Optimal patent policy with externalities

Sylvain Hours (Ph.D) - University of Montpellier

Abstract :

In this paper, we determine the optimal patent policy when the innovation generates a negative externality. In that case, the patent breadth not only influences the deadweight loss induced by the industrial property, but also the size of the external effect. We provide several interpretations of the concept of breadth and we compute the optimal patent structure. Besides, we address the question of how much to reward innovators and we show that a modulated patent policy giving stronger protection to those who have implemented safe Research & Development (R&D) strategies might be an instrument with the potential to prevent the emergence of harmful innovations.

1 Introduction

When free markets are unable to convey sufficient incentives to invest in R&D, industrial property - and the use of patents in particular - may play a significant role in stimulating innovative activities. A patent is an industrial property title that gives a patentee the exclusive and temporary right to exploit her findings commercially. The market power which is awarded to her allows the patentee to make supernormal profits and to recoup R&D investments but it also generates deadweight loss. Hence, the question of how much protection to give innovators should weigh the benefit associated with a higher rate of innovation against the cost of precluding competition.

Nordhaus (1967, 1972) and Scherer (1972) were among the first economists to address this issue and to compute the optimal *length* of a patent, that is, the duration for which innovators should benefit from industrial property. Thereafter, Gilbert & Shapiro (1990) and Klemperer (1990) argued that the value of a patent is not only determined by its length but also by its *breadth*.[1] The economic literature has interpreted the concept of breadth in various ways. As summarized by Denicolò (1996), it can be alternatively understood as (i) the cost of inventing around the patent (Gallini, 1992), (ii) the distance in the space product that must be traveled between the innovation and non-infringing varieties (Klemperer, 1990), (iii) the fraction of the optimal royalty fee that the innovator is allowed to charge by the public agency (Tandon, 1982) or (iv) the fraction of the cost reduction induced by a process innovation that does not spill out and becomes freely available to non-innovating firms (Nordhaus, 1972). Generally speaking, the patent breadth measures the patentee's ability to take advantage of her dominant position and make supernormal profits.

Hence, the design of the optimal patent policy should be guided by two major considerations: first, the reward that is given to innovators (i.e. the value of the patent) for the policy to convey the appropriate amount of incentives to innovate and second, the way in which that reward is allocated over time (i.e. the structure of the patent) for the industrial property to induce the least overall social costs. Indeed, although the patentee is indifferent between all the industrial protection titles that have the same value, the trade-off between the patent length and breadth is not necessarily neutral in terms of welfare.

In this chapter, we assume that the innovation generates a negative externality[2]

[1] Also referred to as *width* or *scope* in the literature.

[2] Throughout this chapter, we focus on negative externalities since our primary interest is the regulation of potentially harmful innovations. Yet, it would be straightforward to extend our

and we discuss the way it influences the design of the optimal patent policy. When it is confined to giving firms incentives to innovate, industrial property is nothing else than a necessary evil. Our main point is that it can also prove to be a powerful instrument for preventing the emergence of negative externalities and for reducing the damage that they may cause.

In particular, if the magnitude of the externality is determined by the innovator's R&D strategy, then it is possible to imagine a patent policy in which firms that have implemented safe R&D strategies are given a stronger protection than those who have not. In that case, the public agency faces an additional trade-off between the decrease in the innovation's harmfulness and the extra deadweight loss.

Besides, while most authors assume that the social costs induced by the industrial property increase with the patent breadth, we argue that in the presence of a negative external effect, they may as well be decreasing. In particular, by giving market power to the patentee, industrial property may lower the equilibrium output and reduce the damage caused by the external effect more effectively. Therefore, the trade-off underlying the design of the optimal patent structure should take into consideration the impact that breadth has on both the deadweight loss and the size of the external effect.

Note that the economic literature has long underlined that imperfectly competitive markets may actually reduce the size of negative external effects. In particular, Buchanan (1969), Baumol & Oates (1975) and Barnett (1980) stress that the market structure for the good or service that generates the externality is a key element in clarifying the issue of taxation for control of external effects. Also, it should be noted that the idea that industrial property can be used as a tool

analysis to a more general class of externalities including those said to be positive.

to correct the market failure associated with external effects has been already put forward: Laxminarayan (2002) addresses the issue of antimicrobial resistance. The author explains that "*antibiotic use creates a negative externality because antibiotic use by one patient may generate resistant bacteria that can infect others*" and he stresses that pharmaceutical companies selling antibiotics do not fully internalize the social cost associated with the depletion of their effectiveness. Likewise, when doctors prescribe antibiotics, they are likely to focus on the short-term benefits to the patient and to overlook the long-term risks to society. The main argument is that stronger industrial property rights may reduce antibiotic use and prevent the emergence of resistant bacteria. Laxminarayan shows that increasing patent breadth - as measured by the number of competing firms within the same class of antibiotics - can be beneficial to society. In that case, the cost of greater monopoly power today is outweighed by the benefit of preserving antibiotic effectiveness tomorrow. Horowitz & Moehring (2004) bring forward another argument in favour of broad patents: when cross-resistance may occur,[3] extending the patent breadth to the whole class of antibiotics could strengthen the incentive for the patentee to internalize this external effect. Finally, the *Office of Technology Assessment* has suggested that a longer patent length could increase the incentives for pharmaceutical companies to contain resistance, since they would enjoy a longer period of monopoly benefits from its antibiotic's effectiveness.

This chapter is organized as follows: in Section 2, we compute the patent optimal length and breadth when the innovation generates a negative external effect. First, we present the general result established by Gilbert & Shapiro (1990) and we show that narrow and infinitely-lived patents may not be optimal even though

[3] Cross-resistance is the resistance to a particular antibiotic that often results in resistance to other antibiotics, usually from a similar chemical class, to which the bacteria may not have been exposed.

breadth is increasingly costly in terms of deadweight loss. Second, we discuss the consequences that the externality may have on the optimal patent structure. Last, we follow Denicolò (1996) and we examine several interpretations of the concept of breadth in order to establish closed-form conditions for the patent's optimal length and breadth. Our conclusions suggest that the optimal patent structure is closely related to the size of the negative external effect and to the extent to which broader patents may make non-infringing imitations less harmful. In Section 3, we compute the optimal value of the patent when the firm's R&D strategy determines the size of the negative external effect. First, we address the case in which the public agency can set a single patent value whatever the firm's R&D strategy. Second, we look into a hypothetical modulated patent policy that gives stronger protection to firms that have undertaken actions aimed at preventive the externality from arising. We establish that the latter policy is strictly more efficient than the former whenever the negative external effect is sufficiently large. In that case, indeed, the public agency effectively takes advantage of the flexibility allowed by the modulated policy and induces firms to undertake preventive actions. Section 4 concludes.

2 Optimal patent length and breadth

2.1 Gilbert & Shapiro (RAND, 1990): the general result

In this seminal article, Richard Gilbert & Carl Shapiro discuss the rule according to which a pre-specified value of the patent V should be allocated over time. The basic trade-off is between broad patents that induce a large deadweight loss for a short period of time and narrow patents that generate a small deadweight loss for a long period of time. Note that the authors leave aside the question of how much to reward innovators so they take the rate of innovation as given and they focus on computing the optimal patent structure. In their framework, T denotes the patent length and π its breadth.[4] The latter is identified to the patentee's flow rate of profits. Once the patent has expired, the firm's flow profits decline to $\overline{\pi}$ (i.e. $\pi \geq \overline{\pi}$). The smaller $\overline{\pi}$, the more competitive the underlying market structure. Hence, the patentee's discounted profit is equal to

$$V(T, \pi) = \int_0^T \pi e^{-rt} dt + \int_T^{+\infty} \overline{\pi} e^{-rt} dt$$

where r is the rate at which future amounts are discounted.[5] The flow social welfare $W(\pi)$ is assumed to be decreasing in breadth (i.e. $W'(\pi) < 0$). The argument is that the broader the patent, the larger the deadweight loss induced by the industrial property. The discounted social welfare is thus equal to

$$\Omega(T, \pi) = \int_0^T W(\pi) e^{-rt} dt + \int_T^{+\infty} W(\overline{\pi}) e^{-rt} dt$$

The problem of the public agency consists in determining the patent length and breadth that maximize the discounted welfare under the constraint that the patentee gets at least V. Formally, the public agency solves

$$\max_{T \geq 0,\ \pi \geq rV} \Omega(T, \pi) \text{ s.t. } V(T, \pi) \geq V$$

The patent breadth must be larger than rV so the desired reward can be achieved through industrial property. Since Ω decreases in both its arguments, it is clear that the constraint binds at the optimum (i.e. $V(T^*, \pi^*) = V$).[6] Therefore, the authors are able to define the function ϕ such that $V(T, \phi(T)) = V$. Differentiating this latter equation with respect to the patent length, they show that

$$\phi'(T) = -r \left[\phi(T) - \overline{\pi}\right] \frac{e^{-rT}}{1 - e^{-rT}} \tag{1}$$

[4] The authors focus on *rectangular patents* whose breadth is constant over time and so will we.

