $\partial c/\partial p_i = x_i$, and by symmetry $c = p$, $x_i = 1/n$ and $\partial x_i/\partial p_i = -(n-1)\sigma/(n^2\, p)$. Therefore, the first order condition becomes:

$$D(p) = -f(p)\,\frac{p-\varepsilon}{n^2} + (1 - F(p))\left(-\frac{\sigma\,(n-1)\,(p-\varepsilon)}{n^2\, p} + \frac{1}{n}\right).$$

Now we prove that the solution cannot be ε nor $\bar{v}$ for $n < \infty$. $p = \varepsilon$ cannot be the equilibrium because $D(\varepsilon) = (1 - F(\varepsilon))/n > 0$. Also, $p = \bar{v}$ cannot be the equilibrium both if $\bar{v}$ is finite or infinite. If $\bar{v} < \infty$, then $D(\bar{v}) = -f(\bar{v})\left(\frac{\bar{v}-\varepsilon}{n}\right) < 0$. On the other hand, $lim_{p\to\infty} D(p) = -\infty < 0$. Therefore, the solution must satisfy $D(p) = 0$. Multiplying $D(p)$ by $-n^2/(1 - F(p))$ we get:

$$h(p)\,(p-\varepsilon) + \sigma\,(n-1)\,\frac{p-\varepsilon}{p} - n = 0. \tag{2}$$

We can be sure that equation (2) has exactly one solution because it is continuously increasing in p by Assumption 1, is negative when $p = \varepsilon$ and is positive when $p \to \bar{v}$ (Assumption 1 implies that $\lim_{p\to\bar{v}} h(p)\, p =$

∞ for finite or infinite $\bar{v}$). Therefore, the solution exists and is unique. Rearranging terms in equation (2) we get the desired result. ∎

Example. We will find it useful to illustrate the results with the help on an example based on the uniform distribution. This example has the advantage of providing an explicit solution for the equilibrium price. Specifically, assume that the value of the innovation, v, is uniformly distributed between 0 and 1. This means that $F(v) = v$ and $h(v) = 1/(1-v)$. The equilibrium price is:

$$p = \frac{a + \sqrt{a^2 + 4\,\sigma\,\varepsilon\,(n-1)\,b}}{2\,b}, \tag{3}$$

where $a = n + \varepsilon - \sigma(n-1)(1+\varepsilon)$ and $b = 1 + n(1-\sigma) + \sigma$. The cost of innovation is equal to the price and the probability of innovation is simply $1 - p$.

3.1. **Elasticity of substitution.** The price of the inputs and the cost of innovation in equilibrium depend on the elasticity of substitution, the complexity of the innovation and the marginal cost of the inputs. In the following subsections we will analyze the comparative statics of the above equilibrium.

Proposition 2. *The cost of innovation is decreasing in σ.*

Proof. Equation (2) provides an implicit function of p in terms of σ. We can calculate $\partial c/\partial\sigma$ using the implicit function theorem (remember that p=c in the symmetric equilibrium):

$$\frac{\partial c}{\partial \sigma} = -\frac{(n-1)(p-\varepsilon)/p}{h(p) + h'(p)\,(p-\varepsilon) + \sigma\,(n-1)\varepsilon/p^2}$$

It is easy to see that this derivative is always negative (the numerator and the denominator are positive). The result follows. ∎

Figure 1 depicts the cost of innovation (i.e. the price of the inputs) as a function of σ for the uniform distribution and for $n = 10$ and $\varepsilon = 0.1$.

The cost of innovation is monotonically decreasing in σ because of increased competition as the inputs become more substitutable. As $\sigma \to \infty$ price converges to marginal cost ε, which is the standard Bertrand price competition result with homogeneous goods.

3.2. **Complements and Substitutes.** We will classify inputs in market complements and substitutes according to the sign of the cross-price derivative of expected demand which, in this setting, is equivalent to

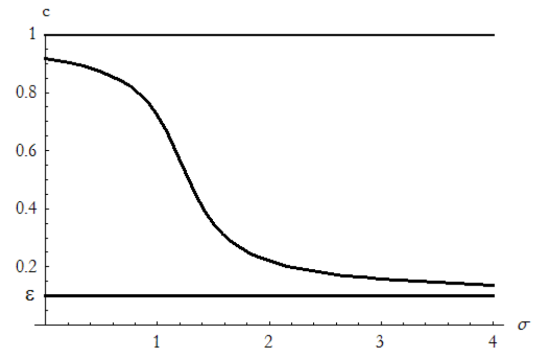


FIGURE 1. Cost of innovation as a function of σ.

analyzing the cross-price derivative of expected profit. This classification is equivalent to the one used in game theory, where the actions of two agents are said to be complements (substitutes) when an increase in the action of one of them implies a decrease (increase) in the payoff of the other agent. In our model, the actions are prices and the payoff is expected profit. Notice that this is an equilibrium definition since it is based on the best response of the innovator.

Definition 1 (Market complements and substitutes)**.** *Input j is a market complement (substitute) of input i if* $\frac{\partial E(x_i)}{\partial p_j} < 0$ $\left(\frac{\partial E(x_i)}{\partial p_j} > 0\right)$.

An increase in the price of input j has two effects on the expected demand of input i. On one hand, the conditional demand of input i increases (substitution effect). On the other hand, the probability of innovation decreases because the inputs are more expensive to the innovator (innovation effect). The sign of the cross-price derivative depends on which of the two effects is stronger. The cross-price derivative is:

$$\frac{\partial E(x_i)}{\partial p_j} = (1 - F(c))\,\frac{\partial x_i}{\partial p_j} + \frac{\partial(1 - F(c))}{\partial p_j}\,x_i.$$

The first effect is related to the standard substitution effect of consumer demand theory. Remember that the Cost Minimization Problem is equivalent to the Expenditure Minimization Problem and in this case there are no wealth effects of price changes (the conditional factor demands are equivalent to Hicksian demands). In principle, the derivative $\partial x_i/\partial p_j$ could be positive, negative or zero. However, the property of negative semidefiniteness of the matrix of cross-price derivatives (which implies that every input must at least have one technical substitute), together with the symmetry of the production function, implies that this derivative is non-negative. The inputs will be technical substitutes ($\partial x_i/\partial p_j > 0$) except in the case of perfect complements, where $\partial x_i/\partial p_j = 0$.

The second effect is due to the fact that the demand for inputs is downward sloping in the cost of innovation. The cost of the inputs used in research affects the profitability of innovation. Therefore, an increase in the price of any input will lower the probability of innovation. This effect is negative, except in the case of perfect substitutes, when it is zero.

Now that our definition of complementarity and substitutability is clear, we can be precise in our exposition. In what follows when we say that inputs are complements or substitutes, we mean that they are *market* complements or substitutes. We will see that the distinction between complements and substitutes is crucial for the predictions of the model.

The following lemma shows the value of σ that makes the cross-price derivative equal to zero.

