UNIVERSIDAD DE SAN ANDRES

Seminario del Departamento de Economía

"Patent Design under the Threat of Litigation"

Gerard Llobet

(University of Rochester)

Martes 15 de Septiembre
11.00 hs.
Aula Chica de Planta Baja

Patent Design under the Threat of Litigation

Gerard Llobet*
Department of Economics
University of Rochester
Rochester, NY 14627

September 1998 - Preliminary and Incomplete

Abstract

This paper introduces a new approach to study patent enforcement under private information. By explicitly modeling the litigation technology, a more realistic concept of patent protection is obtained. This protection decreases on the quality of the competitor's future innovations. It is precisely the existence of competition what leads a firm to disclose the information learned about its own invention. As a result, a revelation mechanism can be implemented to increase social welfare. Interestingly enough, this mechanism resembles the interpretation that courts make of patent legislation.

1 Introduction

Patents and their effect on social welfare have been a concern in economics for a long time. Their use stimulates innovation by granting monopoly power to patentholders, yet this same power creates potential distortions that decrease social welfare.

This paper analyzes what are the effects of litigation - or the threat of using it - as a way to enforce patents. We claim that this is an integral part of the patent system, with relevant consequences in the design of better institutions.

Litigation is also an important source of welfare loss and recent changes in legislation have affected considerably its characteristics.

We also propose a positive way to define the breadth of patents. Rather than thinking about which should be the implications of the breadth of a patent, we focus on the way in which it is actually enforced. The existence of courts and their rulings can be interpreted as a way to

*E-mail: glbt@troi.cc.rochester.edu. I thank Hugo Hopenhayn for his help and advice. I also benefited from conversations with Susanna Esteban and comments by Matt Mitchell and Andrzej Skrzypacz. As usual, all errors are mine.

aggregate the relevant features of a patent. Therefore, if a court decides that an innovation infringes an existing patent, this means that the invention was in fact inside the breadth covered by it.

A patent is a contract with many relevant dimensions that the literature has tried to approach in several ways[1]. There is some consensus, though, about most of its important features: its *length* - or period of time for which the patent is valid -, the *fees* paid, and its *scope*, measured as *vertical scope* - or protection against successive improvements - and *horizontal scope* - or protection against substitute goods.[2] While the length of the patent and the fees paid are elements quite easy to identify and regulate, the scope has proved to be more problematic. There is no general formulation, and the concepts used are difficult to translate into real institutions. These shortcomings limit the recommendations that we can usually make in the design of better ways to reward innovators.

Most of the previous models, with a few exceptions such as Waterson (1990), consider the scope of a patent as being clearly distinguishable, making infringement evident. This setup makes the existence of litigation irrelevant. In reality, boundaries are rather blurry, being this the reason why courts are needed to assess whether inventions infringe or not an existing patent.

Due to this legal system, when a firm decides to produce an invention in an area where some patents are already in place, it has some idea about how likely it is to infringe them, and in this case, to succeed in court. It will also consider the incentives that the patentholder has to settle and litigate the infringement. So, it will decide to enter if the total expected profits, net of litigation costs are positive.

Firms hold private information about their innovations. Legal procedures are also used as a mechanism to reveal this information. For this reason, when firms behave optimally, there is some litigation, even though this is an inefficient way to solve disagreements. If this process were free we would like to use it to disclose all information and decide in a case by case basis. Being this an expensive process[3], it deters entry of some competitors that would produce otherwise, while in other cases it also gives incentives to reach a settlement between the patentholder and the potential infringer.

In order to see if this model is a good representation of the legal system, we study how it matches some of the characteristics of patent litigation highlighted by recent empirical papers. Lanjouw and Schankerman (1997) construct a database that allows them to obtain some stylized facts. One of their most surprising findings is that broader patents are litigated less often than narrower ones. We claim that one of the reasons why this might be the case is that broader patents are in fact deterrent for potential entrants, and so this decreases the amount of infringement and litigation we see.

We also design mechanisms that can improve upon the current patent system even in the presence of asymmetric information. This feature is not new, and some other recent papers,

[1] Gilbert and Shapiro (1990) and Klemperer (1990) are classical examples.

[2] The terminology is not standard, and other terms, such as breadth and width are used. We will use vertical scope and breadth indistinctly.

[3] Lerner (1994) estimates that in the biotechnology sector around 6% of the patents are litigated, with a cost of up to 25% of the total expenditure in basic research.

such as Cornelli and Schankerman (1996) and Scotchmer (1997), have studied the role of these asymmetries. While there they focus on how fees and patent length can be used to screen different inventions, here we use the combination of breadth and fees as a revelation mechanism.

In fact, it is precisely the existence of litigation what allows us to screen among inventions. So, the optimal mechanism must take into account not only the distortion that broader patents create but also the corresponding litigation costs. Our results show that under most circumstances, the patent system should encourage settlement by means of licensing and grant more protection to patents for which this is not possible. This is done by conceding larger breadth to better inventions. Remarkably enough, this result coincides with the interpretation that courts make of patent legislation. Therefore, we propose extending this mechanism, by offering ex-ante different patent breadths at different prices. The mechanism obtained decreases the distortion that patents create by weakening the monopoly power that patentholders have and minimizing litigation costs.

When screening is not possible, we give conditions under which broad or narrow patents are optimal, depending on whether the existence of new firms decreases or increases competition. We also relate this results with papers such as Gallini (1992), where litigation is not explicitly considered.

The model is presented in section 2. Section 3 explores our concept of patent breadth and its implications. In section 4 and 5, the optimal mechanism in this framework is characterized, and section 6 concludes.

2 The Model

Consider a market in which all firms are a priori identical. They obtain ideas of quality θ, according to an exponential distribution with parameter γ.[4] This distribution function depends on the level achieved by the previous ideas, in the following way:

$$\theta \sim \Phi(\theta/\theta \geq \alpha\underline{\theta})$$

The parameter $\underline{\theta}$ is the current state of knowledge and $\alpha \in [0,1]$ represents the importance of *spillovers* created by previous inventions. Clearly, for $\alpha = 0$ the distribution of ideas will be independent of the current state of knowledge, and for $\alpha = 1$ they will build completely on previous notions. This represents a generalization of models, such as Green and Scotchmer(1995) where it is implicitly assumed that $\alpha = 1$. Because the exponential distribution is *memoryless* it is easy to see that:

$$\Phi(\theta/\theta \geq \alpha\underline{\theta}) = 1 - \exp\left[-\gamma(\theta - \alpha\underline{\theta})\right] = \Phi(\theta - \alpha\underline{\theta}).$$

The first firm to obtain an idea will be denoted as the *patentholder* (or *patentee*), p, while we will call i the *potential infringer*, that is, a firm with an alternative idea. We normalize the

[4]Estimates of patent values in Germany reported by Lanjouw (1993) seem to conform quite reasonably with this kind of distribution.

initial state of knowledge $\underline{\theta} = 0$. Hence, the patentholder obtains ideas θ_p from a distribution $\Phi(\theta_p)$ while the infringer builds on this idea, and therefore, it draws from $\Phi(\theta_i - \alpha\theta_p)$.

A firm can invest in order to turn this idea into an *invention.* Hence, obtaining an idea is free, while the subsequent invention requires an expense. Call $\pi_p(\sigma_p, 0)$ the profits that the patentee will obtain without infringement, where σ_p is the level - or size - of invention achieved. A second firm will come up, with probability $\lambda \in [0, 1]$, with an idea that generates an invention of quality σ_i. The quality of the idea will be drawn from the distribution function $\Phi(\theta_i/\theta_i \geq \alpha\theta_p)$. The value of the invention is private information. If this firm decides to produce it, the profits for the patentee and the infringer will be $\pi_p(\sigma_p, \sigma_i)$ and $\pi_i(\sigma_p, \sigma_i)$, respectively. We will denote $\pi(\sigma_p, \sigma_i)$ the sum of both.

We assume that the size of the invention that a firm achieves is increasing in the quality of the idea. Call the size of the invention $\sigma_s(\theta_s)$ where $\sigma_s(0) = 0$ and $\sigma_s'(\theta_s) > 0$ for $s = \{p, i\}$.[5] Therefore we can refer from now on to the parameter θ instead of σ as the invention, as a way to simplify notation.

The patent system is modelled in the following way: The patent is granted to the first inventor. After that, if the potential infringer produces an innovation, the patentholder can decide whether to litigate it or try to reach a settlement. The result of the litigation will depend on the quality of the first invention θ_p, the quality of the infringing one, θ_i, and the propensity of the courts to settle in favor of the patentholder, b.[6]

In order to study this problem we will analyze two aspects: the litigation process and the decision of the patentholder of whether to go to court or settle with the infringer. We start with the first.

2.1 The Litigation Process

When the second inventor obtains the idea, she knows the quality of the patented invention. Afterwards, she has to decide whether to produce the innovation, and take the risk of infringement or not. Once alleged infringement occurs, it is in the patentee's hands to decide whether to prosecute or settle the case.[7]

The Patent Office does not know the quality of the invention, θ_p, when the patent is filed. However, this becomes public before the infringer decides to produce the alternative invention.[8]

[5] This is somehow restrictive, since in many cases it might be reasonable to think that firms can choose what size of innovation $\sigma \in [0, \sigma_s(\theta_s)]$ to implement.

[6] This parameter is related to the concept of *patent breadth* common in the literature. In section 3 we explore some of its implications.

[7] This setup is in some respects similar to Nalebuff (1987). This differs in the choice that the inventor has in this model to decide beforehand whether to infringe or not. Some authors have suggested that the existence and scope of patents tends to distort incentives of firms. In particular, potential infringers take into account not only the profits from their invention but also the expected behavior of the incumbent. See Waterson(1990), for an example.

[8] It could be the case that the first inventor tried to deter entry of competitors by claiming an outstanding level of invention, and therefore almost sure infringement. Meurer (1989) works along these lines. He claims that even though patents imply disclosure of all the relevant information about the invention, some features or the previous related art can remain private information.

The game proceeds as follows: The infringer obtains an idea with probability $\lambda \in [0, 1]$ of size θ_i, which is private information. This innovator can decide whether to implement it or not. In case she does, a cost c will be incurred, independent of θ_i.

If the alternative innovation has been implemented, the patentee can make a take-it-or-leave-it offer of an amount x in order to settle the case. The infringer can accept or reject it. If she accepts, the payoffs will be $\pi(\theta_p, \theta_i) - x$ for the patentee and $x - c$ for the infringer. Otherwise, they will go to court. The litigation process implies constant legal costs for both parts of L_p and L_i respectively, and after that, the expected profit will be $\pi_p^l(\theta_p, \theta_i)$ for the patentee and $\pi_i^l(\theta_p, \theta_i)$ for the infringer. The infringer will succeed in court with a probability $q(\theta_p, \theta_i, b)$, keeping all her profits of the production, $\pi_i(\theta_p, \theta_i)$, and she will obtain nothing otherwise. So, expected profits will take the following form:

$$\begin{aligned} \pi_p^l(\theta_p, \theta_i) &= \pi_p(\theta_p, \theta_i) + (1 - q(\theta_p, \theta_i, b))\pi_i(\theta_p, \theta_i) - L_p \\ \pi_i^l(\theta_p, \theta_i) &= q(\theta_p, \theta_i, b)\pi_i(\theta_p, \theta_i) - c - L_i \end{aligned}$$

That is, the infringer in case of litigation will keep her total profits with probability $q(\theta_p, \theta_i, b)$ and will incur in a sure cost of $c + L_i$, as a result of the cost of inventing and going to court.[9] The patentee will keep the rest of the profits from production net of litigation costs. Therefore, the expected revenue for the infringer of going to court - that we call B_i - will be,

$$B_i(\theta_p, \theta_i, b) = q(\theta_p, \theta_i, b)\pi_i(\theta_p, \theta_i) - L_i \tag{1}$$

Notice that x can be interpreted as how much profit does the patentee allows the infringer to make. Also, this can be considered as a *license.* That is, we defined $\pi(\theta_p, \theta_i) = \pi_p(\theta_p, \theta_i) + \pi_i(\theta_p, \theta_i)$, and so, the profits of a license l will be $\pi_p(\theta_p, \theta_i) + l$ for the patentholder and $\pi_i(\theta_p, \theta_i) - l - c$ for the licensee. Therefore, $l = \pi_i(\theta_p, \theta_i) - x$.[10]

Figure 1 represents the structure and timing of the game.