[5] We assume that the private discount rate equals the social discount rate.

[6] Klemperer (1990) solves an analogous problem in which the constraint is directly written as an equality.

Next, they differentiate $\Omega(T, \phi(T))$ with respect to the patent length and they show that the discounted welfare increases in length whenever

$$re^{-rT}\left[W(\phi(T)) - W(\overline{\pi})\right] + (1 - e^{-rT})W'(\phi(T))\phi'(T) \geq 0$$

Finally, they use (1) and rewrite the former condition as follows:

$$\frac{W(\phi(T)) - W(\overline{\pi})}{\phi(T) - \overline{\pi}} \geq W'(\phi(T)) \tag{2}$$

The concavity of W on the interval $[\overline{\pi}, \phi(T)]$ is thus a sufficient condition for inequality (2) to hold. In the framework of Gilbert & Shapiro, the only source of social costs is the deadweight loss $dwl(\pi)$ with $dwl' > 0$ so $W(\pi) = w_c - dwl(\pi)$ where w_c is the perfectly competitive welfare. Clearly, the convexity of dwl implies the concavity of W so the authors state that if breadth is increasingly costly in terms of deadweight loss, then narrow and infinitely lived-patents are optimal (Proposition 1).[7] In that case, indeed, *"increasing the breadth of the patent [...] is increasingly costly, in terms of deadweight loss, as the patentee's market power grows. When increasing the length of the patent, by contrast, there is a constant trade off between the additional reward to the patentee and the increment to deadweight loss [...]"*. It should be noted that the variations of the flow social welfare are irrelevant to determine the optimal patent structure, only its curvature matters. However, their Proposition would no longer hold if non-linearity were to be introduced in the relation between breadth and the patentee's flow profits so the scope of their results is limited to specific market structures.[8]

[7] Likewise, the concavity of dwl is a sufficient condition for wide and short patents to be optimal. Finally, if dwl is linear, then the mix between length and breadth has no impact on welfare.

[8] On this point, the framework provided by Klemperer (1990) is more general since no restriction is placed on the relation between breadth and the patentee's flow profits. In that case, the optimal patent structure is not merely driven by the curvature of the deadweight loss but by

2.2 Introducing a negative externality

When the innovation generates a negative external effect, the design of the optimal patent structure should not only be driven by the impact of breadth on the deadweight loss but also by its influence on the damage generated by the externality. In that case, breadth is no longer a pure cost for society, as was the case in the framework of Gilbert & Shapiro, since it becomes an instrument with the potential to address the externality. As we shall see, if an increase in breadth leads to a reduction in damage which is larger than the additional induced deadweight loss, then giving the patentee a greater market power might actually be welfare improving.

The idea that broader patents may reduce the negative externality can be explained in two ways: first, broader patents may lessen competition in the innovation market, decrease the equilibrium output and thus reduce the magnitude of the externality. Second, if the concept of breadth is interpreted as the distance that must be traveled, in some product space, away from the innovation to produce a non-infringing variety, then broader patents may make imitations less harmful (i.e. reduce the extent to which they generate a negative externality).

In the presence of a negative external effect, note that the constraint according to which the innovator must be given at least a pre-specified reward might not bind at the optimum. Indeed, the public agency may be better off giving stronger protection, at the cost of a larger deadweight loss, in order to reduce the external effect more effectively. For now, we leave this issue aside and we set $V(T^*, \pi^*) = V$ so our primary interest is to determine the socially efficient way to structure a

---

the *breadth elasticity* of both the profit and the deadweight loss. However, his work may also be considered more restrictive because Klemperer adopts a specific interpretation of the concept of breadth (i.e. the distance that must be traveled, in a vertically differentiated product space, between the variety produced by the patent holder and non-infringing imitations).

patent of given value. We shall discuss how much innovators should be rewarded in a subsequent section.

When introducing a negative external effect into Gilbert & Shapiro framework, the flow social welfare is given by

$$W(\pi) = w_c - dwl(\pi) - d(\pi)$$

where $d(\pi)$ is the damage generated by the externality. Recall that the concavity of W is a sufficient condition for narrow and infinitely-lived patents to be optimal. Yet, in the presence of the external effect, W'' is not only determined by the curvature of dwl but by that of the *total social costs* $dwl + d$. In particular, W and dwl can be both convex when d is sufficiently concave. Therefore, narrow and infinitely-lived patents may not be optimal even though breadth is increasingly costly in terms of deadweight loss. As outlined above, since the variations of the flow social welfare are irrelevant to determine the optimal patent structure, we are able to establish what we refer to as the *extended Gilbert & Shapiro condition*.

Proposition 1 *When introducing a negative externality into the framework of Gilbert & Shapiro, if the total social costs are convex in breadth, then narrow and infinitely-lived patents are optimal.*

Pay attention to the fact that this *extended Gilbert & Shapiro condition* only applies when the patentee's flow profit grows linearly with the patent breadth. Clearly, under the assumption that breadth is increasingly costly in terms of deadweight loss, the introduction of a negative external effect does not challenge the optimality of narrow and infinitely-lived patents whenever the damage it induces is not too concave in breadth.

Throughout this chapter, we assume that the demand for the innovation is constant over time. That is, we disregard the innovation diffusion process. Bass

(1969) provides a theoretical framework in which the probability that a potential buyer purchases the innovation at time t is linearly related to the cumulative number of buyers. If broader patents reduce the equilibrium output in the innovation market, then industrial property may slow the diffusion process down. When the innovation does not generate any negative external effect, this reduction in speed of diffusion entails an additional social cost. However, in the presence of a negative external effect, a slower diffusion process may actually curtail the damage and raise the discounted social welfare. The diffusion of harmful innovations is indeed analogous to the spread of an epidemic and the patent breadth can be assimilated to the extent to which the disease in question is contagious. The benefit of slowing down the innovation diffusion process is even greater if society acquires information about the external effect as time goes by or when the damage induced by the externality is irreversible.

Finally, on a more speculative note, if previous consumption experiences improve the understanding that society has of the innovation and if the industrial property combined with the absence of economies of scale and learning-by-doing productivity gains at the early stages of the diffusion process make the innovation pricy at the time it is first introduced, then we can ourselves whether or not the richest individuals or countries serve as Guinea pigs.

2.3 Examples

In this section, we follow Denicolò (1996) and we provide several interpretations of the concept of breadth. In each case, we describe the market structure and we compute the optimal structure of the patent

2.3.1 Compulsory licensing

To start with, we assume that there is compulsory licensing of a cost reducing innovation as in Tandon (1982). The royalty fee is determined by the public agency and it is understood as the patent breadth. Though this might not be the most intuitive interpretation of the concept of breadth, it allows us to make our point in the clearest possible way.

Consider the market inverse demand $P = a - Q$ for a given good where P is the price at which it is sold and Q is the output. The cost reducing innovation allows the innovator to produce this good at constant unit cost $k < a$. If she were unregulated, and since we assume free-entry, then the innovator would set the royalty fee at $\frac{s}{2}$ where $s = a - k$ so the innovation would be sold (competitively) at monopoly price.[9]

Let $w \in [0, 1]$ be the fraction of that optimal royalty fee that the patentee is allowed to charge by the public agency. We assimilate w to the patent breadth: the smaller w, the less able the innovator to take advantage of her better production efficiency and to make supernormal profits. If $w = 0$, then the market is perfectly competitive whereas it is *equivalent* to a monopoly if $w = 1$.[10]

We depart from the standard framework by assuming that the process innovation generates a production negative externality which induces a damage $d(Q) = \frac{1}{2}\gamma Q^2$ where $\gamma \geq 0$ measures the *size* of the externality. The good is sold at price $p(w) = k + w\frac{s}{2}$ so the equilibrium output is $Q(w) = \frac{s}{2}(2 - w)$. The

[9] When the innovator charges a royalty fee ρ, the equilibrium price and output are, respectively, $k + \rho$ and $s - \rho$. In that case, all licensees make zero profit while the innovator gets $\rho(s - \rho)$ which is maximum at $\rho = \frac{s}{2}$. Hence, the equilibrium price is $k + \frac{s}{2}$ which coincides with the monopoly price.