Lemma 2. *The cross-price derivative $\partial E(x_i)/\partial p_j$ is zero in the symmetric equilibrium if and only if $\sigma = \sigma^*$, where σ^* is the argument that solves $h\left(\frac{\sigma}{\sigma-1}\varepsilon\right) = \sigma - 1$.*

Proof. The cross-price derivative is:

$$\frac{\partial E(x_i)}{\partial p_j} = (1 - F(c))\,\frac{\partial x_i}{\partial p_j} - f(c)\,x_i.$$

By Shepard's Lemma, $\partial c/\partial p_j = x_j$. Imposing symmetry, $x_i = x_j = 1/n$ and $\partial x_i/\partial p_j = \sigma/(n^2\,p)$. Rearranging terms we get:

$$\frac{\partial E(x_i)}{\partial p_j} = \frac{1}{n\,(1 - F(p))}\left(\frac{\sigma}{p} - h(p)\right).$$

This will be zero in the symmetric equilibrium only when $h(p) = \sigma/p$. Introducing this into the first order condition (2) and rearranging we get $(p - \varepsilon)/p = \sigma^{-1}$ or $p = \sigma\varepsilon/(\sigma - 1)$. Plugging this value of p in $h(p) = \sigma/p$ we get the desired result. ■

The following proposition classifies inputs in complements and substitutes according to Definition 1. It is interesting to see that this distinction depends on the values of σ and ε, but not on the value of n.

Proposition 3. *In the symmetric equilibrium, inputs are complements when $\sigma < \sigma^*$ and substitutes when $\sigma > \sigma^*$.*

Proof. We know from Lemma 2 that the cross-price derivative is zero when $\sigma = \sigma^*$ and that its sign depends on $\sigma/p - h(p)$. The latter expression is increasing in σ because p is decreasing in σ from Proposition 2 and h is non-decreasing in p from Assumption 1. The result follows. ■

Interestingly, the value of σ which divides inputs in complements and substitutes has to be larger or equal than 1. To see this, suppose that $\sigma^* < 1$. This means that $h(\varepsilon\sigma^*/(\sigma^* - 1)) < 0$, which is not possible. In the case of the uniform distribution, for example, inputs are complements when $\sigma < (1+\varepsilon)/(1-\varepsilon)$ and substitutes when $\sigma > (1+\varepsilon)/(1-\varepsilon)$.

3.3. **Increasing complexity.** Proposition 4 shows that the sign of the effect of an increase in the complexity of the innovation (n) depends on whether the inputs are complements or substitutes.

Proposition 4. *The cost of innovation increases as innovation becomes more complex if the inputs are complements ($\sigma < \sigma^*$) and decreases if the inputs are substitutes ($\sigma > \sigma^*$).*

Proof. We are looking for the effect of a unit increase in n, but it will suffice to determine the sign of $\partial c/\partial n$. Equation (2) provides an implicit function of c in terms of n. Therefore, we can calculate $\partial c/\partial n$ using the implicit function theorem:

$$\frac{\partial c}{\partial n} = \frac{1 - \sigma\,(p-\varepsilon)/p}{h'(p)\,(p-\varepsilon) + h(p) + \sigma(n-1)\varepsilon/p^2}$$

We know that the denominator is always positive. Therefore, the sign of this derivative depends on the sign of the numerator.

From equation (2) we get the following relation in equilibrium $\sigma\,(p-\varepsilon)/p = (n - h(p)\,(p-\varepsilon))/(n-1)$. Introducing this in the numerator and operating, it becomes $(h(p)\,(p-\varepsilon) - 1)/(n-1)$. We know that $h(p)\,(p-\varepsilon) = 1$ when $\sigma = \sigma^*$ from the proof of Proposition 8. Given that $h(p)\,(p-\varepsilon)$ is increasing in p, it is decreasing in σ. Therefore, the numerator is positive when $\sigma < \sigma^*$ and it is negative when $\sigma > \sigma^*$. The result follows. ■

The probability of innovation is simply $1 - F(c)$, so it moves in an opposite direction to the cost:

$$\frac{dPr}{dn} = -f(c)\,\frac{dc}{dn}.$$

As before, the effect on the probability of innovation of an increase in the complexity of innovation depends on the substitutability between the inputs. If inputs are complements, then the probability decreases as n increases. If inputs are substitutes, then the probability increases as n increases.

Figure 2 shows what happens in the uniform distribution example as the complexity of the innovation increases from $n = 5$ to $n = 15$, for $\varepsilon = 0.1$. The cost schedules cross when $\sigma = 1.22$, which is exactly $\sigma^* = (1+\varepsilon)/(1-\varepsilon)$. This means that the cost of innovation increases if the inputs have low substitutability and decreases in case of high substitutability.

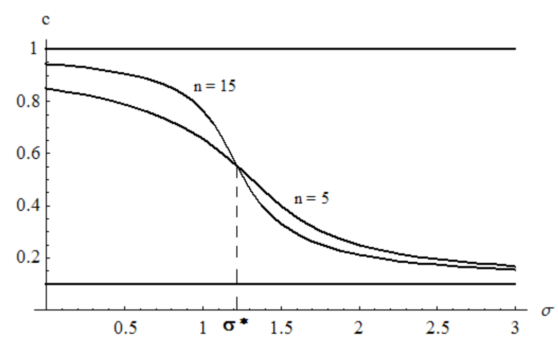


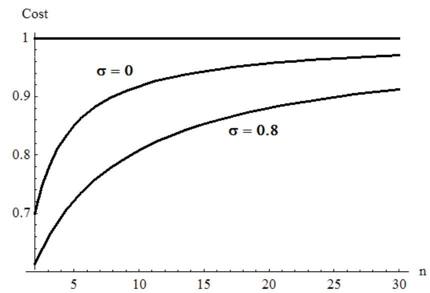
FIGURE 2. Effects of an increase in the complexity of innovation.

Proposition 4 says that patents are very harmful when innovation is sequential and the research inputs are essential or difficult to substitute, but do not pose an important problem when inputs are easily replaceable.

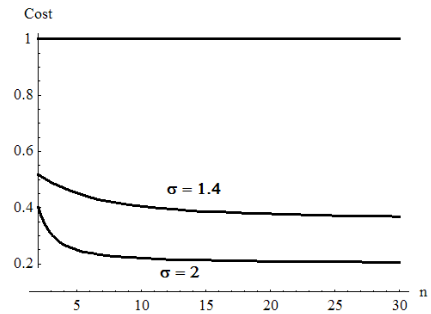
Figure 3a shows cost as a function of n for complementary inputs and $\varepsilon = 0.1$ in the case of the uniform distribution. As innovation becomes more complex, the cost of innovation increases and converges to 1 when $n \to \infty$. This means that the probability of innovation decreases with n and converges to 0. Convergence is faster when σ gets closer to zero. When the substitutability between the inputs is very low (σ close to zero), the probability of innovation is very small even for simple innovations (low n).

Figure 3b shows that the conclusions change when the research inputs are substitutes. In this case the cost of innovation decreases when the complexity of innovation increases (i.e. the probability of innovation increases with n).

3.4. **High complexity and monopolistic competition.** It is interesting to analyze the equilibrium of the economy when $n \to \infty$ for two



(A) Complements.



(B) Substitutes.