In order to solve this game, we restrict ourselves to credible litigation threats. Conceivably the patentee could threat to go to court every time infringement occurred in order to deter potential entrants. However, once infringement happens, she might have incentives to settle in order to avoid costly litigation.

We make the following assumptions,

Assumption 1 $\pi_i(\theta_p, \theta_i) = \pi_i(\Delta)$ and $\pi_p(\theta_p, \theta_i) = \pi_p(\Delta)$, where $\Delta = \theta_i - \theta_p$. Both functions are twice continuously differentiable with $\pi_p'(\Delta) \leq 0$ if $\theta_p > \theta_i$.

[9] Here we assume that after the decision in court, each party pays her own expenses, L_i and L_p. In a more general setup the alleged infringer would pay a proportion $\tau(\theta_p, \theta_i, b)$ of the total litigation cost $L = L_p + L_i$ that could include compensations and other penalties.

[10] This structure is equivalent to the model presented by Meurer (1989). As pointed out before, one of the main differences is where the uncertainty is placed. In his case θ_p remains unknown. Here we assume that θ_i is the only private information once the first invention has been patented. We will see that this difference has dramatic effects on the implications from both models.

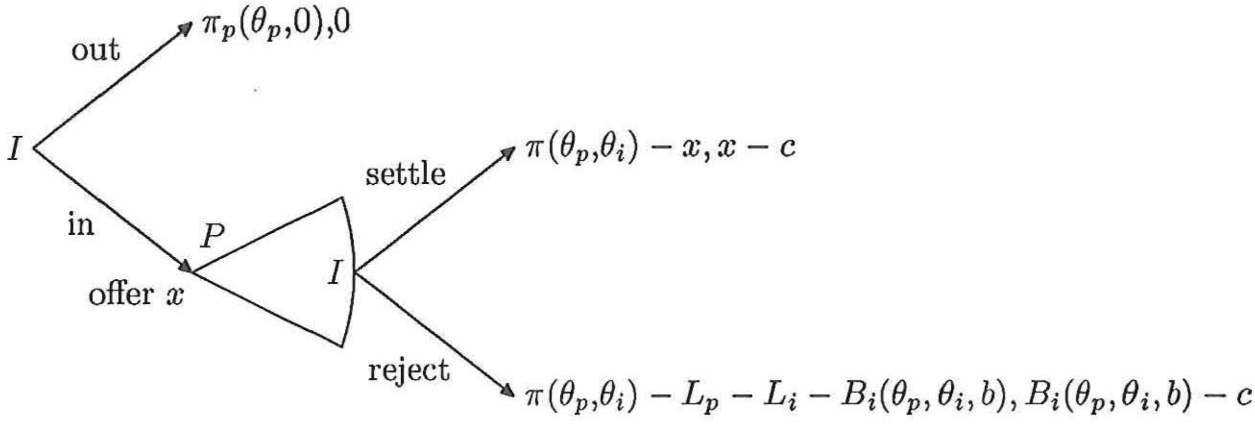

Figure 1: The Entry Game

Assumption 2 $q(\theta_p, \theta_i, b) = q(\Delta, b)$. q is a twice continuously differentiable function in all its arguments with $\frac{\partial q}{\partial b} < 0$ and, for all Δ and b, $0 \leq \frac{\partial B_i}{\partial \theta_i} \leq \exp(\gamma\theta_i)$.

The first assumption, (A.1), implies that profits only depend on the difference in quality among inventions. Also, profits for the patentee are increasing in the quality of her invention when she has the technological lead, and hence they are obviously decreasing as the second firms increases her quality. This structure is consistent with a variety of vertical differentiation models that fit rather well in a market with successive innovations.

Assumption (A.2) means that the probability that the infringer has to obtain a favorable verdict in court is decreasing in the propensity of courts to rule in favor of the patentee, b. Also, whether an invention infringes a patent or not depends essentially on Δ. Some evidence, reported by Chang (1995), tends to support this assumption. Courts seem to rely on the difference between innovations as a criterium to consider infringement. Finally, the expected revenue for the infringer is higher when her innovation differs more from the mere spillovers of the original invention.

Even though inventions are distributed in the $[0, \infty)$ interval, not all of them will be patented. The reason is that small ideas might not be worth to be developed, because they will be very easily superseded, and so the profits obtained from them will not compensate the litigation and development costs involved. We will call $\theta_{\min}$ the threshold level above which it will be profitable to patent an invention.

We require the following additional assumption:

Assumption 3 $\pi_p(-\theta_p) \geq \pi_p(\theta_i - \theta_p) + \pi_i(\theta_i - \theta_p) - c$ for all θ_i. Moreover, for all θ_i and θ_p, $\pi'(\theta_i - \theta_p) \leq \exp(\gamma\theta_i)$ and

$$\int_{\alpha\theta_p}^{\infty} \pi(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p)d\theta_i \leq \pi(-(1-\alpha)\theta_p).$$

Clearly, if the first part of (A.3) did not hold, litigation would most likely not occur, since there still would be room for profitable ex-post negotiation, once infringement had occurred. Moreover, because renegotiation would be expected, any offer made by the patentee would be non-informative and irrelevant. The last condition also imposes an upper bound, but in this case in the ex-ante joint profits accrued by the entrance of the new firm. This bound insures that the expected revenue obtained by the patentee is increasing in the quality of her own invention.

Once infringement occurs, what will be the amount chosen by the patentee to settle? If $x \geq c$ all potential infringers will enter, since by staying out they will obtain a profit of $0 \leq x - c$. On the other hand if $x < c$ nobody will accept the offer and so, some inventions might not be implemented.[11] Hence, we need to consider both cases.

An infringer will accept an offer $x \geq c$ if the expected profits from litigation are smaller than the amount offered to settle. From (A.1) and (A.2) it is clear that only for the best alternative inventions, the infringer will not accept the settlement and take her chances of going to court, since the expected profits from doing it are higher[12]. Hence, the best invention accepting the settlement, θ_s will satisfy,

$$B_i(\theta_s - \theta_p, b) = x \tag{2}$$

By a similar argument, if some firms do not enter - in the case when $x < c$ -, they will be the ones with the smallest inventions. Call the best of those inventions θ_0. Therefore,

$$B_i(\theta_0 - \theta_p, b) = c \tag{3}$$

From (A.1), (A.2) and the definition of both θ_s and θ_0 it is clear that $\theta_s \geq \theta_0$ if and only if $x \geq c$.

If $x \geq c$ the profits for the patentee $W_{x \geq c}$ will be,

$$\begin{aligned} W_{x\geq c}(\theta_p, b, L_i, L_p) &= \max_{x \geq c} \int_{\alpha\theta_p}^{\infty} \pi(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p)d\theta_i - x\Phi(\theta_s - \alpha\theta_p) \\ &\quad - \int_{\theta_s}^{\infty} (L_p + L_i + B_i(\theta_i - \theta_p, b))\phi(\theta_i - \alpha\theta_p)d\theta_i \\ &\quad s.t.\ (2) \end{aligned} \tag{4}$$

That is, the patentholder will make an offer x and all infringers below θ_s will accept. Firms with better inventions will rely on litigation to solve the dispute.[13] In spite of having to pay

[11] To be precise, if $x = c$ the equilibrium is undetermined, since some infringers will be indifferent between entering and setteling or staying out. In order to keep things simple, we assume that *firms will enter unless they make strictly negative profits.* This simplification has no effect on the final result.

[12] We do not allow the patentee to choose whether to go to court or not once the offer is made and rejected. In Appendix B we provide a set of conditions under which this option will keep the sequential equilibrium of the game unaffected.

[13] Waldfogel (1993) estimates the probability of winning of the patentholder before filing the case. He obtains that it is much higher than the probability of inventions that are finally litigated. This restriction suggests that,

an amount x to all inventors below θ_s, this offer might be desirable if it allows to avoid some litigation costs. This will usually be the case for small inventions.

However, if the patentee offers an amount $x < c$, firms with inventions worse than θ_0 will not enter, while the rest will always choose to litigate. Therefore $W_{x<c}$ will be,

$$\begin{aligned} W_{x<c}(\theta_p, b, L_i, L_p) &= \int_{\theta_0}^{\infty} \left[\pi(\theta_i - \theta_p) - L_p - L_i - B_i(\theta_i - \theta_p, b)\right] \phi(\theta_i - \alpha\theta_p) d\theta_i \\ &\quad + \Phi(\theta_0 - \alpha\theta_p)\pi_p(-\theta_p) \\ &\quad s.t.\ (3) \end{aligned} \tag{5}$$

So the patentee will maximize the following function,

$$W(\theta_p, b, L_i, L_p) = \max\left\{W_{x\geq c}(\theta_p, b, L_i, L_p), W_{x<c}(\theta_p, b, L_i, L_p)\right\} \tag{6}$$

We make the following additional assumption:

Assumption 4 $\frac{\partial^2 B_i}{\partial^2 \theta_i} \geq 0$, $\frac{\partial^2 B_i}{\partial \theta_i \partial b} > 0$.

The meaning of (A.4) is clear. The marginal gain that the infringer expects to get from litigation is non-decreasing in the size of θ_i and increasing in b. The first condition is related to the fact that as θ_i raises, it becomes increasingly easy for courts to rule that no infringement occurs. This condition greatly limits the deterrence effect that the patentee can obtain when facing large alternative inventions. The second condition states that the negative effect of b in the profits from litigation dampens as θ_i increases.

The values of θ_0, θ_s and x depend on θ_p, b, L_p and L_i.The next lemma shows how we expect $\theta_0(\theta_p, b, L_i)$, $\theta_s(\theta_p, b, L_p)$ and $x(\theta_p, b, L_i, L_p)$ to move according to changes in those variables.

Lemma 1 *Under (A.1) to (A.4),*

(i) When there is no settlement, $\theta_0(\theta_p, b, L_i)$ is increasing in θ_p, L_i and b. Moreover, $\frac{\partial \theta_0}{\partial \theta_p} = 1$.

(ii) When there is settlement, $\theta_s(\theta_p, b, L_p, L_i)$ is increasing in θ_p, L_p and L_i and decreasing in b. Moreover, $\alpha \leq \frac{\partial \theta_s}{\partial \theta_p} \leq 1$.

(iii) When $x \geq c$, $x(\theta_p, b, L_i, L_p)$ is a decreasing function of θ_p and b.

Proof. See the Appendix. ■

It is interesting to notice that the offer is decreasing in b. This result is similar to the one obtained by Green and Scotchmer (1995). In their case, bigger breadth improves, also unambiguously, the bargaining power of the patentee, allowing her to obtain a higher share of the total profits.[14]

as in the model, low value innovations are settled, while only the more confident infringers will take the risk of going to court.

[14]However, our results come from using a different concept of patent breadth. Here patents do not guarantee absolute protection, because the final outcome depends on a random component related to the size of the improvement. The interpretation that courts make of patent claims is sometimes substantially different to the one intented by the Patent Office.

The previous lemma does not allow us to conclude if it will ever be worth to deter potential infringers. The following two results characterize for α sufficiently large, when settlement will take place.