[10] By *equivalent*, we mean that the innovator behaves as a monopolist when she is unregulated.

innovator's profit is equal to

$$\pi(w) = \frac{s^2}{4}w(2-w)$$

The consumer surplus is equal to

$$CS(w) = \frac{s^2}{8}(2-w)^2$$

the damage associated with the negative externality is equal to

$$d(Q(w)) = \frac{s^2}{8}\gamma(2-w)^2$$

and the social welfare is equal to

$$S(w) = CS(w) + \pi(w) - d(Q(w)) = \frac{1}{8}s^2(2-w)\left[2(1-\gamma) + w(1+\gamma)\right]$$

In this framework, breadth is increasingly costly in terms of deadweight loss.[11] Yet, we cannot conclude which patent structure is the most efficient because the patentee's profit does not grow linearly with the patent breadth and even if it were the case, the existence of a negative external effect is such that the concavity of the flow social welfare is no longer a sufficient condition for narrow and infinitely-lived patents to be optimal.

As mentioned earlier, breadth should not only be seen as a cost for society. Indeed, by reducing the equilibrium output (i.e. $Q'(w) < 0$), broader patents decrease (at a decreasing rate) the damage generated by the negative externality (i.e. $\frac{\partial d(Q(w))}{\partial w} \leq 0$ and $\frac{\partial^2 d(Q(w))}{\partial w^2} \geq 0$). As depicted in Figure 1, if the externality is large (i.e. if $\gamma \geq 1$), then we show that the marginal benefit of a lower damage is always larger than the marginal deadweight loss so increasing the patent breadth is always welfare improving. However, if the externality is small (i.e. if $\gamma \leq 1$), then the social welfare has an inverted U shape: for narrow patents, a marginal

[11] Indeed, $dwl(w) = CS(0) - CS(w) - \pi(w) = \frac{1}{8}w^2s^2$

increase in breadth is such that the additional deadweight loss is smaller than the damage reduction so the welfare is locally increasing. The effect reverses for broader patents.[12] In the absence of any negative external effect (i.e. if $\gamma = 0$), note that the welfare unambiguously decreases with the patent breadth.

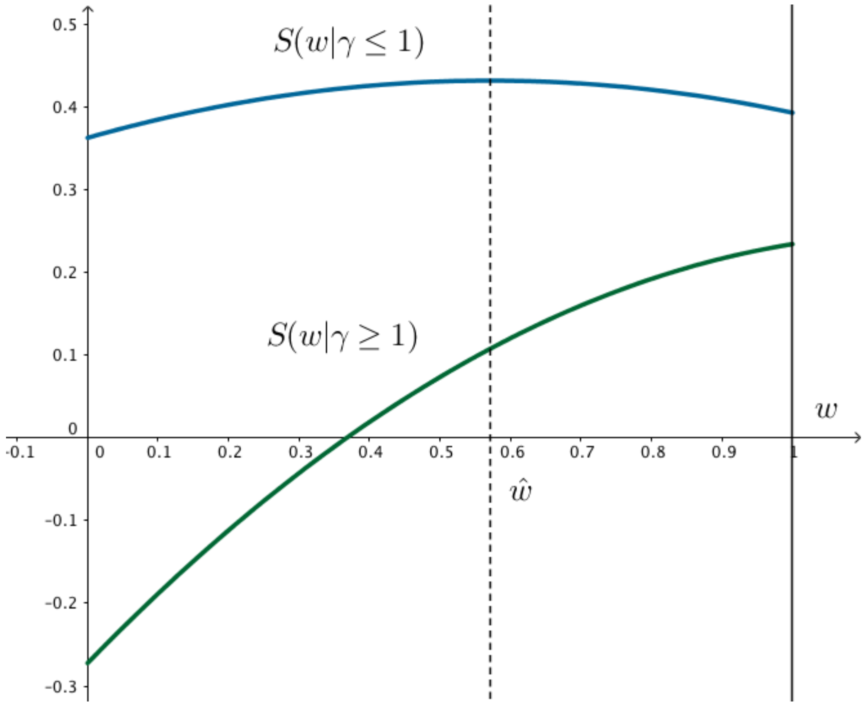

Figure 1: Welfare and size of the negative external effect

Let us now determine the socially efficient way to structure a patent whose value V is given. That is, we compute the patent length $T \geq 0$ and breadth $w \in [0, 1]$ that minimize the discounted social costs for a given rate of technological change. As outlined earlier, when there is no external effect, the basic trade-off is between a smaller deadweight loss for a longer period of time and a larger deadweight loss

[12]The marginal deadweight loss is equal to the marginal damage reduction at $\widehat{w} = \frac{2\gamma}{1+\gamma}$. If $\gamma \geq 1$, then $\widehat{w} > 1$ so breadth unambiguously increases welfare. However, if $\gamma \leq 1$, then $\widehat{w} \in [0, 1]$ so the flow social welfare is non-monotonic with respect to the patent breadth.

for a shorter period of time. In the presence of a negative externality, the trade-off incorporates an additional dimension: long and narrow patents induce a smaller damage reduction for a longer period of time whereas broad and short patents allow a larger damage reduction for a shorter period of time.

Future amounts are discounted at rate r. The public agency chooses the patent structure (T, w) that maximizes

$$\Omega(T, w) = \int_0^T e^{-rt} S(w) dt + \int_T^{+\infty} e^{-rt} S(0) dt$$

under the constraint that

$$V(T, w) = \int_0^T e^{-rt} \pi(w) dt = V$$

Note that V must be smaller than $\overline{V} = V(\infty, 1)$, the maximum reward achievable through industrial property. The constraint $V(T, w) = V$ implies that the patent length and breadth are linked by the relation

$$T = \psi(w) = -\frac{1}{r} \ln(1 - \frac{4rV}{ws^2(2-w)})$$

The discounted welfare is $\Omega(\psi(w), w)$. Differentiating with respect to w, we have

$$\frac{d\Omega}{dw} = \frac{V(\gamma - 1)}{(w-2)^2}$$

so the discounted welfare increases with the patent breadth whenever $\gamma \geq 1$.

Proposition 2 *In the case of compulsory licensing of a cost reducing innovation that generates a negative externality, with linear demand and constant marginal costs, if the size of the negative externality is small (i.e. if $\gamma \leq 1$), then narrow and infinitely-lived patents are optimal.*

Clearly, broad and short patents are optimal if $\gamma \geq 1$ and the mix between length and breadth has no impact on welfare if $\gamma = 1$. In the absence of any negative external effect, Tandon (1982) showed that infinitely-lived patents are optimal. By setting $\gamma = 0$, we come to the same conclusion. This Proposition makes it clear that the optimal patent structure is closely related to the size of the external effect. If the public agency were to overlook the impact that breadth has on the externality, then the rule according to which he would choose to allocate the value of the patent over time might be inefficient.[13]

2.3.2 Costly imitation

Gallini (1992) and Wright (1999) interpret the concept of breadth as the (fixed) cost of *inventing around* the patent: the broader the patent, the more costly it is to develop and to introduce a non-infringing imitation. The authors assume perfect imitation and free-entry into the market for imitations. The inverse demand for the innovation is given by $P = a - Q$ where P is the price at which it is sold and Q is the aggregate output.[14] The innovation is produced at unit cost $k < a$. Finally, the innovation induces a negative external effect which generates a damage $d(Q) = \frac{1}{2}\gamma Q^2$ with $\gamma \geq 0$.