FIGURE 3. Cost of innovation as a function of n.

reasons. First, $n \to \infty$ represents innovations that are highly complex and therefore require a *large number* of inputs to be developed. The innovator faces a patent thicket and has to gather inputs from many patentees. We know how the probability of innovation changes as n increases, but it is interesting to determine in what cases it will go to 0 or $1 - F(\varepsilon)$. Second, in this limiting economy there is an infinite number of input sellers, so the effect of a price change by a single firm has a infinitesimal impact on the cost of innovation, and the market becomes monopolistically competitive.

Proposition 5 characterizes equilibria with positive probability of innovation ($p < \bar{v}$). In this case there are values of σ for which there is no equilibrium with positive probability of innovation.

Proposition 5. *A SSPE with $p < \bar{v}$ exists only when $\sigma > \hat{\sigma}$ where $\hat{\sigma} = \frac{\bar{v}}{\bar{v}-\varepsilon}$. The equilibrium price and cost of innovation are $p = \frac{\sigma}{\sigma-1}\,\varepsilon$.*

Proof. Dividing the first order condition (2) by n, we get:

$$h(p)\,(p-\varepsilon)\,\frac{1}{n} + \sigma\left(\frac{n-1}{n}\right)\left(\frac{p-\varepsilon}{p}\right) - 1 = 0.$$

As $n \to \infty$, the term with the hazard function goes to zero. This is because each firm becomes negligible and does not affect the probability of innovation on its own. It is clear that the equilibrium price of the limiting economy solves:

$$\sigma\left(\frac{p-\varepsilon}{p}\right) - 1 = 0.$$

Therefore, $p = \frac{\sigma}{\sigma-1}\,\varepsilon$, which is between ε and $\bar{v}$ only when $\sigma > \bar{v}/(\bar{v} - \varepsilon)$. ■

It is interesting to comment on three characteristics of the equilibrium. First, any $p \geq \bar{v}$ is an equilibrium for any value of σ in this limiting economy. If $p \geq \bar{v}$ the probability of innovation is zero, but if a single input seller deviates, its impact on the cost of innovation is infinitesimal, so the probability of innovation (i.e. expected profits) remains unchanged. Therefore, there are no profitable deviations when $p \geq \bar{v}$.

Second, the equilibrium quantity x_i goes to zero as $n \to \infty$. This is because the number of inputs is increasing towards infinity but the total quantity of inputs required is keeping constant, given our assumptions on the innovation technology.

Finally, it is easy to show that $1 \leq \hat{\sigma} < \sigma^*$. The first inequality follows trivially from the fact that $\hat{\sigma} = \bar{v}/(\bar{v} - \varepsilon)$. Therefore, $\hat{\sigma} = 1$ only when $\bar{v} \to \infty$ or $\varepsilon = 0$. For the second inequality, it is enough to compare the equilibrium price when $\sigma = \hat{\sigma}$ with the equilibrium price when $\sigma = \sigma^*$, since price is decreasing in σ. When $\sigma = \hat{\sigma}$, price is equal to $\bar{v}$. When $\sigma = \sigma^*$ we know that the equilibrium price solves $h(p)(p - \varepsilon) = 1$. If $p = \bar{v}$, then $h(p)\,(p - \varepsilon) \to \infty$, which is much larger than 1. For $h(p)\,(p - \varepsilon)$ to decrease and approach 1, p has to decrease. This means that equilibrium price is larger with $\hat{\sigma}$ and therefore $\hat{\sigma} < \sigma^*$.

Figure 4 shows the cost schedule as a function of σ when $\bar{v} = 1$ and $\varepsilon = 0.1$. The equilibrium of the limiting economy does not depend on the distribution of v, but it depends on the upper bound of the support of the distribution.

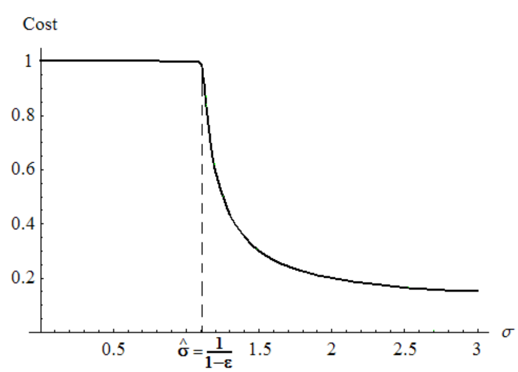


FIGURE 4. Cost of innovation in the limiting economy.

The equilibrium price is the same than Dixit and Stiglitz's (1977) monopolistic competition model. When inputs are substitutes, firms set a mark-up over marginal cost equal to $1/(\sigma - 1)$. This means that the pricing inefficiency decreases as n increases, but it does not disappear even when $n \to \infty$.

For complements, the outcome depends on whether σ is greater or less than $\hat{\sigma} = \bar{v}/(\bar{v} - \varepsilon)$. When $\sigma > \bar{v}/(\bar{v} - \varepsilon)$, firms set a mark-up just like in the substitutes case. When $\sigma \leq \bar{v}/(\bar{v} - \varepsilon)$, the only equilibria have $p \geq \bar{v}$ and so the probability of innovation is zero. In this case, as n increases the inefficiency due to monopoly pricing increases and it is at its maximum when $n \to \infty$.

3.5. **The tragedy of the anticommons revisited.** The model presented in this paper gives a formal treatment to the tragedy of the anticommons in sequential innovation. An anticommon (Heller 1998) arises when multiple owners have the right to exclude each other from using a scarce resource, causing its inefficient under-utilization. This problem is symmetric to the tragedy of the commons, where multiple owners have the right to use a scarce resource, but nobody has exclusion rights and resources tend to be overused.

Heller and Eisenberg (1998) have raised the question whether the tragedy of the anticommons may apply to innovation in biomedical research. They point out that in this sector excessive patenting of research tools might reduce the incentive to innovate, because the innovator has to face a possibly high cost of bundling all the licenses together. This is a concern that might be shared by other hi-tech sectors where innovation follows a similar cumulative process.

The model presented above predicts that anticommons may arise in sequential innovation, but only under certain circumstances. In our model, the scarce resource is the state-of-the-art technology for innovation, over which *each* patent holder of a research input has a claim. Each patent holder decides the selling price of her input. It is interesting to notice that when $\sigma \leq 1$ all the inputs are essential to perform the innovation so all the input sellers can potentially impede the innovation by setting a high price.

Proposition 4 shows that, when the inputs are market complements, the cost of producing an innovation increases as technologies become more complex. This proposition says that when the number of patented and complementary research tools to be used in R&D increases, the probability of innovation is reduced: that is, anticommons applies to sequential innovation when research inputs are market complements. This result holds not only for perfect complementarity between the inputs, but whenever the elasticity of substitution is not sufficiently large to compensate the negative effect of price changes on the probability of innovation. The opposite is true when research tools are market substitutes: in this case there is no anticommons because, as the number of research inputs increases, competitive pressure reduces their price fast enough to reduce also the cost of innovation.

Proposition 5 reinforces the previous result: when the number of patented research tools grows large, and they are highly complementary, the anticommons is so strong that the probability of innovation goes to zero.

Finally, it is important to remark that the anticommons effect arises in the absence of any kind of transaction costs. The anticommons arise as a natural consequence of the uncoordinated market power of the input producers. As we will see in section 6.2, asymmetric information is not essential for our results. All we need is a downward sloping expected demand for the inputs.