In the following lemma we characterize the function $W(\theta_p, b, L_i, L_p)$. As a matter of design, it is important to know how profits are affected by the breadth of the patent. The patentee will be better off when the protection granted by the patent increases. Also, W increases when L_i raises. The reason is that an increase in the litigation costs faced by the infringer weakens her position, allowing the patentee to deter entry more effectively.

Lemma 2 *The function $W(\theta_p, b, L_i, L_p)$ is increasing in θ_p, L_i and b and decreasing in L_p.*

Proof. See the Appendix. ■

The next lemma, shows that when α is large, settlement will be offered only when the innovation obtained by the patentee is small.

Lemma 3 *There exists some $\alpha^* < 1$ such that for all $\alpha \geq \alpha^*$, if there is a value $\theta_p^* < \infty$ for which $W_{x\geq c}(\theta_p^*, b, L_i, L_p) = W_{x<c}(\theta_p^*, b, L_i, L_p)$, it is unique. Moreover, the amount offered to settle, x, will be greater than c only for $\theta_p \leq \theta_p^*$.*

Proof. See the Appendix. ■

According to *Lemmas 2* and *3* the patentee will be able to credibly commit to litigate any infringement when the invention has a quality higher than a certain level θ_p^*. If $\theta_p^* < \infty$ for all b, deterrence will be preferred for θ_p sufficiently large. That is,

$$x(\infty, b, L_i, L_p) < c \text{ for all } b \in (0, \infty)$$

It is important to notice that the patentholder will never offer an amount $x = c$. To see it we solve for $W_{x\geq c}$ when $x = c$ from equation (4).

$$\begin{aligned} W_{x\geq c}(\theta_p, b, L_i, L_p)\big|_{x=c} &= \int_{\alpha\theta_p}^{\infty} \pi(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p)d\theta_i - \Phi(\theta_0 - \alpha\theta_p)c \\ &\quad - \int_{\theta_0}^{\infty} (B_i(\theta_i - \theta_p, b) + L_p)\phi(\theta_i - \alpha\theta_p)d\theta_i \\ &= \int_{\alpha\theta_p}^{\theta_0} [\pi(\theta_i - \theta_p) - c]\, \phi(\theta_i - \alpha\theta_p)d\theta_i \\ &\quad + \int_{\theta_0}^{\infty} (\pi(\theta_i - \theta_p) - L_p - L_i - B_i(\theta_i - \theta_p, b))\phi(\theta_i - \alpha\theta_p)d\theta_i \end{aligned}$$

Due to (A.3) this expression is smaller than $W_{x<c}$, since $\pi_p(-\theta_p) \geq \pi(\theta_i - \theta_p) - c$. This suggests that because the patentee offers a settlement to all inventions below θ_s, the cost of deterring entry is much lower than the price that the patentee is paying to avoid litigation

costs. This settlement involves giving up an amount c for inventions that otherwise will not be produced.[15]

In the remaining part of the paper, we will assume that $\theta_{\min} < \theta_p^* < \infty$ since this is the more general case. Notice, however, that if $\theta_p^* < \theta_{\min}$ settlement will never be offered, and similarly, if $\theta_p^* = \infty$ entry-deterrence will never be optimal.

Lemma 4 *For α sufficiently large, the maximum invention for which it is profit-maximizing to settle, θ_p^*, is increasing in L_p and decreasing in b.*

Proof. See the Appendix. ■

The intuition for this result is that inventions with higher quality are going to be more successful in court. Hence when comparing the sure cost from licensing with the expected loss from litigation, for better inventions it is more likely that it pays off to deter infringement as b increases. On the other hand, higher litigation costs for the patentee will induce more settlement, making the entry-deterrence strategy lose some of its appeal.[16]

Figure 2 shows the relationship between $\theta_0(\theta_p, b, L_p, L_i)$, $\theta_s(\theta_p, b, L_p)$ and $\theta_p^*(b, L_i, L_p)$. The point in which θ_0 and θ_s intercept corresponds to the case where $x = c$. We have seen that such an offer will never be made, and therefore, $\theta_p^*(b, L_i, L_p)$ must be to the left of this point.

This model also offers some predictions about the rates of infringement, settlement and trial. We summarize them in the next proposition.

Proposition 5 *Under (A.1) to (A.4),*

(i) The Infringement rate, defined as

$$In(\theta_p, b, L_i) = \begin{cases} 1 - \Phi(\theta_0 - \alpha\theta_p) & \text{if } \theta_p > \theta_p^* \\ 1 & \text{otherwise} \end{cases}$$

is weakly decreasing in L_i and b and non-increasing in θ_p.

(ii) The Settlement rate, defined as

$$St(\theta_p, b, L_i, L_p) = \begin{cases} \Phi(\theta_s - \alpha\theta_p) & \text{if } \theta_p \leq \theta_p^* \\ 0 & \text{otherwise} \end{cases}$$

is increasing in θ_p, L_p and L_i for $\theta_p \leq \theta_p^*$ and zero elsewhere. It is non-increasing in b.

(iii) The Trial rate, defined as

$$Tr(\theta_p, b, L_i, L_p) = \begin{cases} 1 - \Phi(\theta_s - \alpha\theta_p) & \text{if } \theta_p \leq \theta_p^* \\ 1 - \Phi(\theta_0 - \alpha\theta_p) & \text{if } \theta_p > \theta_p^* \end{cases}$$

is strictly decreasing in θ_p and non-increasing in L_i and L_p everywhere except where $\theta_p = \theta_p^*$. It is increasing in b when there is settlement and decreasing otherwise.

[15] This provides answer to footnote 11. Regardless of what assumption we made on infringers' behavior when $x = c$ the result will not change, since such an offer will never be made.

[16] Notice that the trade-off between entry-deterrence and settlement depends greatly on L_p. The first strategy avoids payments to small inventions at a cost of litigating more of them, since $\theta_0 < \theta_s$.

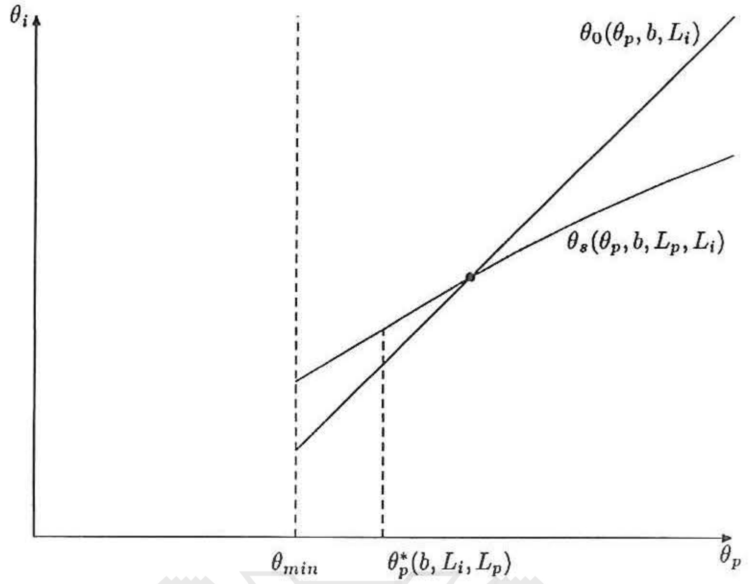

Figure 2: Changes in θ_0 and θ_s as a function of θ_p for a fixed b.

Proof. All (i) to (iii) come directly from the previous lemmas and the definition of θ_0 and θ_s. ■

A very surprising result is the fact that when there is no settlement, broader patents are expected to be litigated less often. This finding is broadly consistent with the evidence presented by Lanjouw and Schankerman (1997). They also obtain that after controlling for variables affecting the value of the invention, the relationship between litigation and breadth is negative. The explanation they give is that it is more difficult to detect infringement in patents with breadth that extends to larger areas. Here, we give an alternative justification. That is, broader patents are litigated less often because infringers see more unlikely to win in court and so, more of them decide not to enter.

This model also predicts that better inventions are litigated less often. To be precise, we observe that the trial rate is decreasing in θ_p everywhere but in the point $\theta_p^*(b, L_i, L_p)$. For innovations smaller than θ_p^* this is due to the increase in settlement, while in the case of better innovations it is the entry-deterrence effect what plays a role.

This relationship is apparently counterfactual. It has been reported in other studies such as Lanjouw and Lerner (1997) that the size of the stakes - estimated using future citations of a patent - has a positive influence in the probability of trial. Nevertheless, some other factors absent in this model are very likely to play an important role. For example, building up a reputation of being tough on infringement is more appealing for better innovations, since competing inventions are more likely to appear.

Other reasons involve for example, the effect of litigation costs. While we have assumed that they are constant, in general they affect the outcome of the trial, and so they are a strategic variable. It is reasonable to think that firms with higher size innovations - and specially big corporations - will have lower litigation costs, and so, might decide to litigate more often.

A third reason is that we have supposed that inventions are somehow correlated, and hence, better innovators will most likely face better infringers. However, the relationship might be stronger than the one modelled here, and might more than compensate the deterrence effect which is here predominant.

Finally, and as we discuss in *Section 3*, it is common practice of courts to assign bigger breadth to better inventions, making it easier to prove infringement. Therefore, the parameter b seems to be somehow correlated with θ_p, affecting the relationship between entry and size of invention.

Other implications of this model are that settlement is only used for small inventions. On one hand any increase in θ_p improves the bargaining position of the patentee. This allows her to offer smaller amounts to settle with infringers. Moreover, as θ_p increases, it is more in the infringer's interest to settle, because it has less chances in court. Therefore, when the quality of the original invention raises, we see an increase in the amount of innovations settled for lower amounts, up to a point, θ_p^*, above which it is not profitable for the patentee to settle anymore.

As we would expect, litigation decreases when L_i or L_p increase, although for different reasons. An increase in L_p gives more incentives to the patentholder to settle, while higher L_i deters more entry. For this last reason, the rate of infringement is also decreasing in L_i.

Finally, we obtain the reasonable result that the probability of settlement is increasing in the costs incurred by the infringer in litigation, at least for small innovations, which is consistent with the empirical findings. It is worth emphasizing that some of the previous models in the literature -see Meurer (1989)- obtain the opposite implication.

2.2 The Decision of the Patentholder

In the first stage, when a firm comes up with an idea, it can decide whether to invest in the production of the invention and go to the Patent Office or not to enter. The patent, by paying an amount f, will offer a certain coverage, interpreted as a probability of succeeding in case of infringement.

The expression for $W(\theta_p, b, L_i, L_p)$ in the previous section, represents the expected *a priori* profits that the patentholder will obtain in the case that an alternative invention might be implemented. Therefore, when a firm decides to patent, it takes into account that profits will be $\pi_p(-\theta_p)$ with probability $1-\lambda$, but with probability λ, potential infringement will occur. In any case, it will incur a cost of $c+f$, due to the expense needed to invent and the price charged for the patent.

$$V_p(\theta_p) = (1-\lambda)\pi_p(-\theta_p) + \lambda W(\theta_p, b, L_i, L_p) - c - f \tag{7}$$

Notice that all inventions might not be patented. We have called before $\theta_{\min}$ the level for which $V_p(\theta_p) = 0$. Obviously, if $V_p(0) > 0$, $\theta_{\min} = 0$. Moreover, due to *Lemma 2*, V is strictly increasing in θ_p, implying that $\theta_{\min}$ is unique.

From the previous results, a natural implication is stated in the following lemma:

Lemma 6 *Higher protection and lower litigation costs incurred by the patentee will increase the amount of patented inventions. That is,* $\frac{\partial\theta_{\min}}{\partial L_p} \geq 0$ *and* $\frac{\partial\theta_{\min}}{\partial b} \leq 0$.

Proof. See the Appendix. ■

The fact that in recent years we have seen an increase in legal costs together with a decrease in the number of patented inventions tends to support this result.