Assume that the original innovator and n perfect imitators compete *à la Cournot*. The symmetric Nash equilibrium is such that each of them produces $q(n) = \frac{s}{2+n}$ with $s = a - k$ so the original innovator gets $\pi(n) = \frac{s^2}{(2+n)^2}$ whereas each imitator gets $\pi(n) - h$ where h is the fixed cost of inventing around the patent. Clearly, the free-entry assumption implies that $n^* = \pi^{-1}(h)$ so all imitators make zero profit and the original innovator gets h. We follow Denicolò (1996) and we assume that the imitation cost is a fraction of the monopoly profit. Namely, we set $h = w\pi(0)$

[13] In our framework, such an inefficient decision would be made whenever $\gamma \geq 1$. In that case, the minimization of the discounted deadweight loss calls for infinitely-lived patents whereas wide are short patents are actually socially efficient.

[14] The aggregate output includes the original innovation as well as its imitations since they are assumed to be perfect substitutes.

where $w \in [0,1]$ is interpreted as the patent breadth. If $w = 1$, then entry is totally blockaded (i.e. $n^* = 0$). If $w = 0$, then imitation is free so the market is perfectly competitive (i.e. $n^* = +\infty$) and all firms make zero profit. Generally speaking, the broader the patent, the more concentrated the market (i.e. $(n^*)'(w) < 0$) so the larger the supernormal profit that the original innovator can make. The entry process stops at

$$n^* = \frac{2(1-\sqrt{w})^2}{w}$$

Each firm produces $q(w) = \frac{s}{2}\sqrt{w}$ so the aggregate output is $Q(w) = \frac{s}{2}(2-\sqrt{w})$. Although the individual output increases with the patent breadth (i.e. $q'(w) > 0$), the aggregate output decreases (i.e. $Q'(w) < 0$) so broader patents do reduce the damage generated by the externality (at a decreasing rate) (i.e. $\frac{\partial d(Q(w))}{\partial w} \leq 0$ and $\frac{\partial^2 d(Q(w))}{\partial w^2} \geq 0$). The consumer surplus is equal to

$$CS(w) = \frac{1}{8}s^2(2-\sqrt{w})^2$$

the damage associated with the negative externality is equal to

$$d(Q(w)) = \frac{1}{8}s^2\gamma(2-\sqrt{w})^2$$

and the flow social welfare is equal to

$$S(w) = h + CS(w) - d(Q(w)) = \frac{1}{8}s^2\left[(3-\gamma)w - 4(\gamma-1)(1-\sqrt{w})\right]$$

Both the marginal damage reduction and the marginal deadweight loss decrease with the patent breadth. If $\gamma \geq 1$, the former is always larger than the latter so the flow social welfare monotonically increases in breadth. If $\gamma \leq 1$, however, the marginal damage reduction outweighs the marginal deadweight for broad patents only, so the social welfare has a U shape. An interesting feature of this model is that even though there is no negative external effect, the social welfare may be increasing in breadth. Indeed, when $\gamma = 0$, we have $S'(w) \geq 0 \Leftrightarrow w \geq \frac{4}{9}$. The reason is that broader patents reduce the number of imitators and thus decrease the total imitation costs which are another type of social cost. However, for all $w \geq \frac{4}{9}$, the imitation cost is larger than the duopoly profit (i.e. $w\pi(0) \geq \pi(1)$) so no entry effectively occurs when the number of entrants is a discrete rather than a continuous variable.

Since the patentee's profit h grows linearly with the patent breadth, the *extended Gilbert & Shapiro condition* applies. In other words, studying the concavity of the flow social welfare is sufficient to establish the optimal patent structure. We show that

$$S''(w) = \frac{1}{8}s^2(1-\gamma)w^{\frac{2}{3}}$$

so the flow social welfare is concave if and only if $\gamma \geq 1$.

Proposition 3 *In a market with linear demand, constant marginal costs and a fixed imitation cost with free-entry, if the size of the negative externality is large (i.e. if $\gamma \geq 1$), then narrow and infinitely-lived patents are optimal.*

Clearly, this Proposition is diametrically opposed to that we established in the case of compulsory licensing.[15] This underscores how important the interpretation of the concept of breadth is when describing the socially efficient way to allocate a given value of the patent over time. Besides, as was already the case in the previous example, we observe that the size of the externality is the key determinant of the optimal patent structure.

2.3.3 Vertical differentiation

In the case of compulsory licensing, the original innovator and the licensees use the same cost-reducing technology. In the model *à la* Gallini (1982), imitations are

[15] We suspect the inverted U curve between breadth and the imitation costs to be behind this reversal but the specific reasons are still unclear to us.

perfect so the original innovator and imitators produced the same good. Hence, the two previous examples were such that the damage generated by the externality was related to a single variety of product or process. Yet, if the patent breadth prevents imitators from using or producing the original innovation without infringing the industrial property title, then multiple varieties will be used or produced. In that case, the damage generated by the externality is determined by each variety's output and *relative rate of conveying*. The relative rate of conveying measures the extent to which a given variety generates a negative externality with respect to the original innovation. Thus, the relative rate of conveying of a perfect imitation is equal to 1. We assume that imitations *convey* an increasing share of the external effect as they get closer to the original innovation.[16]

In order to understand better the concept of relative rate of conveying and to explain our latter assumption, let us consider the example of a tobacco company that patents a new additive. The innovation improves the taste of the cigarette but it also makes more addictive so its consumption generates increased negative externalities (e.g. passive smoking, cost of treating smoking-related-diseases, street litter, etc.). If the patent is broadly defined, then imitators cannot capture a large fraction of the taste improvement without infringing the patent so the variety they produce *conveys* smaller negative external effects in the sense that it is less addictive.

Here, we assume that a product innovation allows a quality improvement $\widehat{\theta} > 0$. The patent breadth $w \in [0, 1]$ is interpreted as the fraction of $\widehat{\theta}$ that cannot be captured by imitators without infringing the patents. Clearly, all imitators produce the non-infringing variety that has the highest quality. Therefore, there are only two varieties that compete in the market: the original innovation of

[16]For simplicity, we do not consider the (plausible) case in which imitations may be more harmful than the original innovation.

quality $\overline{\theta}+\widehat{\theta}$ where $\overline{\theta}$ is the quality of the traditional product, and imitations of quality $\overline{\theta}+(1-w)\widehat{\theta}$. Production costs are normalized to zero. We further assume free-entry into the market for imitations so competition drives their price to zero.

We consider a mass-one continuum of consumers. Let $m \in [0,1]$ be their willingness to pay for quality. It is assumed to be distributed with density f and cumulative F. Each consumer buys either one unit of the innovation or one unit of the imitation. In the former case, the consumer obtains utility $m(\overline{\theta}+\widehat{\theta})-p$ where p is the price at which the innovation is sold. In the latter case, the consumer obtains utility $m(\overline{\theta}+(1-w)\widehat{\theta})$. Hence, the consumer $\overline{m}$ which is indifferent between buying one unit of the innovation and one unit of the imitation is such that

$$\overline{m}=\frac{p}{w\widehat{\theta}} \tag{3}$$

Therefore, the demand for the innovation is given by $1-F(\overline{m})$ so the patentee's profit is equal to $\pi(\overline{m})=p(1-F(\overline{m}))$. Using (3):

$$\pi(\overline{m})=\overline{m}w\widehat{\theta}(1-F(\overline{m}))$$

Choosing at which price p to sell the innovation is equivalent to choosing $\overline{m}$. We differentiate π with respect to $\overline{m}$ and we show that the indifferent consumer m^* that maximizes $\pi(\overline{m})$ is independent of w. Therefore, the patentee's optimal profit is given by

$$\pi(w)=m^*w\widehat{\theta}(1-F(m^*))$$

This has two important consequences: first, it implies that the innovator's profit grows linearly with respect to the patent breadth. Hence, the *extended Gilbert & Shapiro* condition applies so studying the concavity of the flow social welfare function is sufficient to determine the optimal patent structure. Second, it implies that the innovator's market share is not affected by the patent breadth.

That is, unlike our two previous examples, broader patents do not reduce the equilibrium output of the original innovation. It might then be thought that industrial property has no influence on the damage generated by the negative externality. Yet, although the patent breadth does not alter the equilibrium output of each variety, it determines the extent to which imitations are harmful with respect to the original innovation.