4. Patent policy

We have shown that strong patents and fragmentation of ownership lead to a low probability of innovation when inputs used in research are complementary. The government could reduce the patent thicket by granting less patents, or by reducing the breadth of the patents. However, weaker patents imply a reduced incentive to discover research inputs in the first place. The problem of the division of profit between sequential innovators has been studied by Scotchmer (1991), Green and Scotchmer (1995), Chang (1995), Scotchmer (1996) and Hopenhayn, Llobet, and Mitchell (2006). In this section we complement existing literature by analyzing the optimal division of profits between sequential innovators when the final innovation requires multiple inputs to be performed. Our objective is to determine the effects of higher complexity on the optimal patent policy.

The innovation technology is the same as the one in section 2. For tractability, we focus on the perfect complements case, so the final good can only be introduced only if all inputs are invented. Without loss of generality (for the $\sigma = 0$ case), we assume that once the inputs are invented they can be reproduced at zero marginal cost ($\varepsilon = 0$). Finally, we will concentrate on the uniform distribution case, $v \sim U[0, 1]$, which provides linear demands for the innovation.

An important difference with Section 2 is that now the n inputs must be invented at an earlier stage. We assume that there is a fixed (sunk) cost K/n of inventing the inputs. Each input will be introduced if expected revenues are larger than the fixed cost. As is standard in the literature of sequential innovation, the fixed cost of the input sellers is unknown to the policy maker. All the policymaker knows is that the sunk cost has a distribution $K \sim U[0, \bar{K}]$. Therefore, the patent policy cannot depend on the realization of K.

Our assumptions imply that the social cost and value of the innovation do not change as n increases. All that changes is the number of input producers with which the final innovator has to agree to perform her innovation.

We introduce a patent policy parameter $\phi \in [0, 1]$. This policy parameter can be interpreted as the patent breadth, the novelty requirement, or the strength with which Intellectual Property law is enforced in courts: there is a probability ϕ that the input innovator is granted a patent and that the patent can be successfully defended in court. We also assume that imitation is costless: Bertrand competition will ensure that inputs not protected by patents are sold at marginal cost. Consequently the innovator only has to pay a non-competitive license fees for inputs that are covered by patents, and can buy the rest at marginal cost.

We will first analyze the case in which patent policy affects only input innovators (first stage innovations). In this section therefore we abstract from the fact that in some cases the same patent policy should apply to the final (second stage) innovation. This is standard in the literature on sequential innovation where the focus is on the effect of patents on the optimal division of profits between first and second stage innovators. For completeness, in Section 4.2 we will analyze what happens when patents apply equally to first and second stage innovations (symmetric patent policy). We will show that the qualitative results do not change.

The timing of the game is as follows: (i) input innovators observe the sunk cost of innovation, and decide whether to invent their input or not, (ii) if all inputs are invented, Nature decides which inputs are protected by patents, (iii) patent holders set a price for their patented inputs, (iv) Nature decides the value of the innovation, (v) the final innovator decides whether to innovate or not.

Research inputs will be invented only if the expected revenue from selling the input is higher than the sunk cost of inventing it. Expected revenue depends on whether they are granted a patent and on how many other patents are granted. Remember that each input innovator is granted a patent with probability ϕ. Suppose m patents are granted in the second stage. Then, the price of the inputs, the cost of innovation and the probability of innovation are, respectively: $p_m = \frac{n}{m+1}$, $c_m = \frac{m}{m+1}$ and $Pr_m = \frac{1}{m+1}$.

Consider an input innovator who is granted a patent. In the third stage, her revenues depend on how many other patents have been granted. Let $k = m - 1$ denote the number of patents granted, in addition to the patent of the input innovator we are considering. Actual revenues (after uncertainty is resolved) are $\Pi_k = 1/(k+2)^2$. Expected revenues are:

$$E(\Pi) = \phi \sum_{k=0}^{n-1} \left(\frac{1}{k+2}\right)^2 \frac{(n-1)!}{(n-1-k)!k!} \phi^k (1-\phi)^{n-1-k} \tag{4}$$

There are two effects of increases in ϕ on $E(\Pi)$. First, a higher ϕ increases the probability of being granted a patent, which increases $E(\Pi)$. Second, the increase on ϕ increases the probability that more patents are granted in addition to mine, which decreases $E(\Pi)$ because of the anticommons effect. Our simulations indicate that the first effect always dominates the second effect for n small, so that $E(\Pi)$ are increasing in ϕ. For larger n, though, the second effect dominates the first for large ϕ, so $E(\Pi)$ first increases and then decreases with ϕ. Figure 5 shows the expected revenue of input producers as a function of patent strength for $n = 5$.

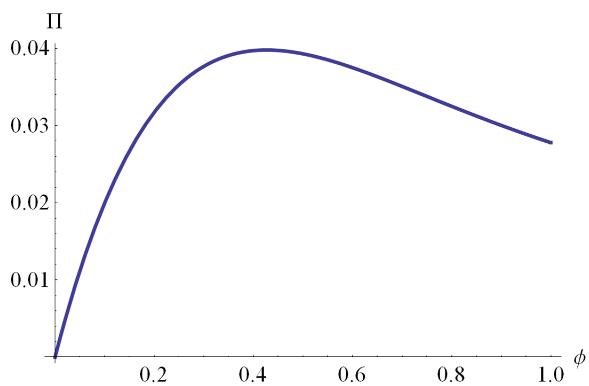


FIGURE 5. Expected revenues for inputs. n=5

Let us now focus on how patent policy affects the probability of innovation. In order for the final innovation to be introduced, two things must happen: (i) expected revenues have to be larger than the fixed cost for the input sellers, and (ii) the value of the innovation for the final innovator has to be larger than the cost of paying the inputs protected by patents. The probability that (i) happens is $Pr(E(\Pi) >$

$K/n) = n\,E(\Pi)/\bar{K}$. The expected probability that (ii) happens is:

$$E(Pr(v > c)) = \left(\sum_{m=0}^{n} \frac{1}{m+1} \frac{n!}{(n-m)!\,m!} \phi^{m} (1-\phi)^{n-m} \right).$$

It is easy to see that $E(Pr(v > c))$ is decreasing in ϕ. This is because, assuming the inputs are invented, a higher ϕ increases the probability that more patents are granted, which implies a lower probability of introducing the final innovation due to the anticommons problem.

The probability of introducing the final innovation is simply the product of the previous two probabilities: $Pr = Pr(E(\Pi) > K/n)\,E(Pr(v > c))$.

$$Pr = \frac{nE(\Pi)}{\overline{K}} \left(\sum_{k=0}^{n} \frac{1}{k+1} \frac{n!}{(n-k)!k!} \phi^{k} (1-\phi)^{n-k} \right) \tag{5}$$

Figure 6 shows that the probability increases with respect to ϕ, reaches a maximum (for the optimal policy ϕ^*), and then decreases, for $n = 5$. There are two effects pulling in opposite directions: on one hand increasing ϕ increases the probability of inventing the inputs, on the other hand it increases the cost of the final innovator.