Remark 1 *We implicitly assume that if the inventor does not patent her invention, other firms can produce and sell the same good - without incurring in any cost -, until profits go to zero, and therefore they do not cover the invention costs. Therefore, without patents, the firm will not invest in turning her idea into an invention. Alternatively, we could think that the patentee could produce the good without the protection of a patent, but any other firm could obtain an idea and turn it into an invention without worrying about litigation. In this case, profits would be* $\pi_p(\theta_i - \theta_p)$ *if the potential infringer with quality* θ_i *enters and* $\pi_p(-\theta_p)$ *otherwise. It is easy to prove that as long as* $\pi_p'(\theta_i - \theta_p) \leq \exp(\gamma\theta_i)$ *the previous lemma still holds.*

3 Discussion: The Concept of Patent Breadth

This model differs from the previous literature in the concept of patent breadth used. While most of the models assume that the scope of a patent is exogenous and absolute, here we introduce an *endogenous* concept of breadth. This means that the protection that patents concede is partial and in general depends on the quality of the alternative invention considered.

The judicial doctrine on Infringement is based in several concepts.[17] An invention can infringe literally the text of a patent, and in this case direct infringement is called. However, most of the cases do not fall in this category. In those situations, courts rely essentially on two concepts: the *Doctrine of Equivalents* and the *Legitimate Design-Around.* The first states that infringement also applies to inventions that although are different in some respects to the original one, accomplish essentially the same result. The second concept is embedded in the very essence of patents: to give incentives to innovators to find different - potentially better - ways to attain the same goal. As the next paragraph explains, the combination of both defines whether there is infringement or not.

> "*A court has to measure infringement by a yardstick with the doctrine of equivalents on the near end of the stick and the doctrine of legitimate design-around on the far end of the stick. The length of the yardstick taken up by the doctrine of equivalents depends on the particular patented invention. A pioneer or basic patent may be entitled to a very long portion of the stick; a small improvement in a crowded art may be entitled to only a short distance on the near end of the stick*" (Kintner and Lahr(1975), pp.77-78).

[17] This part is inspired in Kintner and Lahr (1975).

Along this paper, we have used the parameter b to denote the propensity of courts to settle in favor of the patentholder. This can be interpreted, now from a legal viewpoint, as the portion of the stick that is covered by the Doctrine of the Equivalents. There is some systematic empirical evidence that courts use a narrower range of equivalents in very crowded areas, that here are represented by small improvements with respect to the previous knowledge. Therefore, courts seem to use different values of b, at least to a certain extent.

Because the concept of patent breadth used in this paper is substantially different from the one common in the literature, we need a comparable measure on which we can rely in order to confront them. We introduce next a concept that we call *effective breadth of patent*, involving both the profits that the infringer obtains with and without patents.

Definition 1 *Call the effective breadth of a patent to the ratio*

$$\beta = 1 - \frac{B_i(\theta_i - \theta_p, b)}{\pi_i(\theta_i - \theta_p)}$$

Notice that this ratio must be between 0 and 1. The value β measures what share of profits obtained by the infringer when no patents are enforced, $\pi_i(\theta_i - \theta_p)$, is conceded to the patentholder.

In previous models of sequential innovation, as for example O'Donoghue, Scotchmer and Thisse (1995) it is assumed that an invention with quality θ_p will have a total lagging breadth - that is, protection against worse inventions - of $\Delta_{lagging}$, and a total leading breadth - or with respect to improvements - of $\Delta_{leading}$. In this case, β will take values,

$$\beta = \begin{cases} 1 & \text{if } \theta_i \in [\theta_p - \Delta_{lagging}, \theta_p + \Delta_{leading}] \\ 0 & \text{otherwise.} \end{cases}$$

In contrast, the next lemma shows the form of the function β in this model:

Lemma 7 *Under (A.1) and (A.2), the function β is characterized by:*

(i) $\beta = 1$ if $\theta_i < \theta_0$.
(ii) $\beta = 1 - \frac{c}{\pi_i(\theta_i - \theta_p)} \leq 1$ if $\theta_i = \theta_0$.
(ii) β will be decreasing in θ_i if $\theta_i > \theta_0$.

Proof. Part (i) and (ii) are direct implications from the definition of θ_0. For the last part, we can compute using the definition of B_i in equation (1):

$$\frac{\partial \beta}{\partial \theta_i} = -\frac{\partial q}{\partial \theta_i} - \frac{L_i}{\pi_i(\theta_i - \theta_p)^2}\pi_i'(\theta_i - \theta_p)$$

Since $\frac{\partial B_i}{\partial \theta_i} \geq 0$ we can verify that,

$$-\frac{\pi_i'(\theta_i - \theta_p)}{\pi_i(\theta_i - \theta_p)} \leq \frac{-\partial q / \partial \theta_i}{q(\theta_i - \theta_p, b)}$$

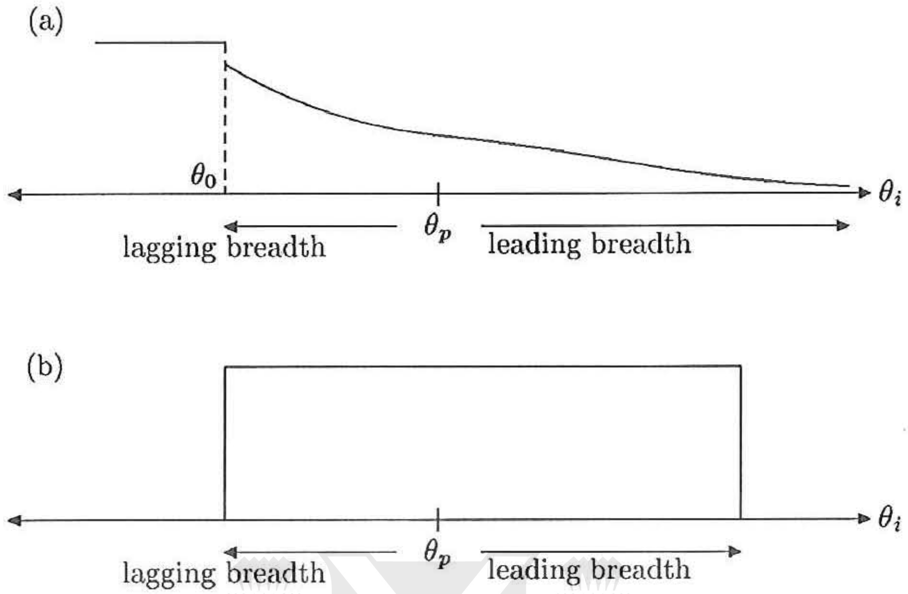

Figure 3: Effective Breadth of a Patent: (a) The concept used in this paper, (b) the concept in O'Donoghue, Scotchmer and Thisse (1995).

And therefore,

$$\frac{\partial \beta}{\partial \theta_i} \leq -\frac{\partial q}{\partial \theta_i} + \frac{L_i}{\pi_i(\theta_i - \theta_p)} \frac{-\partial q/\partial \theta_i}{q(\theta_i - \theta_p, b)} = -\frac{\partial q}{\partial \theta_i}\left[1 - \frac{L_i}{B_i(\theta_i - \theta_p, b) + L_i}\right] < 0$$

■

Figure 3 shows an example of this function. Notice, however, that it does not need to be convex, although it will always be decreasing. We observe that forward breadth is unlimited but weakening as the quality of the alternative invention improves. The reason is that any litigation process induces a cost L_i, part of which will be paid by the infringer, and so, it will always be the case that $B_i(\theta_i - \theta_p, b) < \pi_i(\theta_p, \theta_i)$ even when $\theta_i - \theta_p$ tends to infinity.

4 An Implementable Mechanism

In this model the Patent Office has two mechanisms to regulate inventions: the propensity of courts to settle in favor of the patentee, measured by the parameter b, and the fees charged, f. We want to study if a mechanism consisting in the pairs $\{b(\theta_p), f(\theta_p)\}$ can be implemented. That is, to offer different breadths of patent or protection at different prices.

We will see that as long as the spillover from previous inventions is important enough, a revelation mechanism will be implementable. In the next section we will explore under which circumstances this will be optimal from a social viewpoint.

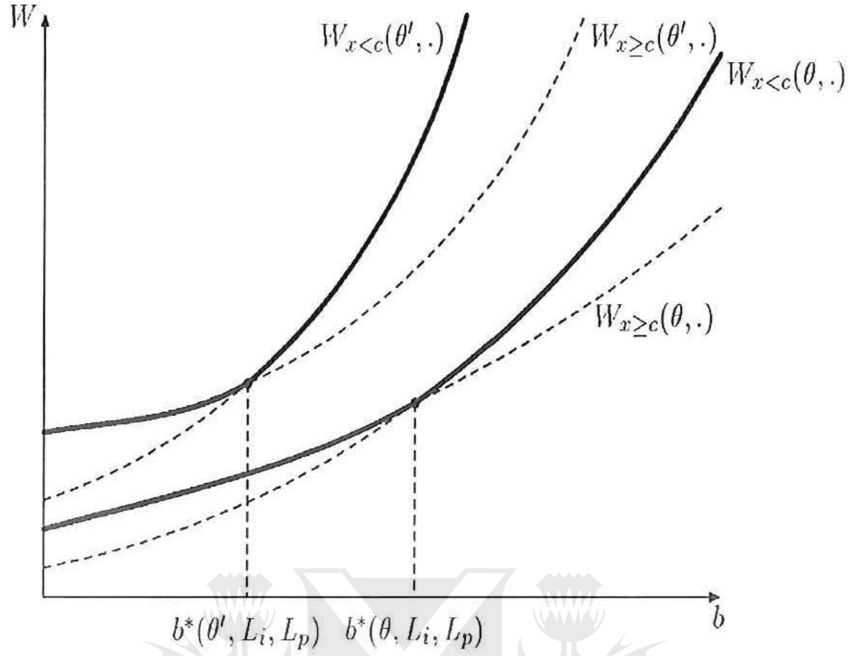

Figure 4: $W(\theta_p, b, L_i, L_p)$ as a function of b for two values of θ_p: θ and θ' where $\theta' > \theta$. Since both $\frac{\partial^2 W_{x\geq c}}{\partial b \partial \theta_p} > 0$ and $\frac{\partial^2 W_{x<c}}{\partial b \partial \theta_p} > 0$, the function is steeper when evaluated at θ'. Therefore, a sufficient condition for the maximum of both functions to be steeper is $b^*(\theta, L_i, L_p) \geq b^*(\theta', L_i, L_p)$.

Proposition 8 *Under (A.1) to (A.4), For α close enough to* 1*, in any implementable mechanism $\{b(\theta_p), f(\theta_p)\}$, b and f are non-decreasing in θ_p.*

Proof. See the Appendix for the whole proof. Here we only provide a sketch. We need to show that $\frac{\partial^2 V_p}{\partial b \partial \theta_p} \geq 0$. In order to do it, we prove that both $\frac{\partial^2 W_{x\geq c}}{\partial b \partial \theta_p} \geq 0$ and $\frac{\partial^2 W_{x<c}}{\partial b \partial \theta_p} \geq 0$, and as *Figure 4* shows, if $\frac{\partial W_{x<c}}{\partial b} \geq \frac{\partial W_{x\geq c}}{\partial b}$ and $b^*(\theta_p, L_i, L_p)$ - that is, the inverse of $\theta_p^*(b, L_i, L_p)$ - is decreasing in θ_p, this becomes a sufficient condition. *Lemma 4* shows that this will be the case for a large α. ■

The interpretation of this result is that in order to distinguish among different inventions, the patent scheme has to give higher protection to better inventions. The reason is that firms with higher inventions can make a better use of their protection and so, they are willing to pay a higher price for it. On the other hand, firms with smaller inventions will be unable to deter most infringers and when facing entry, they will offer higher settlement amounts, decreasing the profitability of their patent. A higher α tends to accentuate this effect, by making settlement more and more likely to occur.