Let $R(w)$ be the imitation relative rate of conveying for a given patent breadth. As mentioned above, we have $R(0) = 1$ and $R' < 0$. We define

$$Q(R(w)) = F(m^*) + R(w)(1 - F(m^*))$$

as the *weighted sum of output,* that is, the actual amount of output that causes social harm. In particular, the original innovation and imitations are assumed to induce a negative externality that generates a damage given by $d(Q(R(w)))$ with $d' > 0$ and $d'' \geq 0$.[17] Note that Q decreases with the patent breadth so broader patents do reduce the negative externality more effectively because they make imitations less harmful. The consumer surplus is equal to

$$CS(w) = F(m^*)\left[\overline{\theta} + (1-w)\widehat{\theta}\right]E(m|m \leq m^*) + (1-F(m^*))\left[(\overline{\theta} + \widehat{\theta})E(m|m \geq m^*) - m^* w\widehat{\theta}\right]$$

Note that it is linearly related to the patent breadth. Finally, the social welfare is equal to

$$S(w) = \pi(w) + CS(w) - d(Q(R(w)))$$

Therefore, since π and CS are both linear functions, the concavity of the social welfare is determined by that of the damage. Clearly, if R is either linear or convex, then the damage is convex in breadth. When R is concave, however, the curvature of the damage function is undetermined.

[17] In this example, there is no need to assume a specific form for the damage generated by the external effect.

Proposition 4 *In the case of product innovation in a vertically differentiated industry, if the relative rate of conveying is not too concave, then narrow and infinitely-lived patents are optimal.*

Unlike the two previous examples, note that the size of the externality is no longer the key determinant of the optimal patent structure. Here, only the curvature of the relative rate of conveying matters. Unfortunately, economic intuition provides very little guidance on this specific point. Intuitively, we would be more inclined to assume convexity between the patent breadth and the relative rate of conveying so only imitations that are sufficiently close to the original innovation *conveys* this external effect. Indeed, in the case of concavity, even *remote* imitations *convey* a large share of the external effect, which seems less relevant.

2.3.4 Horizontal differentiation

In the previous example, we showed that the patentee's market share was unaffected by the patent breadth so broader patents reduced the damage more effectively only because they made imitations less harmful. On that particular point, the case in which the market is horizontally differentiated is more interesting because the patent breadth impacts the damage through two channels. Namely, an increase in breadth improves the patentee's market share and it reduces the relative rate of conveying of imitations. Hence, it is no longer straightforward to determine whether or not broader patents decrease the damage generated by the externality. Also, this market structure tends to be more delicate to address because the patentee's profit is non-linearly related to the patent breadth and because firms make supernormal profits after the patent expires. In order to deal with those additional complications, we shall assume that both the relative rate of conveying and the damage are linear functions.

We consider a mass-one continuum of consumers whose location x is uniformly distributed on the line $[0,1]$. There are two exogenously located firms: firm 1 is located at 0 and firm 2 is located at 1. They initially produce the same good of quality $\overline{\theta}$. We define t as the unit transport cost. Firm 1 innovates and raises the quality of her good to $\overline{\theta}+\widehat{\theta}$. Again, the patent breadth $w \in [0,1]$ is interpreted as the fraction of the quality improvement that cannot be captured by firm 2 without infringing the patent. Therefore, firm 2 produces a good of quality $\overline{\theta}+(1-w)\widehat{\theta}$, the non-infringing variety of highest quality. Production costs are normalized to zero. We assume that the market is covered so all consumers buy one unit of the innovation or one unit of the imitation. We further assume that the innovation is not drastic so the patentee does not monopolize the market.[18] Finally, let p_1 be the price of the innovation and p_2 be that of the imitation. If a consumer located at x buys one unit of the innovation, then he obtains utility

$$U_1(x) = \overline{\theta}+\widehat{\theta} - p_1 - tx$$

whereas he obtains utility

$$U_2(x) = \overline{\theta}+(1-w)\widehat{\theta} - p_2 - t(1-x)$$

when he purchases one unit of the imitation. The consumer $\overline{x}$ which is indifferent between buying one unit of the innovation and one unit of the imitation is such that

$$\overline{x} = \frac{1}{2} + \frac{w\widehat{\theta}}{2t} - \frac{p_1 - p_2}{2t}$$

Hence, the patentee's profit is equal to $\pi_1(p_1,p_2) = p_1\overline{x}$ whereas the imitator's profit is equal to $\pi_2(p_1,p_2) = p_2(1-\overline{x})$. Firms simultaneously choose at which price to sell their good. It is straightforward to show that there exists a single Nash Equilibrium (NE) such that

$$p_1^*(w) = t + \frac{1}{3}w\widehat{\theta} \text{ and } p_2^*(w) = t - \frac{1}{3}w\widehat{\theta}$$

[18] The market is covered if $\widehat{\theta} \geq 3t - 2\overline{\theta}$ and it is not monopolized by the innovator so long as $\widehat{\theta} \leq 3t$.

which yields

$$\pi_1^*(w) = \frac{1}{18}\frac{(3t + w\widehat{\theta})^2}{t} \text{ and } \pi_2^*(w) = \frac{1}{18}\frac{(3t - w\widehat{\theta})^2}{t}$$

As mentioned above, the patentee's profit is non-linearly related to the patent breadth so the *extended Gilbert & Shapiro condition* does not apply. The equilibrium market share of the patentee is equal to

$$x^*(w) = \frac{1}{2} + \frac{1}{6}\frac{w\widehat{\theta}}{t}$$

Hence, the broader the patent the greater the equilibrium output of the original innovation. The consumer surplus is given by

$$CS(w) = \int_0^{x^*} U_1(x)dx + \int_{\overline{x}^*}^1 U_2(x)dx$$

with $p_1 = p_1^*$ and $p_2 = p_2^*$. Let $R(w)$ be the relative rate of conveying of the imitation. Hence, the weighted sum of output is such that $Q(x^*(w), R(w)) = \overline{x}^*(w) + R(w)\,[1 - x^*(w)]$. To simplify, we assume that $R(w) = 1 - w$ so the relative rate of conveying is linearly related to the patent breadth.[19] Under such assumption,

$$Q(w, x^*(w)) = 1 - w\,[1 - x^*(w)]$$

The original innovation and the imitation induce a negative externality that generates a damage $d(Q(w))$. Again, for simplicity, we assume that

$$d(Q(w, x^*(w))) = \gamma Q(w, x^*(w))$$

where $\gamma \geq 0$ is the constant marginal damage measuring the size of the external effect. Unlike our previous examples, an increase in breadth does not always reduce the damage. Two opposite forces are involved here: on the one hand,

[19] Also, this specification implies that the traditional product does not generate any negative externality.

broader patents make imitations less harmful; on the other hand, they increase the market share of the most harmful variety (i.e. the original innovation). We show that the damage decreases monotonically with the patent breadth whenever the innovation is incremental (i.e. if $\widehat{\theta} \leq \frac{3t}{2}$). In that case, indeed, the quality gap between both varieties is small so the patent breadth has little impact on their respective market shares. However, if the quality improvement is significant, then the damage generated by the externality has a U shape: for patents narrower than $\widehat{w} = \frac{3}{2}\frac{t}{\widehat{\theta}}$, a marginal increase in breadth is such that the effect of a lower imitation's relative rate of conveying outweighs the impact of an increased market share of the original innovation so the damage is locally decreasing. The balance of power reverses for patents broader than $\widehat{w}$.

Finally, the flow social welfare is given by

$$S(w) = \pi_1^*(w) + \pi_2^*(w) + CS(w) - d(Q(w, x^*(w)))$$

Unlike our previous examples, the market is imperfectly competitive even though there is no industrial property. That is, before the patent expires, the deadweight loss has two components: one part is induced by the industrial property while another is generated by the underlying market structure.

Now that we have described the market equilibrium for a given patent breadth, we turn to computing the socially efficient way of structuring a patent whose value V is pre-specified. The patent (imperfectly) protects the innovator for $T \geq 0$ periods. Future amounts are discounted at rate r. The patentee's discounted profit is equal to

$$V(T, w) = \int_0^T e^{-rt}\pi_1^*(w)dt + \int_T^{+\infty} e^{-rt}\pi_1^*(0)dt$$

Note that V cannot be smaller than $V(0,0) = \frac{t}{2r}$, the competitive discounted profit. Likewise, V cannot exceed $V(+\infty, 1) = \frac{1}{18}\frac{(3t+\widehat{\theta})^2}{rt}$, the maximum reward achievable through industrial property. The discounted social welfare is equal to[20]

$$\Omega(T, w) = \int_0^T e^{-rs} S(w) ds + \int_T^{+\infty} e^{-rs} S(0) ds \tag{4}$$

The constraint $V(T, w) = V$ implies that $T = \psi(w)$ with

$$\psi(w) = -\frac{1}{r} \ln \left[1 - \frac{9t(2Vr - t)}{\widehat{\theta} w(\widehat{\theta} w + 6t)} \right]$$

The discounted welfare is $\Omega(\psi(w), w)$. Differentiating with respect to w, we have

$$\frac{d\Omega}{dw} = \frac{3}{2} \frac{t(2Vr - t)(8\widehat{\theta} - 9\gamma)}{r(w\widehat{\theta} + 6t)^2}$$

so the discounted welfare increases with the patent breadth whenever $\gamma \leq \frac{8}{9}\widehat{\theta}$.