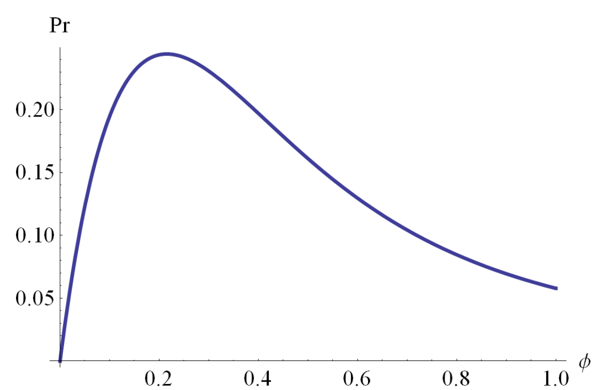


FIGURE 6. Probability of final innovation. n=5

Interestingly, the policy that maximizes Pr is smaller than the policy that maximizes $E(\Pi)$. This is because is always decreasing in ϕ. This has an implication for rent seeking activities: at the optimal policy, input innovators would push for stronger patents on research inputs, and final innovators would push for weaker patents.

The following proposition presents the most interesting result of this section.

Proposition 6. *The optimal patent strength is decreasing in the complexity of the innovation.*

Proof. Rather than going though an analytical proof, it suffices to analyze Figure 7. This figure shows the policy that maximizes the probability of innovation, as a function of n. Given that the optimal policy does not depend on $\bar{K}$ or any other parameter, this actually shows the effect of increases on n on ϕ^*, which is clearly decreasing. ∎

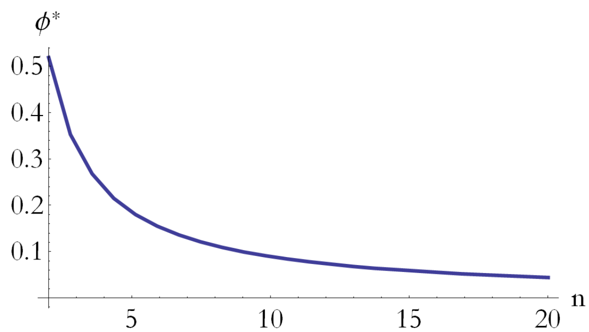


FIGURE 7. Increasing complexity and optimal patent strength

In models of sequential innovation, the degree of patent protection determines the division of profits between sequential innovators. Stronger patents increase the protection for early innovators, as they grant them a claim over following innovations. When innovation builds on several prior inventions, the uncoordinated market power of earlier patentees generates an anticommons effect. Proposition 6 shows that as the number of claims increase, the anticommons gets worse, and the optimal response is to reduce the degree of patent protection.

It is important to remark that in our model a minimum amount of protection is always needed, otherwise input producers would have no incentives to invent the inputs in the first place. The important result, however, is that the degree of patent protection should decrease as n increases. In section 4.2 we will show that the same result holds in a more sophisticated model where patent policy also affect the revenues of the final innovator.

4.1. **Patent policy with imperfect substitutability.** In this section we relax any constraint on the substitutability between research inputs, and allow it to vary between zero and infinity. Any degree of substitutability higher than one requires a positive marginal cost ε for the inputs, otherwise the innovator would only use unpatented inputs and would be able to innovate at zero cost. As before, inputs require a fixed cost of K/n to be invented, with $K \sim U[0, \bar{K}]$, and $v \sim U[0, 1]$.

This setting represents a generalization of the analysis of Section 4. However, the problem becomes analytically untractable, and we have to resort to numerical simulations.

The results in this section are in line with those of Section 4: increasing complexity reduces the optimal patent strength. Figure 8 shows the optimal patent breadth ϕ^* as a function of the elasticity of substitution σ, with $\varepsilon = 0.1$ and $\bar{K} = 0.4$, for $n = 5$ and $n = 10$. Increasing n reduces ϕ^* for any σ, and the result holds for any $n \geq 2$.

Another interesting result is that the optimal patent strength ϕ^* is increasing in σ. This is because substitutability increases competition among input producers, thus reducing their ability to set a price above the marginal cost. This is equivalent to redistribute revenues from input producers to the final innovator. In order to compensate for this redistribution, and return to its optimal level, patents must be strengthened. Notice however that, for this same reason, patents in this case are less harmful as they provide a more limited market power. Finally, it is interesting to note that there is some value of σ, larger than 1, for which the optimal patent policy is 1, i.e. inputs have to be protected with strong patents when the substitutability is very large.

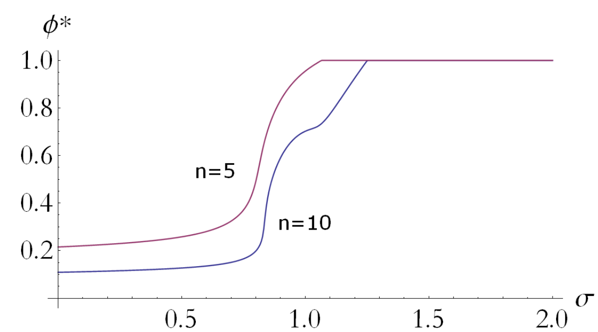


FIGURE 8. Increasing complexity and optimal patent strength for imperfect substitutes ($\varepsilon = 0.1, \overline{K} = 0.4$)

4.2. **Symmetric patent policy.** Here we complete the analysis of section 4 by assuming that the patent policy also applies to the final good innovator: there is a probability $\phi \in [0, 1]$ that, if the final innovation is introduced, it will be protected by a patent. The expected revenue of the final innovator is $\phi\, v$, where $v \sim U[0, 1]$ is the gross social value of the innovation, just as before. This change introduces an additional factor in favor of strong patents: the expected private value of the innovation ϕv, to be shared between the final innovator and the input producers, it is now increasing in patent strength.

The rest of the model is the same as in Section 4, so we focus on the perfect complements case with linear demand for the final innovation. Research inputs need to be invented in a previous stage, with a fixed cost $K \sim U[0, \overline{K}]$, and the necessary condition for this to happen is that expected profits from selling the input $E(\Pi_i) = (1 - F(c/\phi))\, x_i\, p_i$ are higher than this sunk cost of innovation. The timing of the game is also left unchanged.

Interestingly enough, introducing this extension does not change our results. Input producers completely internalize the effect of ϕ on the expected profits of the final innovator, and leave the probability of innovation unaffected. When m inputs receive patent protection, inputs price and cost of innovation become respectively $p = \frac{\phi n}{m+1}$ and $c = \frac{\phi m}{m+1}$. As has been said, the probability of innovation is the same: $Pr(\phi v > c) = \frac{1}{m+1}$.