5 The Optimal Mechanism

In order to describe the optimal mechanism, we need to characterize the social welfare obtained from the inventions and the patent system. We use the consumer surplus and producer surplus as a measure of this welfare. While in this case, the producer surplus is the sum of the profits of the patentee and the infringer - when there is entry -, some assumptions about the consumer surplus need to be made. This surplus, $cs(\theta, \pi)$, depends only on the level of improvement achieved and the total profits.

In particular, we assume that consumers obtain a level of welfare that is increasing in the size of the invention. The reason is that better innovations, keeping constant the level of profits, will never harm the agents. An important simplification that we will make is the assumption that consumer surplus is increasing in the *highest* available innovation - or technological frontier -, independently of whether it belongs to the patentholder or the infringer.

Also, the higher the profits that firms obtain from the innovation, the lower will be the welfare of the consumers. Moreover, the decrease is more than proportional. This assumption implies not only that the consumer surplus is decreasing in the profits obtained by the firms, but also that the total surplus, defined as $ts(\theta, \pi) = cs(\theta, \pi) + \pi$, is decreasing in π. The consequence is quite standard: the Dead Weight Loss is non-decreasing in the amount of profits.

The next assumption summarizes the previous ideas:

Assumption 5 $cs(\theta, \pi) = cs(\max_s \theta_s, \pi)$, $\frac{\partial cs}{\partial \theta_p} \geq 0$ and $\frac{\partial cs}{\partial \pi} \leq -1$.

Notice that the consumer surplus can be affected in different ways by the arrival of a second innovation, depending on whether total profits increase or not. In the next section we will impose additional conditions to see how they affect the optimal patent.

When infringement occurs but the patentee does not offer settlement - and therefore $\theta_0 > 0$ -, the total welfare from a patent of quality θ_p becomes:

$$\begin{aligned} S_{x<c}(\theta_p, b, L) &= (1 - \lambda(1 - \Phi(\theta_0 - \alpha\theta_p)))\left[\pi_p(-\theta_p) + cs(\theta_p, \pi_p(-\theta_p))\right] - \lambda(1 - \Phi(\theta_0 - \alpha\theta_p)) \\ &\quad \times (L + c) + \lambda \int_{\theta_0}^{\theta_p} \left[\pi(s - \theta_p) + cs(\theta_p, \pi(s - \theta_p))\right] \phi(s - \alpha\theta_p) ds \\ &\quad + \lambda \int_{\max(\theta_p, \theta_0)}^{\infty} \left[\pi(s - \theta_p) + cs(s, \pi(s - \theta_p))\right] \phi(s - \alpha\theta_p) ds \end{aligned}$$

Here θ_0 is as defined in equation (3) and $L = L_i + L_p$. That is, the total surplus obtained will be resulting from the monopoly of the patentholder with probability $(1 - \lambda(1 - \Phi(\theta_0 - \alpha\theta_p)))$ while in the other cases the infringer will implement her idea and hence there will be litigation, with the corresponding cost L, plus the cost of inventing θ_i. Changes in the litigation parameters will affect this function indirectly through changes in θ_0.

When there is settlement, that is $x \geq c$, the welfare function can be obtained as,

$$S_{x \geq c}(\theta_p, b, L) = (1 - \lambda)\left[\pi_p(-\theta_p) + cs(\theta_p, \pi_p(-\theta_p))\right] - \lambda(1 - \Phi(\theta_s - \alpha\theta_p))L$$

$$-\lambda c \ + \lambda \int_0^{\theta_p} \left[\pi(s-\theta_p) + cs(\theta_p, \pi(s-\theta_p))\right] \phi(s-\alpha\theta_p)ds$$
$$+\lambda \int_{\theta_p}^{\infty} \left[\pi(s-\theta_p) + cs(s, \pi(s-\theta_p))\right] \phi(s-\alpha\theta_p)ds$$

In this case any alternative innovations will be implemented, regardless of its quality. With probability $(1-\lambda)$ the patentee will retain a monopoly power, while in the rest of the occasions, both inventions will coexist. Notice that in this case θ_s is the only element that can be affected by changes in b.

We will study the choice of the optimal b in two stages. First, we will derive what is the optimal b when we restrict it to be constant. Afterwards, we will characterize which is the optimal implementable mechanism.

5.1 Optimal Mechanism with a fixed b

The best that we could achieve when the spillover from one invention to the next is not very important is a constant choice of b. The spillover α might not be sufficiently large to even guarantee that there will be only one cutoff value θ_p^* so that all innovators with inventions above that quality decide to deter entry, as shown in *Lemma 3*. Therefore, we define $\Omega_{x\geq c} \in \Re_+$ as the size of innovations for which it is optimal for the patentee to settle with some infringers. In a similar way, $\Omega_{x<c} \in \Re_+$ will correspond to innovators that avoid any settlement.

The planner's program to be maximized will be the following[18]:

$$\begin{aligned} S(L) \quad = \quad & \max_{b,\theta_{\min},\widetilde{\theta}} \int_{\Omega_{x\geq c}} S_{x\geq c}(\theta_p, b, L)\phi(\theta_p)d\theta_p + \int_{\Omega_{x<c}} S_{x<c}(\theta_p, b, L)\phi(\theta_p) \qquad (8) \\ & +\Phi(\theta_{\min})\lambda \int_{\widetilde{\theta}}^{\infty} \left[\pi(s) + cs(s, \pi(s))\right] \phi(s)ds \\ & s.t. V(\theta_p) \geq 0 \text{ for all } \theta_p \geq \theta_{\min} \end{aligned}$$

Therefore, the optimal b will result from maximizing the social surplus given the individual rationality constraint.

The value $\widetilde{\theta}$ represents the minimum invention allowed to be patented once the first one has been rejected.[19] Since we have assumed that the infringer is the last to achieve an invention, $\widetilde{\theta}$ will be set so that the marginal increase in total welfare is 0.

What will be the optimal b? It is difficult to answer this question without additional assumptions. In the remaining of this section, we will add some conditions under which b will be either 0 or ∞. This will give us some intuition on how b depends on the parameters of the model. First, we can obtain the corresponding first order condition for the problem in (8) as,

[18] The choice variables in this problem will be $b, \theta_{\min}$ and $\widetilde{\theta}$. The parameter f could have been used instead, but it is easy to see that the choice of b and $\theta_{\min}$ already determines it.

[19] We have not explicitly defined the continuation game. We assume that all inventions are patentable. In fact, an invention infringing a patent can itself be patentable.

$$\int_{\Omega_{x\geq c}} \frac{\partial S_{x\geq c}}{\partial \theta_s}(\theta_p, b, L)\frac{\partial \theta_s}{\partial b}\phi(\theta_p)d\theta_p + \int_{\Omega_{x<c}} \frac{\partial S_{x<c}}{\partial \theta_0}(\theta_p, b, L)\frac{\partial \theta_0}{\partial b}\phi(\theta_p) = 0 \tag{9}$$

It is easy to see that

$$\frac{\partial S_{x\geq c}}{\partial \theta_s}(\theta_p, b, L) = \lambda\phi(\theta_s - \alpha\theta_p)L \geq 0$$

Together, this expression and *Lemma 1* imply that the first term in equation (9) will be non-positive. The second part is more involved. We obtain that,

$$\begin{aligned} \frac{S_{x<c}}{\partial \theta_0}(\theta_p, b^*, L) &= \lambda\phi(\theta_0 - \alpha\theta_p)\left[\pi_p(-\theta_p) - \pi(\theta_0 - \theta_p) + c + L + cs(\theta_p, \pi_p(-\theta_p))\right. \\ &\quad \left. - cs(\max(\theta_0, \theta_p), \pi(\theta_0 - \theta_p))\right] \end{aligned}$$

We will make two different sets of assumptions, 6(a) and 6(b):

Assumption 6 (a) $\pi_p(-\theta_p) < \pi(\theta_i - \theta_p)$ for all $\theta_i \leq \theta_p$.

Under Assumption 6 (a), the existence of an infringer increases the total profits from innovations. This case will arise for example if the competition between the infringer and the patentee is rather weak and the innovation created by the infringer is a lower quality substitute of the original one, allowing to discriminate among consumers.

The next proposition provides us with an intuition about when it will be optimal for courts to allow for imitation, and to which extent. It also helps us to compare this model with the existing literature.

Proposition 9 *If in the optimal mechanism with a fixed b there is no settlement and A.6(a) holds, then either $b = \infty$ or $\theta_0 \geq \theta_p$ for all θ_p.*

Proof. Suppose towards a contradiction that the optimal mechanism $b^* < \infty$ is such that for a certain θ_p, $\theta_0(\theta_p, b^*) \leq \theta_p$. Since $\frac{\partial \theta_0}{\partial \theta_p} = 1$, this will be true for any other level of θ_p.The first order condition corresponding to the problem in equation (8) now becomes,

$$\int_{\theta_{\min}}^{\infty} \frac{S_{x<c}}{\partial \theta_0}(\theta_p, b^*, L)\frac{\partial \theta_0}{\partial b}\phi(\theta_p)d\theta_p = 0$$

By assumption (A.3),

$$\frac{S_{x<c}}{\partial \theta_0}(\theta_p, b^*, L) \geq \lambda\phi(\theta_0 - \alpha\theta_p)\left[cs(\theta_p, \pi_p(-\theta_p)) - cs(\theta_p, \pi(\theta_0 - \theta_p)) + L\right] > 0$$

where the last inequality comes from A.6 (a). Therefore, the first order condition will be positive, meaning that the social welfare can be increased by raising b. And this contradicts b^* being optimal in the first place. ■

The interpretation of the previous proposition is quite straightforward. The only reason for which courts should allow imitation would be in the case where the entry of a new firm increases competition, making joint profits decrease and therefore increase the total surplus. If this is not the case, to allow imitators decreases social welfare because resources are wasted in developing products of inferior quality, and litigation costs are incurred. This result is similar to Gallini (1992) where broad patents granted for a short time are optimal because they discourage imitation, and competitors wait for the patent to expire in order to introduce new inventions.

Alternatively, one would think that if the cost of innovation c is not very big, the optimal mechanism should encourage all innovations, even in the case they infringe an existing patent. The reason, is that the lower are these costs, the more difficult it is for a patentee to deter entry, and in this case settlement must be encouraged. To prove this intuition, we will need an alternative assumption:

Assumption 6 (b) $\frac{\partial^2 cs}{\partial^2 \pi} < 0$ and $\pi_p(-\theta_p) \geq \pi(\theta_i - \theta_p)$ for all θ_i.

Proposition 10 *In the optimal mechanism with a fixed b, if A.6 (b) holds, then $b = 0$ for c sufficiently small.*

Proof. In order to prove this proposition, we only need to show that the equation (9) is negative for all b. We already know that this will be the case for $x \geq c$. In the other case, notice that making use of (A.5) and (A.6 (b)), and after a Taylor expansion,

$$\begin{aligned} cs(\theta_p, \pi_p(-\theta_p)) &< cs(\theta_p, \pi(\theta_0 - \theta_p)) + \frac{\partial cs}{\partial \pi}\left[\pi_p(-\theta_p) - \pi(\theta_i - \theta_p)\right] \\ &\leq cs(\max(\theta_p, \theta_0), \pi(\theta_0 - \theta_p)) - \pi_p(-\theta_p) + \pi(\theta_i - \theta_p) \end{aligned}$$

Therefore, we obtain that,

$$\frac{S_{x<c}}{\partial \theta_0}(\theta_p, b^*, L) \leq \lambda\phi(\theta_0 - \alpha\theta_p)\left[c + L\right] \leq \lambda\phi(0)\left[c + L\right]$$

Finally, notice that the set $\Omega_{x \geq c}$ increases as c approaches 0, and so, for c sufficiently small,

$$\begin{aligned} \int_{\Omega_{x\geq c}} \lambda\phi(\theta_s - \alpha\theta_p) L \frac{\partial \theta_s}{\partial b}\phi(\theta_p)d\theta_p &< -\int_{\Omega_{x<c}} \lambda\phi(0)\left[c + L\right]\frac{\partial \theta_0}{\partial b}\phi(\theta_p)d\theta_p \\ &\leq -\int_{\Omega_{x<c}} \frac{S_{x<c}}{\partial \theta_0}(\theta_p, b^*, L)\frac{\partial \theta_0}{\partial b}\phi(\theta_p)d\theta_p \end{aligned}$$

And the First Order Condition becomes negative. ■

5.2 Optimal Revelation Mechanism

If spillovers are important enough, - that is, α is close enough to 1 - we can in general do better than to choose a constant b. In particular, the planner will be able to choose a profile $\{b(\theta_p), f(\theta_p)\}$ in order to maximize social welfare.