Proposition 5 *In a market with linear transport costs and horizontal differentiation, if the size of the externality is large (i.e. if $\gamma \geq \frac{8}{9}\widehat{\theta}$), then narrow and infinitely-lived patents are optimal.*

Hence, we find similar results to the example in which we interpreted the patent breadth as the cost of inventing around the patent. Namely, the public agency should prefer length to breadth whenever the negative externality is large. Note that the size of the external effect above which narrow and infinitely-lived patents are optimal increases with the magnitude of the quality improvement. Therefore, the more incremental the innovation, the wider the *range* of externalities for which length should be preferred to breadth.

2.3.5 Cournot competition

The last example we discuss is the case for which two firms compete *à la Cournot* in a homogenous product market. One of them (the patentee) achieves an innovative production process and becomes more efficient than her rival (the imitator). The

[20] We use the dummy variable s so no confusion can be made with the transport cost t.

new process generates a production negative externality. Here, we interpret the concept of patent breadth as the fraction of the cost reducing technology that can not be used by the imitator without infringing the patent.

Let $P = a - Q$ be the (inverse) demand function for the product where Q is the aggregate output and P is the price at which it is sold. Firms initially produce at unit cost $k \leq a$. The new process allows the patentee to save $d \leq k$ per unit of output. This cost reduction is assumed to be small enough so she does not monopolize the market before the patent expires.[21] The imitator can incorporate a fraction $1 - w$ of the new technology into his own production process without infringing the patent so he can produce at unit cost $k - (1 - w)d$ where $w \in [0, 1]$ is the patent breadth. Assume that firm 1 is the patentee and firm 2 is the imitator and let q_1 and q_2 denote their respective output. Hence, the patentee's profit is equal to $\pi_1(q_1, q_2) = q_1(s - q_1 - q_2 + d)$ with $s = a - k$ whereas the imitator's profit is equal to $\pi_2(q_1, q_2) = q_2(s - q_1 - q_2 + (1 - w)d)$. Firms simultaneously choose their output. The Cournot-Nash equilibrium is such that $q_1^*(w) = \frac{1}{3}[s + d(1 + w)]$ and $q_2^*(w) = \frac{1}{3}[s + d(1 - 2w)]$. Typically, the wider the cost-gap between the two competitors, the larger the equilibrium output of the low-cost firm (i.e. the patentee), the smaller that of the high-cost firm (i.e. the imitator) and the lower the aggregate output. The equilibrium profits are such that

$$\pi_1^*(w) = \frac{1}{9}(s + d(1 + w))^2 \text{ and } \pi_2^*(w) = \frac{1}{9}(s + d(1 - 2w))^2$$

As was the case in the previous example, note that firms make supernormal profits after the patent expires because the market is imperfectly competitive. Also, we observe that the patentee's profit is non-linearly related to the patent breadth so the *extended Gilbert & Shapiro* condition does not apply. The consumer surplus

[21] It implies that the cost reduction d must be smaller than $a - k$.

is given by

$$CS(w) = \frac{1}{18}(2s + d(2-w))^2$$

Again, let $R(w)$ denote the imitation's relative rate of conveying so the weighted sum of output $Q(q_1^*(w), q_2^*(w), R(w))$ is equal to $q_1^*(w) + R(w)q_2^*(w)$. To simplify, we assume that $R(w) = 1 - w$. We thus have

$$Q(w, q_1^*(w), q_2^*(w)) = q_1^*(w) + (1-w)q_2^*(w)$$

The original innovation and the imitation induce a negative externality that generates a damage $d(Q(w, q_1^*(w), q_2^*(w)))$. For tractability reasons, we assume that the damage is linearly related to the weighted sum of output. That is,

$$d(Q(w, q_1^*(w), q_2^*(w))) = \gamma Q(w, q_1^*(w), q_2^*(w))$$

with $\gamma \geq 0$. For the same reasons as those explained in the previous example, the damage may increase in breadth if the cost reducing innovation is incremental. Finally, the social welfare is given by

$$S(w) = \pi_1^*(w) + \pi_2^*(w) + CS(w) - d(Q(w, q_1^*(w), q_2^*(w)))$$

We now discuss the rule according to which a pre-specified value of the patent V should be allocated over time. Let $V(T, w)$ be the patentee's discounted profit and let $\Omega(T, w)$ be the discounted social welfare as defined in (4). Here, the desired reward V must be comprised between $V(0,0) = \frac{1}{9}\frac{(s+d)^2}{r}$ and $V(+\infty, 1) = \frac{1}{9}\frac{(2d+s)^2}{r}$. The constraint $V(T, w) = V$ implies that $T = \psi(w)$ with

$$\psi(w) = -\frac{1}{r}\ln\left[1 - \frac{9rV - (s+d)^2}{dw(dw + 2(s+d))}\right]$$

The discounted welfare is $\Omega(\psi(w), w)$. Differentiating with respect to w, we have

$$\frac{d\Omega}{dw} = \frac{1}{3}\frac{(5d(d+s) - \gamma(6d+5s))(9Vr - (d+s)^2)}{r(dw + 2(d+s))^2}$$

Thus, the discounted welfare increases with the patent breadth whenever $\gamma \leq \underline{\gamma} = \frac{5d(s+d)}{6d+5s}$.

Proposition 6 *In the case of a cost reducing innovation in a linear homogenous Cournot duopoly with constant marginal costs, if the size of the externality is large (i.e. if* $\gamma \geq \underline{\gamma}$*), then narrow and infinitely-lived patents are optimal.*

This result is familiar to us from the horizontal differentiation case and from that of costly imitation. It is possible to check that $\underline{\gamma}$ increases with the magnitude of the cost reduction. Hence, as was the case in the previous example, our findings suggest that if the innovation is incremental, then narrow and infinitely-lived patents are optimal for a wider *range* of externalities.

3 Optimal value of the patent

In the previous section, we focused on determining the optimal length and breadth of a patent whose value was given. That is, we assumed that the rate of innovation was pre-specified and we discussed the socially efficient way of structuring the patent. Here, we address the question of how much innovators should be rewarded when the innovation generates a negative external effect. Greater rewards induce stronger incentives to innovate. Since negative externalities undermine the appeal of innovative activities, one might expect that the larger those effects, the smaller the optimal value of the patent. As we shall see, this is not necessarily the case.

In this section, we illustrate our point by focusing on the case of compulsory licensing in which the patent breadth is interpreted as the fraction of the optimal royalty fee that the patentee is allowed to charge by the public agency. This example is indeed the easiest to work with since broader patents unambiguously generate additional deadweight loss and reduce the damage more effectively. Besides, since the original innovator and the licensees use the same innovative cost reducing technology, complications related to the relative rate of conveying are ruled out.