On the other hand, the revenue of a patent holder when $k = m - 1$ additional patents are granted is now ϕ times lower at $\frac{\phi}{(1+k)^2}$. This obviously has a negative effect on the overall probability of getting the final innovation $Pr = Pr(E(\Pi) > K/n)\, E(Pr(v > c))$. It also has a positive effect on the *level* of the optimal patent strength (the one that maximizes Pr). Still it does not affect the main result: optimal patent strength is decreasing in the complexity of the innovation, as shown if Figure 9

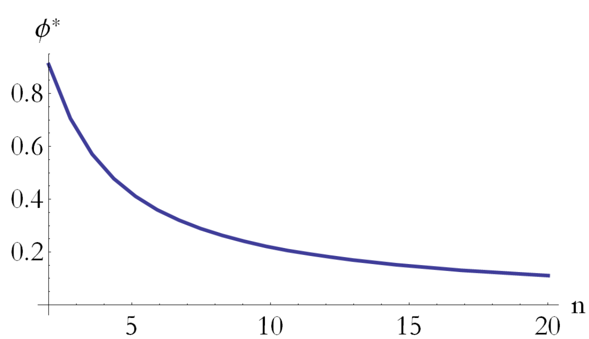


FIGURE 9. Increasing complexity and optimal patent strength: the case of symmetric patent policy

This result reinforces the one in section 4, and gives an idea how strong the anticommons effect is as technological complexity increases. The result is even more remarkable when taking into account that, in this model, the final innovator is a perfect price discriminator. This implies that reducing patent strength reduces the expected producer revenue (ϕv) on a one-to-one basis, which we see as a realistic upper bound on the impact of patents on the incentive to innovate.

5. Patent pools

Even though our objective is to analyze the optimal patent policy, it is interesting to analyze what would happen if research inputs were priced cooperatively, either by a collective institution such as a patent pool or by a single patent holder (monopolist) that owns all the patents. This analysis is interesting because the USPTO (US Patents and Trademarks Office) itself has recommended the creation of patent pools to ease the access to biotechnology research tools (Clark, Piccolo, Stanton, and Tyson 2000).

For this analysis, we come back to the basic model of Section 2 (the innovator requires n inputs, which are already invented and protected by patents of infinite breadth). Proposition 7 shows the equilibrium price in this case. The difference with the previous case is that now the patent holder maximizes joint-profits and therefore takes into account the cross-price effects between expected demands.

Proposition 7 (Patent Pool)**.** *The equilibrium price when all the inputs are priced cooperatively, p^*, is the argument that solves $h(p)(p - \varepsilon) = 1$.*

Proof. Given the symmetric input demands, the pool wants to sell a symmetric bundle. Therefore $x_i = 1/n$ and $p_i = p$ for all i and the pool wants to maximize total profits $n\,(1 - F(p))\,(p - \varepsilon)$. The first order condition is $n(-f(p)\,(p - \varepsilon) + 1 - F(p)) = 0$. Rearranging terms we get the desired result. ∎

Notice that p^* depends only on the functional form of h and the value of ε, but not on the values of σ or n. The following proposition compares the cost of innovation when the inputs are priced individually, c, with that of a patent pool, p^*.

Proposition 8. *The cost of innovation when the inputs are priced non-cooperatively, c, is equal to that of a patent pool, p^*, when the cross-price derivative is zero ($\sigma = \sigma^*$), it is larger when the inputs are complements ($\sigma < \sigma^*$) and it is smaller when the inputs are substitutes ($\sigma > \sigma^*$).*

Proof. We know from the proof of Lemma 2 that when $\sigma = \sigma^*$, the cross-price derivative is zero and $\sigma = p\,h(p)$. Replacing this in (2) and rearranging we get $h(p)(p - \varepsilon) = 1$, which is the cooperative result.

Given that p is decreasing in σ, whereas p^* is independent of σ, $p > p^*$ when $\sigma < \sigma^*$ and $p < p^*$ when $\sigma > \sigma^*$. ■

The difference between cooperative and non-cooperative pricing is that in the first case the firms take into account the effect of an increase in the price of one input on the demand for the rest. When $\sigma = \sigma^*$ this effect is zero so the price of the pool coincides with that of the non-cooperative equilibrium. When $\sigma < \sigma^*$ the effect is negative, so the pool knows that an increase in price will decrease the demand for the rest and will set a price smaller than the uncoordinated input sellers. The opposite happens when $\sigma > \sigma^*$.

In the case of the uniform distribution, the pool price is $p^* = (1 + \varepsilon)/2$. Figure 10 compares this price with the non-cooperative price for $\varepsilon = 0.1$ and $n = 5$.

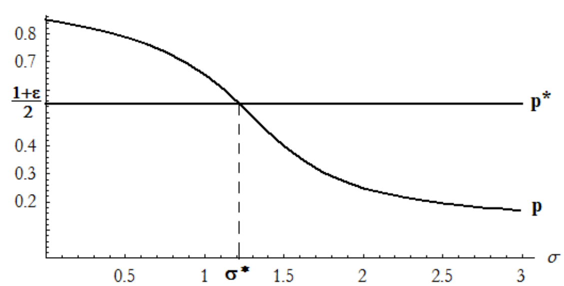


FIGURE 10. Cooperative and non-cooperative pricing.

Our results are similar to those found in Lerner and Tirole (2004). As we discussed in Section 2.4, the difference between the two papers is that we have assumed that inputs are used in research in a continuous fashion, while Lerner and Tirole (2004) consider discrete input choices (1 if the input is used and 0 otherwise).

Under the latter approach, the equilibrium will depend on whether the competition margin or demand margin bind. When the competition margin binds, if the input seller raises her price, her input would be evicted from the bundle of patents bought by the final innovator. When the demand margin binds, the input seller can raise price without excluding its input from the basket, but the overall demand for the bundle would decrease.

These competition and demand margins are related to our substitution and innovation effects. The substitution effect says that an increase in the price of one of the inputs lowers the demand for that input and increases the demand for the rest (holding overall demand constant). The innovation effect says that an increase in the price of an input lowers the overall demand for the basket of inputs (holding the relative demand of the inputs constant).

Therefore, our model can be interpreted as a "smooth" version of Lerner and Tirole's. In Lerner and Tirole, as inputs become more complementary, it is more likely that the demand margin will bind. In our model, both margins always bind (except when $\sigma = 0$ or $\sigma \to \infty$), but as inputs become more complementary, the innovation effect becomes more important relative to the substitution effect.

In this sense, the contribution of our model is twofold: (i) we show that Lerner and Tirole's results extend to the continuous innovation technology case, and (ii) we explain the effects of patent pools on innovation using the traditional definition of complementarity based on cross-price derivatives (precisely because the effects of price changes are smooth), while Lerner and Tirole base their definition of complementarity on the shape of the revenue function of the innovator.

6. Extensions

In this section we analyze the consequences of relaxing some of the basic assumptions of the model.

6.1. **Social value and cost depend on complexity.** Until now, we have assumed that the distribution of values of the innovation and the social cost of the inputs do not depend on n, and that there are no returns from specialization. Under these assumptions, a change in n only changes the number of producers from whom the innovator has to buy the research inputs in order to innovate, but does not change the probability that the innovation is socially valuable.

However, it could be argued that the revenues of the innovator or the cost of the inputs are increasing or decreasing in n, or that a higher number of inputs has a positive impact in the R&D technology due to a higher division of labor. All these changes have equivalent effects on the probability of innovation so we will concentrate on changes in the distribution of returns of the innovation.