Using the Revelation Principle, we only need to focus on revelation mechanisms. Applying *Lemma 3* the social planner's program will be as follows,

$$\begin{aligned} S(L) &= \max_{b,\theta_{\min},\widetilde{\theta}} \int_{\theta_{\min}}^{\theta_p^*(b)} S_{x\geq c}(\theta_p, b, L)\phi(\theta_p)d\theta_p + \int_{\theta_p^*(b)}^{\infty} S_{x<c}(\theta_p, b, L)\phi(\theta_p)d\theta_p \\ &+\Phi(\theta_{\min})\lambda \int_{\widetilde{\theta}}^{\infty} \left[\pi(s) + cs(s,\pi(s)) - c\right]\phi(s)ds \end{aligned} \tag{10}$$

$$\begin{aligned} s.t. \quad \frac{\partial b}{\partial \theta_p} &\geq 0 \\ V(\theta_p) &\geq 0 \text{ for all } \theta_p \geq \theta_{\min} \end{aligned}$$

With the choice of b, the planner can affect through the values of θ_s and θ_0 the offer that a certain patentholder will make in order to settle. The first restriction corresponds to the Incentive Compatibility constraint. As obtained in *Proposition 8*, in order to provide incentives, the breadth of the patent has to be non-decreasing in the quality of the invention. The second condition states that all innovators should obtain at least as much as they would get by not producing the invention.

In the case where there is settlement, the best that the planner can do (contingent on $x \geq c$) is

$$\max_b \ \theta_s(\theta_p, b, L_p)$$

$$\text{s.t. } \theta_p \leq \theta_p^*(b, L_i, L_p)$$

The optimal $b(\theta_p)$ will be in this case the maximum b such that the constraint is satisfied. According to *Lemma 4*, the constraint will be tighten for smaller b.

This leads to the following proposition:

Proposition 11 *If $\alpha < 1$, and if in the optimal mechanism $b > 0$, b must be an increasing function of θ_p, at least when there is settlement.*

Proof. Assume towards a contradiction that the optimal mechanism involves a constant $b = \widehat{b}$. We could actually consider the case in which b is decreasing in θ_p but we have seen in *Proposition 8* that this will never be implementable.

Given this level $\widehat{b}$ all patentholders with innovations below $\theta_p^*(\widehat{b}, L_i, L_p)$ will settle, offering values of x above c. For the rest, all entrants will be litigated.

From the previous analysis, we can construct another mechanism $\{b(\theta_p), f(\theta_p)\}$ such that for all $\theta_p > \theta_p^*(\widehat{b}, L_i, L_p)$ we have $b(\theta_p) = \widehat{b}$ and for lower θ_p, $b(\theta_p) < \widehat{b}$. Because $\widehat{b}$ is still available, the cutoff value of θ_p^* and $\theta_{\min}$ will not increase. It is easy to see that with the appropriate choice of $f(\theta_p)$ increasing in θ_p, $\theta_{\min}$ will remain unchanged. Clearly, this mechanism improves social welfare with respect to the original one, whenever settlement was optimal, leaving the rest unchanged. Therefore, $\widehat{b}$ could not be optimal in the first place. ■

This result is somehow different to the recommendations made by Chang (1995). He claims that small inventions should have a broader coverage, and his argument relies on the fact that this allows the patentee to have a better bargaining position and therefore to internalize part of the welfare increase created by the invention.

In this case, there are two reasons why the opposite sort of mechanism is optimal. On one hand, the existence of settlement is preferred to deterrence for most of the inventions below θ_p, since this increases competition among firms.[20] On the other hand narrower breadth makes inventions of smaller size more vulnerable to infringement, and therefore it gives them more incentives to settle, avoiding costly litigation.

Clearly, we can achieve the optimal level of R&D through the right choice of fees $f(\theta_p)$. In particular, and if b is very low for small inventions, the fees required could be negative, subsidiazing then, part of the research cost. In comparison, currently these fees are very low, and are mainly used to cover the expedition cost and other bureaucratic charges.

Remarkably enough, the mechanism proposed here is precisely the way in which courts handle infringement suits. As we pointed out before, courts require the patentholder to prove imitation to a higher extent for smaller inventions in order to rule that the patent has been infringed. The present model would also suggest that such a discretion should be used beforehand in order to screen among inventions.

6 Conclusion

The goal of this paper was to present a model of patent design that could take explicitly into account the effect of litigation, or the threat of using it. We have shown that with private information on the quality of the innovation there is in equilibrium a certain amount of litigation, together with the settlement of low quality inventions.

We also proposed a new way of analyzing the protection that a patent concedes, that we call *the effective breadth of a patent.* This concept differs from the previous approaches on

[20]To be specific, inventions below θ_p such that

$$\left(\frac{\partial cs}{\partial \pi}(\theta_p, \pi(\theta_i - \theta_p)) + 1\right) \frac{\partial \pi}{\partial \theta_i}(\theta_i - \theta_p) > c$$

will increase social welfare.

patents in the consideration that the breadth is not an absolute concept. In fact, it depends on institutions, in this case courts, that enforce this notion in ways somehow different to the one intended by the Patent Office.

The structure matches some of the stylized facts regarding patent litigation. Most remarkably it is able to reproduce some recent findings by Lanjouw and Schankerman (1997), who show that the breadth of the patent is negatively related with the probability of litigation. They speculate with the possibility that with broader patents it is more difficult to detect infringement, since the applications might extend to a diversity of areas. According to this model, though, this fact might also be due to the entry-deterrence effect that broad patents have on bad inventions.

The model, as in Nalebuff (1987), predicts that bad inventions will be settled, while only good quality infringers will go to court. This explains why, according to Waldfogel (1993), the probability that the patentee wins an infringement suit is smaller when we condition on having filed the case.

However, the model fails to reproduce the empirical result that good inventions are litigated more often. This might be due to several reasons. First, this model does not take into account effects such as reputation, which have proved to be relevant. Second, the assumption that the quality of the patented invention is known to the infringer might also be driving this result. Finally even though we have assumed that infringers build up to a certain extend on the quality achieved by previous inventors, the structure we have used is somehow restrictive. Changing it could help reconcile the model with this empirical fact.

We also study the combination of patent breadth and fees paid by innovators that maximizes social welfare. We find that the optimal patent should have a smaller breadth for patents with low quality, and this protection should be increasing as quality raises. The reason is that for small inventions it is very likely to have infringement, and in this case, the optimal mechanism should be designed in order to maximize the possibility of settlement.

This sort of models could be used to address some questions regarding policy, and in particular how changes in the legal environment could affect the allocation of resources in research.

One such a question is the effect of *preliminary injunction* on social welfare. According to Lanjouw and Lerner (1996) this legal motion is becoming popular in recent years. Clearly this changes the structure of the game studied in this paper and can affect the incentives of both the patentee and the potential infringer to settle or litigate.

Also, some literature has focussed in the allocation of litigation costs and their effect on inventions. This model takes explicitly into account the legal costs that both the patentholder and the infringer bear, making easy to compare different allocation systems. Preliminary results show that in some cases, the way in which costs are allocated has similar effects as our parameter b. Whether a revelation mechanism concerning those costs can be implemented or not remains to be explored.

A Appendix

Here we reproduce the proofs of most of the results in the paper.

Proof to Lemma 1:

For (i) it is enough to notice that $\frac{\partial B_i}{\partial \theta_i} = -\frac{\partial B_i}{\partial \theta_p} > 0$ $\frac{\partial B_i}{\partial b} < 0$ and $\frac{\partial B_i}{\partial L_i} = -1$. Therefore, by standard use of the Implicit Function Theorem on equation (3) we obtain $\frac{\partial \theta_0}{\partial b} \geq 0$, $\frac{\partial \theta_0}{\partial L_i} \geq 0$ and $\frac{\partial \theta_0}{\partial \theta_p} = 1$.

For the second part, notice that in order to have settlement, $x \geq c$ and so, we only need to consider the last case. From equation (4) we can obtain the first order condition regarding x. However, due to the one-to-one relationship between x and θ_s, we can instead solve for θ_s which turns out to be easier. It can be shown that given (A.4) the following first order condition defines the solution to the problem.

$$\frac{\partial B_i}{\partial \theta_i} = (L_p + L_i)\,\Gamma(\theta_s - \alpha\theta_p)$$

where $\Gamma(z) = \frac{\phi(z)}{\Phi(z)}$.

From this condition, and using again the Implicit Function Theorem, it we obtain the following results:

$$\begin{aligned}
\frac{\partial \theta_s}{\partial \theta_p} &= \frac{\frac{\partial^2 B_i}{\partial^2 \theta_i} - \alpha\,(L_p + L_i)\,\Gamma'(\theta_s - \alpha\theta_p)}{\frac{\partial^2 B_i}{\partial^2 \theta_i} - (L_p + L_i)\,\Gamma'(\theta_s - \alpha\theta_p)} \geq 0 \\
\frac{\partial \theta_s}{\partial b} &= \frac{-\frac{\partial^2 B_i}{\partial \theta_i \partial b}}{\frac{\partial^2 B_i}{\partial^2 \theta_i} - (L_p + L_i)\,\Gamma'(\theta_s - \alpha\theta_p)} < 0 \\
\frac{\partial \theta_s}{\partial L_i} &= \frac{\partial \theta_s}{\partial L_p} = \frac{\Gamma(\theta_s - \alpha\theta_p)}{\frac{\partial^2 B_i}{\partial^2 \theta_i} - (L_p + L_i)\,\Gamma'(\theta_s - \alpha\theta_p)} > 0
\end{aligned}$$

From these expressions one can observe that $\alpha \leq \frac{\partial \theta_s}{\partial \theta_p} \leq 1$, $\frac{\partial \theta_s}{\partial b} < 0$ and $\frac{\partial \theta_s}{\partial L_p} = \frac{\partial \theta_s}{\partial L_i} > 0$.

Finally, for part (iii), $x(\theta_p, b, L_i, L_p)$ is defined as

$$x(\theta_p, b, L_i, L_p) = B_i(\theta_s - \theta_p, b)$$

Therefore,

$$\begin{aligned}
\frac{\partial x}{\partial \theta_p} &= \frac{\partial B_i}{\partial \theta_i}\left[\frac{\partial \theta_s}{\partial \theta_p} - 1\right] < 0 \\
\frac{\partial x}{\partial b} &= \frac{\partial B_i}{\partial b} + \frac{\partial B_i}{\partial \theta_i}\frac{\partial \theta_s}{\partial b} < 0
\end{aligned}$$

Proof to Lemma 2:

From equation (6) we only need to prove that this is true for $W_{x<c}$ and $W_{x\geq c}$. That is, they are increasing in θ_p, b and L_i and decreasing in L_p.