In the case of compulsory licensing, recall that if $\gamma \leq 1$, then narrow and infinitely-lived patents are optimal so $T^* = +\infty$ and $w^* : V(+\infty, w^*) = V$. However, if $\gamma \geq 1$, then patents of maximum breadth are optimal so $w^* = 1$ and $T^* : V(T^*, 1) = V$. In the former case, the *optimal* discounted welfare generated by the innovation[22] is given by $\underline{\Omega}(V) = \Omega(+\infty, w^*)$ so

$$\underline{\Omega}(V) = \frac{1}{4}\left[\frac{s(1-\gamma)(s+\sqrt{s^2-4rV})}{r} + 2(1+\gamma)V\right]$$

In the latter, it is defined by $\overline{\Omega}(V) = \Omega(T^*, 1)$ so

$$\overline{\Omega}(V) = \frac{1}{2}\left[\frac{s^2(1-\gamma)}{r} + (3\gamma-1)V\right]$$

We thus have

$$\Omega(V) = \begin{cases} \underline{\Omega}(V) & \text{if } \gamma \leq 1 \\ \overline{\Omega}(V) & \phantom{\text{if }} \gamma \geq 1 \end{cases}$$

In the case for which the patent length is maximum, breadth is the only adjustement variable to a change in the value of the patent. For small rewards, the patent is narrow so an increase in the strength of protection raises the optimal discounted welfare. Indeed, recall that in the case for which $\gamma \leq 1$, the marginal deadweight loss was outweighed by the marginal damage reduction for narrow patents. However, since the balance of power reversed for broader patents, the optimal discounted welfare starts declining beyond a certain value of the patent.[23]

In the case for which the patent breadth is maximum, the optimal discounted welfare increases linearly with the value of the patent. Indeed, recall that when the negative externality is large, the welfare declines once the patent expires. Therefore, by delaying the onset of perfect competition, stronger industrial property raises the optimal discounted social welfare.

[22] Here, we use the term *optimal* to indicate that the value of the patent is allocated over time in the most socially efficient way.

[23] It is immediate to show that $\underline{\Omega}(V)$ is srictly concave and that it is maximum at $V = \frac{\gamma s^2}{r(1+\gamma)^2}$.

Let us now describe the problem of a rational, profit-maximizing and risk-neutral firm that is endowed with an idea for a cost-reducing innovation.[24] Let $x \in [0,1]$ be the amount of resources that she devotes to R&D at cost

$$C(x) = \frac{1}{2}cx^2$$

where $c \geq \overline{V} = V(+\infty, 1)$ so corner solutions are ruled out. Apart from choosing how much to invest in that process, the firm decides whether or not to undertake actions aimed at preventing the negative externality from arising. As mentioned in chapter one, such actions may include the implementation of strict safety protocols or the employment of highly-skilled researchers. Let $y \in \{0,1\}$ be the binary variable that determines whether or not the firm undertakes preventive actions: $y = 0$ means that she does not, while $y = 1$ indicates that she does. The negative externality arises in the former case only.[25] Preventive actions are costly in the sense that they divert a fraction of the firm's resources from their primary objectives so they undermine the chances that R&D succeeds. In particular, we assume that the firm achieves innovation with probability

$$p(x,y) = \begin{cases} x & \text{if } y = 0 \\ ax & \phantom{\text{if }} y = 1 \end{cases}$$

where $a \in]0,1[$ measures the cost of undertaking such actions.[26] Clearly, the greater the amount of resources that the firm devotes to R&D, the more likely it is that innovation is achieved. If the process fails, then the firm implements the *status quo* action which is assumed to generate zero profit and zero social welfare.[27]

[24] Unlike Denicolò (1996), we do not address the case in which several firms are engaged in a patent race. The issue of rivalry in R&D will be addressed in the next chapter.

[25] If γ is understood as the expected size of the externality, then our framework may include risk in the relation between the firm's R&D strategy and the innovation's actual impact on welfare.

[26] Pay attention to the fact that preventive actions induce a large cost when a is small.

[27] The assumption that the *status quo* action generates zero social welfare suggests that the initial unit cost of production is larger than a.

We now consider two types of policy: the *standard patent policy* and the *modulated patent policy*. In the *standard policy* the value of the patent is the same whatever the firm's R&D strategy. In the *modulated policy*, the public agency is able to reward preventive actions by giving stronger protection to the firm. Obviously, this latter policy is hardly implementable and remains mainly theoretical. The modulated patent policy is indeed confronted to the *one size fits all* problem and the informational requirements are likely not to be met.

In both cases, the timing of the game is as follows: first the public agency determines and announces the value(s) of the patent; second, the firm chooses her R&D strategy. We thus face a dynamic game whose information is perfect and complete. We do not not allow players to randomize their strategies and we search for Subgame Perfect Nash Equilibria (SPNE). Thus, we shall first compute the firm's optimal R&D strategy and then, we will determine the optimal value(s) of the patent.

3.1 Standard patent policy

In this first type of policy, the firm is given a protection whose value V is the same regardless of her R&D strategy. She thus solves

$$\max_{\{x,y\}} \Pi(x,y) = p(x,y)V - C(x)$$

Clearly, undertaking preventive actions can never be optimal because it induces a cost and it is not rewarded by stronger protection. Hence, it is straightforward to show that the firm's optimal R&D strategy is such that $(x^*, y^*) = (x_0, 0)$ with $x_0 = \frac{V}{c}$. Her optimal profit is equal to $\Pi(x_0, 0) = \frac{1}{2}\frac{V^2}{c}$. Note that it is always positive so the firm's participation constraint is always satisfied.

We now go back to the first stage of the game in which the public agency chooses V^*, the optimal value of the patent. We define

$$W_0(V) = p(x_0, 0)\Omega(V) - C(x_0)$$

as the social welfare induced by the R&D strategy $(x_0, 0)$. Formally, the public agency solves

$$\max_{V \leq \overline{V}} W_0(V)$$

where $\overline{V}$ is the maximum reward achievable through industrial property

If $\gamma \geq 1$, then $\Omega = \overline{\Omega}$. In that case, as depicted in Figure 2, we show that the problem has no interior solution. In that case, indeed, the size of the negative external effect would justify a strong protection but the existence of an upper bound for the value of the patent prevents the public agency from giving the firm the protection that is socially optimal in the unconstrained problem. If the size of the externality is moderate (i.e. if $\gamma \in [1, 2]$), then maximum protection yields a positive welfare so the constraint binds at the optimum (i.e. $V^* = \overline{V}$). However, if the externality is too large (i.e. if $\gamma \geq 2$), then the welfare is negative for all admissible values of the patent so the public agency should give the firm no protection at all (i.e. $V^* = 0$).

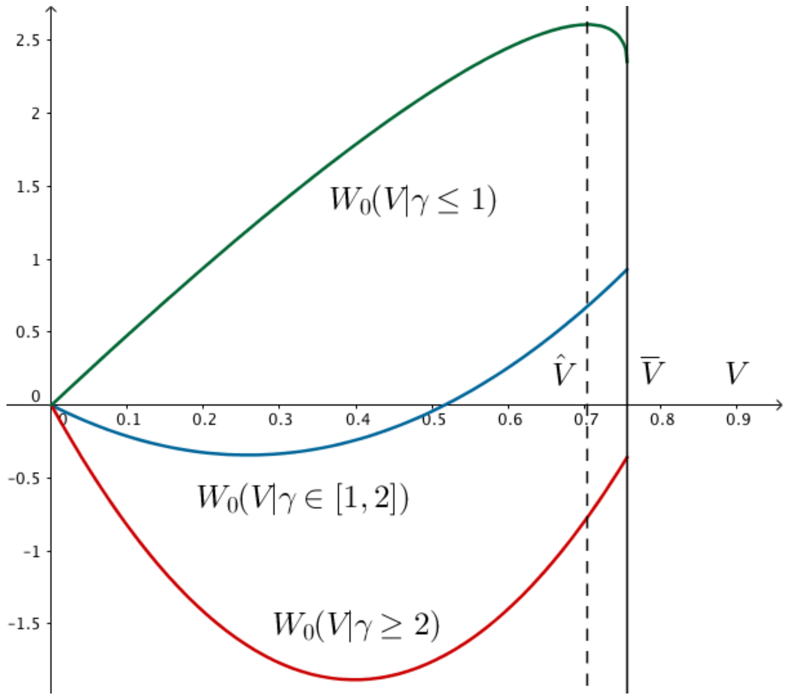

Figure 2: Welfare and value of the patent

If $\gamma \leq 1$, then $\Omega = \underline{\Omega}$. In that case, we show that the problem has an interior solution such that $V^* = \widehat{V}$ with

$$\widehat{V} = \frac{1}{32}s^2 \left[\frac{(3-\gamma)\sqrt{9(1-\gamma)^2 + 4\gamma} + 3\gamma^2 + 10\gamma - 9}{r\gamma^2} \right]$$