Let the return of the innovation be $a(n)\,v$, with $a'(n) \geq 0$ or $a'(n) \leq 0$ and $\lim_{n\to\infty} a(n) = a_\infty > 0$. v has a cumulative distribution $F(v)$ as before. Notice that we are not setting an upper bound on a_∞. All we require is that if a is non-increasing it does not go to zero as $n \to \infty$. This is because if $a_\infty = 0$ then the distribution of values of the innovation will collapse to zero and the innovation will never be profitable when n is very large by assumption.

The probability of innovation is $1 - F(c/a)$, and in the symmetric equilibrium $c = p$ and $x = 1/n$. The equilibrium price of the inputs (i.e. the cost of innovation) solves:

$$(p - \varepsilon)\, h(p/a)/a = n - \sigma\,(n-1)\,(p-\varepsilon)/p$$

but we are more interested in the ratio $k = c/a$, which enters in the probability of innovation. Replacing in the previous equation we have:

$$(k - \varepsilon/a)\, h(k) = n - \sigma\,(n-1)\,(k-\varepsilon/a)/k. \tag{6}$$

This equilibrium is equivalent to the one in Proposition 1, thinking of $k = c/a$ as the cost of innovation and ε/a as the social cost of the inputs. We can prove the same theorems as before with respect to the difference between complements and substitutes, the welfare effects of patent pools and $\partial k/\partial \sigma$. However, $\partial k/\partial n$ will be different because now ε/a is a function of n.

Using the implicit function theorem on the equilibrium relation (6) we get:

$$\frac{\partial k}{\partial n} = \frac{\frac{h(k)\,(k-\varepsilon/a)-1}{n-1} - \frac{n\,a'}{a}\,\frac{\varepsilon/a}{k-\varepsilon/a}}{h'(k)\,(k-\varepsilon/a) + h(k) + \sigma(n-1)(\varepsilon/a)/k^2}$$

As before, the sign of this derivative depends only on the sign of the numerator, but now there is an additional term which shifts the threshold value of σ that divides positive and negative changes in k. This threshold value will be to the left of σ^* when $a'(n) > 0$ and to the right of σ^* when $a'(n) < 0$.

Two important remarks are in order. First, if $a'(n)$ is large then the last term in the numerator will determine the sign of the derivative. In this case, the effect of changes in n on revenues completely overcomes the effect on the pricing of inputs, and $\partial k/\partial n$ has the opposite sign of $a'(n)$ irrespective of the value of σ. Second, even for small $a'(n)$, when $a'(n) > 0$ and $\sigma \to \infty$ the derivative is always positive. Therefore when $a'(n)$ is small and positive, there are two regions where $\partial k/\partial n$ is positive: one with low values of σ and another with large values of σ.

According to the previous analysis, assuming that the return of the innovation depends on n has an effect on the derivative of the probability of innovation with respect to n. Next, we will show that this assumption has no significant effect on the analysis of the equilibrium as $n \to \infty$.

The equilibrium price solves:

$$h(p)\,(p-\varepsilon)\,\frac{1}{a\,n} + \sigma\left(\frac{n-1}{n}\right)\left(\frac{p-\varepsilon}{p}\right) - 1 = 0.$$

When $n \to \infty$, the first term will go to zero because $a_\infty > 0$. Therefore, the equilibrium price is the same as before, $p = \frac{\sigma}{\sigma-1}\varepsilon$, which is less than the maximum possible revenue $(a_\infty \bar{v})$ only if $\sigma > a_\infty\bar{v}/(a_\infty\bar{v} - \varepsilon)$. When $\sigma \leq a_\infty\bar{v}/(a_\infty\bar{v} - \varepsilon)$, on the other hand, there is no equilibrium price such that the probability of innovation is positive.

The probability of innovation is $1 - F(p/a_\infty)$. There are two possible cases. If $a_\infty < \infty$ then the probability of innovation is less than optimal, just as in the basic model. If $a_\infty = \infty$ then the probability of innovation will go to 1 for $\sigma > 1$ and 0 for $\sigma \leq 1$, which is the same as assuming $\varepsilon = 0$ in the basic model.

6.2. **No asymmetric information.** Another assumption of the basic model is that there is asymmetric information on the value of the innovation (read Gallini and Wright 1990, Bessen 2004, for good discussions of why this assumption makes sense). However, our results do not depend on the existence of asymmetric information. All that is needed for the results is a downward sloping demand for innovations.

An alternative interpretation could be that there is a continuum of innovators with decreasing returns to their innovations. Suppose that the innovators are indexed by the return to their innovations, which ranges between $\underline{v}$ and $\bar{v}$. Now, $F(v)$ is the measure of innovations with a return less or equal to v. Also, assume that the innovations do not compete against each other in the final goods market and that the input sellers cannot price discriminate between the innovators. It is easy to see that all the previous results translate directly into this setting. All that changes is that now $1 - F(c)$ is not the probability of innovation but the measure of innovations performed.

6.3. **No price discrimination.** We can also relax the assumption that the innovator is a perfect price discriminator. Dropping this assumption introduces a wedge between the social and private values of the innovation. This means that the distribution of values of innovation changes, and that now there is also an inefficiency in the final goods sector. Assume that the social value of the innovation is still distributed according to $F(v)$, with probability density function $f(v)$. The private value of the innovation is now v_p, which is less than the social value of the innovation. With a linear demand for the final good, for example, the private return of the innovation would be $v_p = v/2$, which has a probability density function given by $2\,f(2v_p)$. The qualitative results are the same as before. All that changes is that now the probability of innovation decreases for each value of σ, and so the values of σ^*, $\hat{\sigma}$ and

$\bar{\sigma}$ increase. Also, the optimal patent protection is lower for each value of σ and n than in the case of perfect price discrimination.

6.4. **Uncertain return of the innovation.** We have also assumed that the innovator is the only one that knows the value of the innovation. In this section we ask what happens if v is also unknown to the innovator. Formally, we do this by changing the timing of the game: (i) the input producers simultaneously set the price of their inputs, (ii) given prices, the innovator calculates the input mix that minimizes the cost of innovation and decides whether to innovate or not, and (iii) Nature extracts a value v for the innovation from the distribution $F(v)$.

We begin by solving the second stage of the game. The innovator decides what would be the optimal combination of inputs to perform the innovation in case he decides to perform it. This leads to the same cost of innovation and conditional demands as before. Then, the innovator decides whether to perform the innovation or not, in order to maximize expected profits $E(v) - c$. The innovation will be performed if $E(v) \geq c$ and will not be performed otherwise. If $E(v) < \varepsilon$, then the innovation will never be performed, so we assume that $E(v) \geq \varepsilon$. We also assume that the innovator will perform the innovation if $E(v) = c$.

The uncertainty has now passed from the input sellers to the input producer. The problem of the input sellers is deterministic, they know $E(v)$ and they know that if the price is higher than $E(v)$ the innovation will not be performed (i.e. their demands will be zero). Now, the inputs are always market substitutes unless $\sigma = 0$. It is easy to show that the innovation will always be performed, and that the elasticity of substitution only affects the distribution of payoffs between the input sellers and the innovator.

Lemma 3 shows that input demands are discontinuous at a certain price, and Proposition 9 proves that in the symmetric equilibrium $c \leq E(v)$ so the innovation is always performed.