In the first case, when $x \geq c$, and after integrating by parts, we obtain,

$$\begin{aligned}\frac{\partial W_{x\geq c}}{\partial \theta_p} &= (1-\alpha)\Big\{(L_p+L_i)\,\phi(\theta_s-\alpha\theta_p) - \int_{\theta_s}^{\infty}\left[B_i(\theta_i-\theta_p,b)-B_i(\theta_s-\theta_p,b)\right]\phi'(\theta_i-\alpha\theta_p)d\theta_i \\ &\quad + \int_{\alpha\theta_p}^{\infty}\left[\pi(\theta_i-\theta_p)-\pi(-(1-\alpha)\theta_p)\right]\phi'(\theta_i-\alpha\theta_p)d\theta_i\Big\}\end{aligned}$$

This is positive using assumption (A.2) for the second term and (A.3) for the last one. Notice that in this final term we use the fact that $\phi'(x) = -\gamma\phi(x)$.

For the case where $x < c$, and again after some manipulation,

$$\begin{aligned}\frac{\partial W_{x<c}}{\partial \theta_p} &= (1-\alpha)\left\{\int_{\theta_0}^{\infty}\left[\pi(\theta_i-\theta_p)-\pi_p(-\theta_p)-B_i(\theta_i-\theta_p)-L_p-L_i\right]\phi'(\theta_i-\alpha\theta_p)d\theta_i\right\} \\ &\quad -\pi_p'(-\theta_p)\Phi(\theta_0-\alpha\theta_p)\end{aligned}$$

And this expression is positive given (A.1) and (A.3).

In a similar way, we compute the effect of b as,

$$\frac{\partial W_{x\geq c}}{\partial b} = -\frac{\partial B_i}{\partial b}\Phi(\theta_s-\alpha\theta_p) - \int_{\theta_s}^{\infty}\frac{\partial B_i}{\partial b}\phi(\theta_i-\alpha\theta_p)d\theta_i \geq 0$$

and

$$\frac{\partial W_{x<c}}{\partial b} = \frac{\partial \theta_0}{\partial b}\phi(\theta_0-\alpha\theta_p)\left[\pi_p(-\theta_p)-\pi(\theta_0-\theta_p)+c+L_p+L_i\right] - \int_{\theta_0}^{\infty}\frac{\partial B_i}{\partial b}\phi(\theta_i-\alpha\theta_p)d\theta_i$$

Again, this is positive due to (A.1), (A.3) and *Lemma 1.*

The effect of L_i is positive, as shown below.

$$\frac{\partial W_{x\geq c}}{\partial L_i} = \Phi(\theta_s-\alpha\theta_p)$$

$$\frac{\partial W_{x<c}}{\partial L_i} = -\frac{\partial \theta_0}{\partial L_i}\phi(\theta_0-\alpha\theta_p)\left[\pi_p(-\theta_p)-\pi(\theta_0-\theta_p)+c+L_p+L_i\right] + \Phi(\theta_0-\alpha\theta_p) > 0$$

Finally, the derivative with respect to L_p is computed as,

$$\frac{\partial W_{x\geq c}}{\partial L_p} = \Phi(\theta_s - \alpha\theta_p) - 1 < 0$$
$$\frac{\partial W_{x<c}}{\partial L_p} = \Phi(\theta_0 - \alpha\theta_p) - 1 < 0$$

Proof to Lemma 3:

In order for the functions $W_{x\geq c}$ and $W_{x<c}$ to cross only once, it is sufficient to show that $\frac{\partial W_{x\geq c}}{\partial \theta_p}(\theta_p, b, L_i, L_p) < \frac{\partial W_{x<c}}{\partial \theta_p}(\theta_p, b, L_i, L_p)$. With that and the continuity of both functions we obtain the desired result. We can compute, using the expressions in *Lemma 2,* the following:

$$\left.\frac{\partial W_{x<c}}{\partial \theta_p} - \frac{\partial W_{x\geq c}}{\partial \theta_p}\right|_{(\theta_p,b,L_i,L_p)} = -\pi_p'(-\theta_p)\Phi(\theta_0 - \alpha\theta_p) + (1-\alpha)\left\{\pi_p(-\theta_p)\phi(\theta_0 - \alpha\theta_p) - \int_{\theta_s}^{\infty} \frac{\partial B_i}{\partial \theta_i}\phi(\theta_i - \alpha\theta_p)d\theta_i + \int_{\alpha\theta_p}^{\infty} \pi'(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p)d\theta_i + \int_{\theta_0}^{\infty} \left[\pi(\theta_i - \theta_p) - B_i(\theta_i - \theta_p)\right]\phi(\theta_i - \alpha\theta_p)d\theta_i + (L_p + L_i)\left[\phi(\theta_0 - \alpha\theta_p) - \phi(\theta_s - \alpha\theta_p)\right]\right\}$$

This expression is positive for α big enough or alternatively, if L_p is big enough.

Proof to Lemma 4:

We define $\theta_p^* = g(b, L_i, L_p)$ as the level of θ_p such that $W_{x\geq c}(\theta_p, b, L_i, L_p) = W_{x<c}(\theta_p, b, L_i, L_p)$. Therefore,

$$\frac{\partial g}{\partial L_p} = \frac{\frac{\partial W_{x<c}}{\partial L_p} - \frac{\partial W_{x\geq c}}{\partial L_p}}{\frac{\partial W_{x\geq c}}{\partial \theta_p} - \frac{\partial W_{x<c}}{\partial \theta_p}}$$
$$\frac{\partial g}{\partial b} = \frac{\frac{\partial W_{x<c}}{\partial b} - \frac{\partial W_{x\geq c}}{\partial b}}{\frac{\partial W_{x\geq c}}{\partial \theta_p} - \frac{\partial W_{x<c}}{\partial \theta_p}}$$

The denominator in both expressions is negative when L_p or α are big enough, using *Lemma 3*. For the numerator,

$$\frac{\partial W_{x<c}}{\partial L_p} - \frac{W_{x\geq c}}{\partial L_p} = \Phi(\theta_0 - \alpha\theta_p) - \Phi(\theta_s - \alpha\theta_p) > 0$$

So, we obtain directly that $\frac{\partial g}{\partial L_p} < 0$. For the other case,

$$\begin{aligned}\frac{\partial W_{x<c}}{\partial b}-\frac{\partial W_{x\geq c}}{\partial b} &\geq \frac{\partial B_i}{\partial b}(\theta_s-\theta_p,b)\Phi(\theta_s-\alpha\theta_p)+\frac{\partial \theta_0}{\partial b}\left(L_p+L_i\right)\phi(\theta_0-\alpha\theta_p)\\ &= \frac{\partial B_i}{\partial b}(\theta_s-\theta_p,b)\Phi(\theta_s-\alpha\theta_p)-\frac{\frac{\partial B_i}{\partial b}(\theta_0-\theta_p,b)}{\frac{\partial B_i}{\partial \theta_i}(\theta_0-\theta_p,b)}\left(L_p+L_i\right)\phi(\theta_0-\alpha\theta_p)\\ &\geq \frac{\partial B_i}{\partial b}(\theta_s-\theta_p,b)\Phi(\theta_s-\alpha\theta_p)\left[1-\frac{\phi(\theta_0-\theta_p,b)}{\phi(\theta_s-\theta_p,b)}\right]\geq 0\end{aligned}$$

The second equality comes from the definition $\frac{\partial\theta_0}{\partial b}$ in *Lemma 1.* Finally, the last inequality results from using (A.4) and the first order condition obtained for $W_{x\geq c}$ in *Lemma 1.*

Proof to Lemma 6:

The amount of patented inventions is given by $1-\Phi(\theta_{\min})$ since ideas appear with an exogenous probability. Therefore, to prove that the amount of inventions is increasing in b and decreasing in L_p we only need to check that $\frac{\partial\theta_{\min}}{\partial b}<0$ and $\frac{\partial\theta_{\min}}{\partial L_p}>0$.

From the definition of $\theta_{\min}$,

$$V(\theta_{\min})=(1-\lambda)\pi_p(-\theta_{\min})+\lambda W(\theta_{\min},b,L_i,L_p)-c-f=0$$

And by use of the Implicit Function Theorem and *Lemma 2*,

$$\begin{aligned}\frac{\partial\theta_{\min}}{\partial b} &= \frac{-\lambda\frac{\partial W}{\partial b}}{\frac{\partial\pi_p}{\partial\theta_p}+\lambda\frac{\partial W}{\partial\theta_p}}<0\\ \frac{\partial\theta_{\min}}{\partial L_p} &= \frac{-\lambda\frac{\partial W}{\partial L_p}}{\frac{\partial\pi_p}{\partial\theta_p}+\lambda\frac{\partial W}{\partial\theta_p}}>0\end{aligned}$$

Proof to Proposition 8:

We only need to check that the Spence-Mirlees condition holds. In this case, that

$$\frac{\partial^2 V_p}{\partial b\partial\theta_p}\geq 0$$

We first check that $\frac{\partial^2 W}{\partial b\partial\theta_p}(\theta_p,b,L)\geq 0$. In order to do that we will prove that $\frac{\partial^2 W_{x\geq c}}{\partial b\partial\theta_p}\geq 0$ and $\frac{\partial^2 W_{x<c}}{\partial b\partial\theta_p}\geq 0$ and we will show that the composition of both functions inherits this single-crossing property. From equation (4) it follows that,

$$\frac{\partial^2 W_{x\geq c}}{\partial b \partial \theta_p} = (1-\alpha)\left\{-\frac{\partial \theta_s}{\partial b}\left[\frac{\partial B_i}{\partial \theta_i}\phi(\theta_s-\alpha\theta_p)-(L_p+L_i)\,\phi'(\theta_s-\alpha\theta_p)\right] + \int_{\theta_s}^{\infty}\frac{\partial^2 B_i}{\partial b \partial \theta_i}\phi(\theta_i-\alpha\theta_p)d\theta_i\right\}$$

This expression is positive, given the assumptions and *Lemma 1*. To show that this is also true when $x < c$, we use equation (5), obtaining:

$$\frac{\partial^2 W_{x<c}}{\partial b \partial \theta_p} = (1-\alpha)\left\{\left[\pi_p(-\theta_p)-\pi(\theta_0-\theta_p)+c+(L_p+L_i)\right]\phi'(\theta_0-\alpha\theta_p) - \int_{\theta_0}^{\infty}\frac{\partial B_i}{\partial b}\phi'(\theta_i-\alpha\theta_p)d\theta_i\right\} - \pi_p'(-\theta_p)\phi(\theta_0-\alpha\theta_p)$$

Clearly, since the last term is positive, the whole expression will also be positive for α close enough to 1.

As we know, second order properties are not directly inherited by the use of the maximum operator. However, as can be seen in *Figure 4*, a sufficient condition for $\frac{\partial^2 W}{\partial b \partial \theta_p} \geq 0$, is $\frac{\partial W_{x<c}}{\partial b} \geq \frac{\partial W_{x\geq c}}{\partial b}$ and $\theta^*(b, L_i, L_p)$ being decreasing in b. As has been proven in *Lemma 4* the first is always true. For the second, *Lemma 4* also shows that the crossing point of $W_{x<c}$ and $W_{x\geq c}$, $\theta^*(b, L_i, L_p)$ is decreasing in b for α sufficiently large.

It is clear from the definition of V_p that for θ other than θ^*,

$$\frac{\partial V_p}{\partial b} = \lambda\frac{\partial W}{\partial b}$$

and similarly,

$$\frac{\partial^2 V_p}{\partial b \partial \theta_p} = \lambda\frac{\partial^2 W}{\partial b \partial \theta_p} \geq 0$$

Therefore, the condition necessary for a mechanism to be implementable is $b'(\theta_p) \geq 0$. Obviously, this implies that $f'(\theta_p) \geq 0$.

B Appendix

In the structure of the game so far we have assumed that once the infringer rejects the offer made, x, the patentholder has no other choice but going to court to solve the matter. However, at first sight, this action does not need to be optimal, and in particular, the patentee might prefer to accommodate the entrant. The purpose of this section is to provide a set of conditions under which this will never be true. We will show that the equilibrium of such a game would in general be equivalent to the one considered in the rest of the paper[21]. Therefore, there is no much loss of generality in assuming that the patentee will always go to court when the offer is rejected.

First, we need to define the payoffs for both players in case they do not go to court. It is natural to assume that they will be $\pi_p(\theta_i - \theta_p)$ for the patentholder, and $\pi_i(\theta_i - \theta_p) - c$ for the infringer, since in this case the existence of patents is vacuous.

We will next use a concept defined by Nalebuff (1987) and adapted to this environment:

Definition 2 *A case has merit if the patentholder's expected value of litigation is higher than the value of not going to court, given the prior distribution of infringers.*

This condition implies that the unconditional value of going to court is higher than the value of not exercising this right.

$$\int_{\alpha\theta_p}^{\infty} \left[\pi(\theta_i - \theta_p) - L_p - B_i(\theta_i - \theta_p, b)\right] \phi(\theta_i - \alpha\theta_p) d\theta_i \geq \int_{\alpha\theta_p}^{\infty} \pi_p(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p) d\theta_i$$

If we define the function $J(\theta, \theta_p)$ as,

$$J(\theta, \theta_p) = \int_{\theta}^{\infty} \left[(1 - q(\theta_i - \theta_p, b))\pi_i(\theta_i - \theta_p) - L_p\right] \phi(\theta_i - \alpha\theta_p) d\theta_i \tag{11}$$

a case has merit if

$$J(\alpha\theta_p, \theta_p) \geq 0 \tag{12}$$

In this new game, the patentholder has to choose for each level of θ_p not only a settlement offer x, but also a potentially mixed strategy, in case the offer is rejected, between going to court or not. We define the function $s(\theta_p, \theta)$ as the probability that a patentee with an invention of size θ_p will go to court when the worse infringer that decides to reject the settlement has quality θ (in particular θ_s if there is settlement, and θ_0 otherwise). The patentee will go to court if

$$\int_{\theta}^{\infty} \left[\pi(\theta_i - \theta_p) - L_p - B_i(\theta_i - \theta_p, b)\right] \phi(\theta_i - \alpha\theta_p) d\theta_i \geq \int_{\alpha\theta}^{\infty} \pi_p(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p) d\theta_i$$

or, using equation (11), if $J(\theta, \theta_p) \geq 0$.

Therefore, a patentee behaving optimally will choose:

$$s(\theta_p, \theta) \begin{cases} = 1 & if \quad J(\theta, \theta_p) > 0 \\ \in [0,1] & if \quad J(\theta, \theta_p) = 0 \\ = 0 & if \quad J(\theta, \theta_p) < 0 \end{cases}$$

We make the following assumption:

Assumption 7 *For every θ_p there is a unique value $\widehat{\theta}$ such that $J(\widehat{\theta}, \theta_p) > 0$ if $\theta < \widehat{\theta}$ and $J(\widehat{\theta}, \theta_p) < 0$ if $\theta > \widehat{\theta}$.*

[21] This part is a generalization of Nalebuff (1987) from which we borrow some of the concepts used.

If a case has merit and Assumption 7 holds, it is immediate that, the optimal strategy for a patentholder is:

$$s(\theta_p, \theta) \begin{cases} = 1 & if \quad \theta < \widehat{\theta} \\ \in [0,1] & if \quad \theta = \widehat{\theta} \\ = 0 & if \quad \theta > \widehat{\theta} \end{cases} \tag{13}$$

Two values of θ are relevant: θ_s when $x \geq c$ and θ_0 when $x < c$. The next lemma will show that in all sequential equilibria of this enhanced game, whenever $x \geq c$, litigating will be preferred to not exercising this right.

Lemma 12 *If a case has merit and (A.7) is satisfied, there is no sequential equilibrium in which $x \geq c$ and litigation is not used.*

Proof. Suppose towards a contradiction that this is the case. Therefore, there exist some θ_p for which $s(\theta_p, \theta_s) = 0$. The worse inventor not accepting the settlement, θ_s, will be such that,

$$\pi_i(\theta_s - \theta_p) = x$$

The optimal level of x will be obtained by maximizing the following function:

$$W_{x \geq c}(\theta_p) = \max \int_{\alpha\theta_p}^{\theta_s} \left[\pi(\theta_i - \theta_p) - x\right] \phi(\theta_i - \alpha\theta_p) d\theta_i + \int_{\theta_s}^{\infty} \pi_p(\theta_i - \theta_p) \phi(\theta_i - \alpha\theta_p) d\theta_i$$

With First Order Condition:

$$\frac{\partial W}{\partial x} = -\Phi(\theta_s - \alpha\theta_p) < 0$$

Therefore, the solution to this problem is $x = c$. We will now show that this choice is dominated by offering $x < c$ and not litigating any infringer.

Because infringers do not face any litigation when they enter, the level θ_0 will be defined as,

$$\pi_i(\theta_0 - \theta_p) = c$$

And therefore, $\theta_0 = \theta_s$. However, profits are in this case,

$$W_{x<c}(\theta_p) = \pi_p(-\theta_p)\phi(\theta_s - \alpha\theta_p) + \int_{\theta_s}^{\infty} \pi_p(\theta_i - \theta_p)\phi(\theta_i - \alpha\theta_p) d\theta_i$$

And so,

$$W_{x \geq c}(\theta_p) - W_{x<c}(\theta_p) = \int_{\alpha\theta_p}^{\theta_s} \left[\pi(\theta_i - \theta_p) - c - \pi_p(-\theta_p)\right] \phi(\theta_i - \alpha\theta_p) d\theta_i \leq 0$$

where the last inequality comes from (A.3).Thus, settlement was not optimal in the first place. ■

The previous lemma shows that an acceptable settlement and litigation go together. Hence, there is no loss of generality in assuming that when an offer $x \geq c$ is made and rejected, litigation is the only possible choice. When $x < c$ different cases might arise.

Call, to be consistent with the rest of the paper, the worse entrant θ_0. If $\theta_0 < \widehat{\theta}$ the patentee will litigate all infringements, and therefore, the level of θ_0, that we will call θ_0^L, will be defined as usual - see equation (3)-:

$$B_i(\theta_0^L - \theta_p, b) = c$$

If, on the other hand, $\theta_0 > \widehat{\theta}$, the level of θ_0 is defined as θ_0^N, where

$$\pi_i(\theta_0^N - \theta_p) = c$$

Two important consequences of the previous definitions are that (i) $\theta_0^L \geq \theta_0^N$ and (ii) θ_0^N and θ_0^L are increasing in θ_p at a one to one rate.

According to those definitions, for every level of θ_p three cases might arise:

Case 1: $\widehat{\theta} \leq \theta_0^N \leq \theta_0^L$

In this case the patentee will never be able to credibly commit to litigate the infringers. Therefore, in the corresponding sequential equilibrium, all innovations with size larger than θ_0^N will be implemented, and litigation will not be used.

Case 2: $\theta_0^N \leq \theta_0^L \leq \widehat{\theta}$

The patentholder can credibly commit to fight any infringement. Therefore, in the sequential equilibrium of the game, only infringers above θ_0^L will enter.

Case 3: $\theta_0^N \leq \widehat{\theta} \leq \theta_0^L$

In this case, in the only sequential equilibrium of the game, the patentee will randomize between litigating and not. That is, $0 < s(\theta_p, \theta) < 1$.

The next lemma gives a sufficient condition for Case 2 - litigating any infringer - to arise:

Lemma 13 *If a case has merit and (A.7) holds, there is a sequential equilibrium of this game in which the patentee will choose to litigate the infringer if,*

$$B_i(\widehat{\theta}(\theta_{\min}) - \theta_{\min}, b) \geq c$$

Proof. $\widehat{\theta}(\theta_p)$ is defined as $J(\widehat{\theta}, \theta_p) = 0$. By a standard use of the implicit function theorem, we obtain that,

$$\frac{\partial \widehat{\theta}}{\partial \theta_p} = \frac{\pi_i(\widehat{\theta} - \theta_p) - B_i(\widehat{\theta} - \theta_p, b)}{\pi_i(\widehat{\theta} - \theta_p) - B_i(\widehat{\theta} - \theta_p, b) - L_p} \geq 1$$

where the denominator is positive because $J(\widehat{\theta}, \theta_p) = 0$, and $J(\theta, \theta_p) < 0$ for all $\theta > \widehat{\theta}$.

Therefore, we only need $\theta_0^L(\theta_{\min}) \leq \widehat{\theta}(\theta_{\min})$, because this would imply $\theta_0^L(\theta_p) \leq \widehat{\theta}(\theta_p)$ for all $\theta_p > \theta_{\min}$. By definition of θ_0^L, this means that,

$$B_i(\widehat{\theta}(\theta_{\min}) - \theta_{\min}, b) \geq B_i(\theta_0^L(\theta_{\min}) - \theta_{\min}, b) = c$$

■

References

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

[2] Cornelli, F. and M. Schankerman (1996) 'Patent Renewals and R&D Incentives', The Economics of Industry Group Discussion Paper EI/13. STICERD. London School of Economics.

[3] Gallini, N. T. (1992) 'Patent Policy and Costly Imitation', RAND Journal of Economics, **23**: 52-63.

[4] Gilbert, R. and C. Shapiro (1990) 'Optimal Patent Length and Breadth', *RAND Journal of Economics*, **23**: 52-63.

[5] Green, J. R. and S. Scotchmer (1995) 'On The Division of Profit in Sequential Innovation', RAND Journal of Economics, **26**: 20-33.

[6] Kintner, E.W. and J.L. Lahr (1975) 'An Intellectual Property Law Primer', Macmillan Publishing Co., Inc.

[7] Klemperer, P. (1990) 'How Broad Should the Scope of Patent Protection Be?', *RAND Journal of Economics*, **21**: 113-130.

[8] Lanjouw, J. (1993) 'Patent Protection: Of What Value and for How Long?', NBER working paper No. 4475.

[9] Lanjouw, J. (1994) ' Economic Consequences of a Changing Litigation Environment: The case of Patents', NBER working paper No. 4835.

[10] Lanjouw, J. and J. Lerner (1996) 'Preliminary Injunctive Relief: Theory and Evidence from Patent Litigation', NBER working paper No. 5689.

[11] Lanjouw, J. and J. Lerner (1997) 'The enforcement of Intellectual Property Rights: A Survey of the Empirical Literature', NBER working paper No.6296.

[12] Lanjouw, J. and M. Schankerman (1997) 'Stylized Facts of Patent Litigation: Value, Scope and Ownership', NBER working paper No. 6297.

[13] Lanjouw, J., A. Pakes and J. Putnam (1996) 'How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data', NBER working paper No. 5741.

[14] Lerner, J. (1994) 'The importance of patent scope: an empirical analysis', RAND Journal of Economics, **25**: 319-333.

[15] Meurer, M. (1989) 'The Settlement of Patent Litigation', *RAND Journal of Economics*, **20**: 77-91.

[16] Nalebuff, B. (1987) 'Credible Pretrial Negotiation', *RAND Journal of Economics*, **18**: 198-210.

[17] O'Donoghue T. S. Scotchmer and J. T. Thisse (1995) 'Patent Breadth, Patent Life and the Pace of Technological Progress', Department of Economics Working Paper No. 65-242, IBER University of California, Berkeley.

[18] Scotchmer, S. (1997) 'On the Optimality of the Patent System', GSPP Working Paper 236, University of California, Berkeley.

[19] Waldfogel, J. (1993) 'The Selection Hypothesis and the Relationship Between Trial and Plaintiff Victory', NBER working paper No. 4508.

[20] Waterson, M. (1990) 'The Economics of Product Patents', *American Economic Review*, **80**: 860-869.