Lemma 3. *Input demands are discontinuous at*

$$p_i = \left(nE(v)^{1-\sigma} - \sum_{j \neq i} p_j{}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} .$$

Proof. The demand for inputs is positive if the cost of innovation is not larger than the expected value of the innovation, that is:

$$n^{-\frac{1}{1-\sigma}} \left(\sum_{i=1}^{n} p_i{}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \leq E(v).$$

Rearranging terms, we get the condition on the price of the input:

$$p_i \leq \left(nE(v)^{1-\sigma} - \sum_{j \neq i} p_j{}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

If p_i is larger than this value, then the innovation is not performed and the demand for all inputs is zero. ■

The input sellers want to maximize profits $x_i\,(p_i - \varepsilon)$. Proposition 9 states the solution of the game.

Proposition 9. *The equilibrium price when the return of the innovation is uncertain for the innovator is:*

$$p = \begin{cases} \frac{\sigma(n-1)}{\sigma(n-1)-n}\varepsilon & \text{if } \sigma > \frac{n}{n-1}\frac{E(v)}{E(v)-\varepsilon}, \\ E(v) & \textit{otherwise.} \end{cases}$$

Proof. After imposing symmetry, the derivative of $x_i\,(p_i - \varepsilon)$ with respect to p_i becomes:

$$D(p) = \frac{1}{n}\left(-\frac{\sigma(n-1)}{n}\frac{p-\varepsilon}{p} + 1 \right).$$

Lemma 3 implies that if the derivative with respect to price is positive at $p = E(v)$, this is a symmetric equilibrium, as firms are making positive profit, do not want to lower price ($D \geq 0$), and would have a zero profit if they would rise price. This happens when $\sigma \leq \frac{n}{n-1}\frac{E(v)}{E(v)-\varepsilon}$.

When $\sigma > \frac{n}{n-1}\frac{E(v)}{E(v)-\varepsilon}$, on the other hand, the equilibrium price solves the unrestricted first order condition $D(p) = 0$. ■

7. Conclusion

In this paper we extend the literature of sequential innovation in two directions. First, we study how the probability that an innovation is privately profitable changes as technologies become more complex, and the inputs used in research are patented. We find that the results depend on the substitutability between these research inputs.

When the inputs are complements, the profitability of the innovation is decreasing in the technological complexity. In the limit (when $n \to \infty$), when the degree of substitutability is below a threshold level, which is higher than 1, the innovation is never profitable. This paper therefore gives a formal treatment of the tragedy of the anticommons.

On the other hand, when the inputs are substitutes, the profitability of the innovation is increasing in technological complexity. Even in this case, when $n \to \infty$, the cost of gathering all the inputs for the innovation is always too high from a social point of view and thus the probability of innovation is suboptimal.

Second, we study the optimal response of patent policy to increasing complexity of innovation. We find that, because of the anticommons effect, the optimal degree of patent protection is decreasing in the complexity of the innovation. The degree of patent protection determines the division of profits between sequential innovators. Stronger patents distribute more revenues from the last innovator to the previous innovators. This result says that increasing complexity of innovation reduces the optimal amount of protection granted to earlier innovators. This result is very robust: numerical simulations suggest that it holds for any degree of substitutability between previous innovations, and also holds when patent protection affects the revenues of the final innovator (i.e patents also affect the amount of profits to be distributed between successive innovators).

These results are at odds with respect to what we observe in the real world: the complexity of technology is increasing but patents are becoming stronger. Not only they have been recently extended to sectors previously lacking protection (sexually reproduced plants, software, business methods, products and processes of biotechnology, including plants and animals). Also patent length has been increasing over the years, and patent systems are being created in countries where they did not previously exist. We think this is a contradiction worth being studied further.

Finally, we also study what happens when inputs are priced cooperatively, either by a collective organization as a patent pool or by a single owner of all the inputs. We find that the cost of the innovation decreases with respect to the non-cooperative pricing, when inputs are market complements, while it increases when inputs are market substitutes. This result is in line with the intuition of Shapiro (2001) and the model of Lerner and Tirole (2004). Still, we believe it complements these previous papers by using a standard definition of complementarity/substitutability, allowing us to exactly identify the economic forces driving the result.

References

Bergstrom, T. C. (1978): "Cournot Equilibrium in Factor Markets," Discussion paper, University of Michigan working paper.

Bessen, J. (2004): "Holdup and Licensing of Cumulative Innovations with Private Information," *Economics Letters*, 82(3), 321–26.

Boldrin, M., and D. K. Levine (2002): "The Case against Intellectual Property," *American Economic Review (Papers and Proceedings)*, 92(2), 209–212.

——— (2005a): "The economics of ideas and intellectual property," *Proceedings of the National Academy of Sciences*, 102(4), 1252–1256.

——— (2005b): "Intellectual Property and the Efficient Allocation of Surplus from Creation," *Review of Economic Research on Copyright Issues*, 2(1), 45–66.

Chang, H. F. (1995): "Patent Scope, Antitrust Policy, and Cumulative Innovation," *RAND Journal of Economics*, 26(1), 34–57.

Clark, J., J. Piccolo, B. Stanton, and K. Tyson (2000): "Patent pools: a solution to the problem of access in biotechnology patents?," Discussion paper, United States Patent and Trademark Office.

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).

Dixit, A. K., and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297–308.

Gallini, N. T., and B. D. Wright (1990): "Technology Transfer under Asymmetric Information," *RAND Journal of Economics*, 21(1), 147–60.

Garfinkel, S., R. Stallman, and M. Kapor (1991): "Why patents are bad for software.," *Issues in Science and Technology*, 8(1), 50–55.

Graff, G. D., S. E. Cullen, K. J. Bradford, D. Zilberman, and A. B. Bennett (2003): "The public-private structure of intellectual property ownership in agricultural biotechnology," *Nature Biotechnology*, 21(9), 989–995.

Green, J. R., and S. Scotchmer (1995): "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*, 26(1), 20–33.

Heller, M. A. (1998): "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets," *Harvard Law Review*, 111(3), 621–688.

Heller, M. A., and R. S. Eisenberg (1998): "Can Patents Deter Innovation? The Anticommons in Biomedical Research," *Science*, 280(5364), 698–701.

Hopenhayn, H., G. Llobet, and M. Mitchell (2006): "Rewarding Sequential Innovators: Prizes, Patents, and Buyouts," *Journal of Political Economy*, 114(6), 1041–1068.

Lerner, J., and J. Tirole (2004): "Efficient Patent Pools," *American Economic Review*, 94(3), 691–711.

Romer, P. M. (1987): "Growth Based on Increasing Returns Due to Specialization," *American Economic Review*, 77(2), 56–62.

——— (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), 71–102.

Scotchmer, S. (1991): "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives*, 5(1), 29–41.

——— (1996): "Protecting Early Innovators: Should Second-Generation Products Be Patentable?," *RAND Journal of Economics*, 27(2), 322–331.

Shapiro, C. (2001): "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting," in *Innovation policy and the economy*, ed. by A. B. Jaffe, J. Lerner, and S. Stern, vol. 1, pp. 119–150. MIT Press for the NBER.

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