Note that $\widehat{V}$ - and thus the rate of technological change - increases with γ so the public agency should give the firm stronger protection as the size of the negative externality goes up. The reason is that the benefit of reducing the damage through broader patents increases with γ. Let us define V_0 as the optimal value of the patent in the standard policy. As depicted in Figure 3, We thus have

$$V_0 = \begin{cases} \widehat{V} \text{ if } \gamma \leq 1 \\ \overline{V} \text{ if } \gamma \in [1,2] \\ 0 \text{ if } \gamma > 2 \end{cases}$$

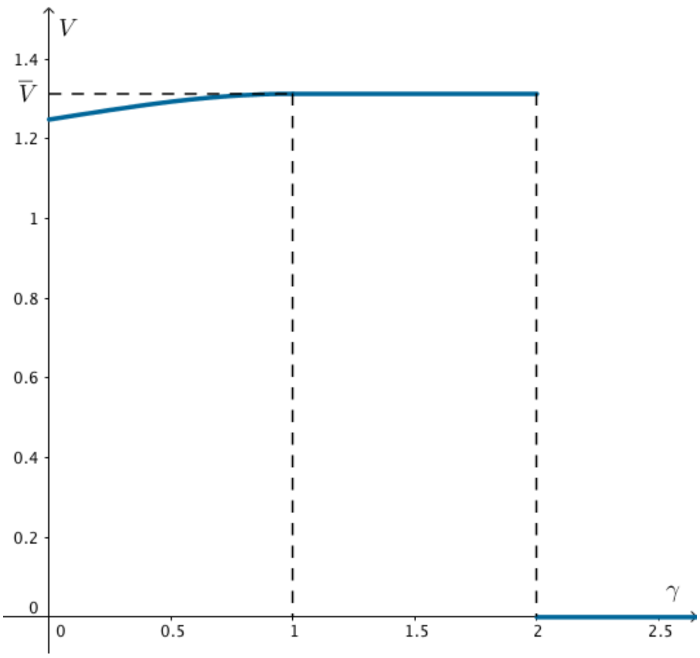

Figure 3: Optimal value of the patent as a function of the size of the negative external effect

The following Proposition summarizes our main findings:

Proposition 7 *In the standard patent policy, as the size of the externality goes up, the optimal value of the patent increases, reaches the maximum strength of protection, and then falls to zero.*

3.2 Modulated patent policy

In this section, we examine a policy in which the public agency may reward preventive actions by giving stronger protection to the firm. In that case, the patent policy has the potential to give incentives to innovate and to reduce the negative externality.

To some extent, this type of policy echoes patent laws that incorporate a precautionary principle. Usually, new products or processes qualify for industrial property whenever they fulfill three patentability standards: novelty, usefulness and non-obviousness. As emphasized by Kolitch (2006), there are numerous examples of nations whose patent policies are more stringent and exclude potentially harmful innovations from industrial property. In that case, preventive actions may be rewarded in the sense that they make the public agency less inclined to trigger the precautionary principle and deny patentability.

In the modulated policy, the value of the patent is defined as follows:

$$V(y) = \begin{cases} V & \text{if } y = 0 \\ V + B & \phantom{\text{if }} y = 1 \end{cases}$$

where $B \geq 0$ is the additional value that is given to the firm when she undertakes preventive actions. Note that $B \leq \overline{B} = \overline{V} - V$ since the value of the patent cannot exceed the maximum reward achievable through industrial property.[28] The firm then solves

$$\max_{\{x,y\}} \Pi(x, y) = p(x, y)V(y) - C(x)$$

If she does not undertake preventive actions, then the firm's investment decision is identical to that in the standard policy case so she implements the R&D strategy $(x_0, 0)$ and she gets $\Pi(x_0, 0)$. However, if she does undertake preventive actions, then her optimal R&D strategy is such that $(x^*, y^*) = (x_1, 1)$ with $x_1 = \frac{a(V+B)}{c}$ so she gets $\Pi(x_1, 1) = \frac{1}{2}\frac{a^2(V+B)^2}{c}$. Again, the firm's participation constraint is trivially satisfied. Clearly, the firm undertakes preventive actions if $\Pi(x_1, 1) \geq \Pi(x_0, 0)$. The inequality is true whenever

$$B \geq \underline{B} = \frac{(1-a)V}{a}$$

That is, the extra value of the patent must be sufficiently large to induce the firm to undertake preventive actions. As depicted in Figure 4, note that $\underline{B} \leq \overline{B} \Leftrightarrow$

[28] In the standard policy, we had $B = 0$.

$V \leq a\overline{V}$. Indeed, since there is an upper bound on the value of the patent, the public agency cannot give the firm a sufficient extra protection when V is too large. Let us define Δ as the set of *modulated patent policies* (V, B) that induce the implementation of preventive actions. We then have

$$\Delta = \left\{ V \leq a\overline{V}, B \in \left[\underline{B}, \overline{B} \right] \right\}$$

Hence, the firm's optimal R&D strategy is such that

$$(x^*, y^*) = \begin{cases} (x_0, 0) & \text{if} \quad (V, B) \notin \Delta \\ (x_1, 1) & \phantom{\text{if}} \quad (V, B) \in \Delta \end{cases}$$

Let us now go back to the first stage of the game in which the public agency chooses the optimal modulated patent policy (V^*, B^*). To start with, we focus on the set of policies that does not induce the implementation of preventive actions (i.e. $(V, B) \notin \Delta$). In that case, the problem of the public agency is identical to that we solved in the standard policy case. We thus have $(V^*, B^*) = (V_0, B)$. Clearly, the value of B is irrelevant - so long as $(V, B) \notin \Delta$ - because the firm does not undertake preventive actions so she will not be given the extra value of the patent.

We now turn our attention to the set of modulated policies that make the firm undertake preventive actions (i.e. $(V, B) \in \Delta$). In that case, the negative externality does not arise so the patent should be narrow and infinitely-lived. We define

$$W_1(V, B) = p(x_1, 1)\underline{\Omega}(V + B|\gamma = 0) - C(x_1)$$

as the social welfare induced by the R&D strategy $(x_1, 1)$. Therefore, the public agency solves

$$\max_{(V,B)\in\Delta} W_1(V, B)$$

We show that $(V^*, B^*) = (V, B) \in \Delta$ such that $V + B = Z$ with

$$Z = \frac{1}{32}s^2 \left[\frac{\sqrt{4a^2 - 12a + 17}(5 - 2a) + 4a^2 - 13}{(1 - a)^2 r} \right]$$

Note that $Z \leq \overline{V}$ so the set of solutions is nonempty.

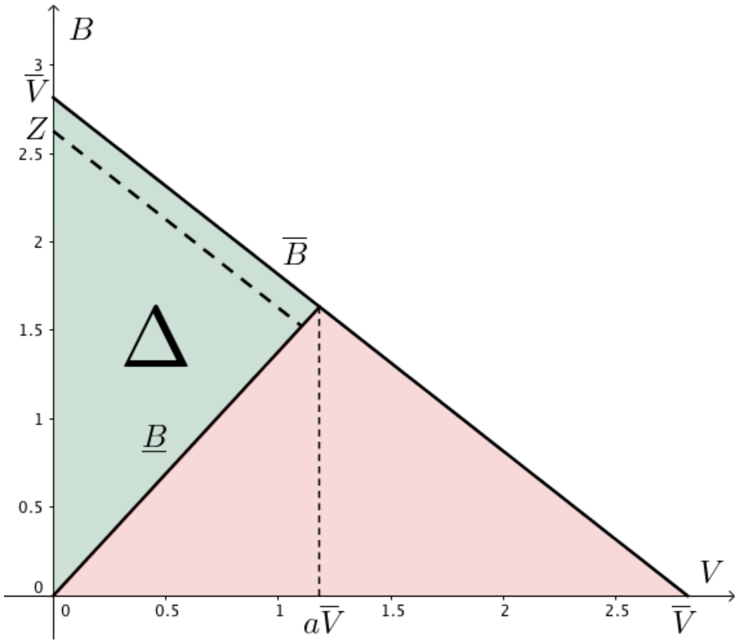

Figure 4: Modulated patent policies

Now that we have computed the optimal policy on each subset, we turn to comparing the welfare that they generate in order to determine under which conditions the optimal modulated patent policy effectively improves welfare by inducing the firm to undertake preventive actions. Clearly, policy $(V, Z - V) \in \Delta$ should be preferred to policy $(V_0, B) \notin \Delta$ whenever

$$W_1(V, Z - V) \geq W_0(V_0, B)$$

As depicted in Figure 5, we show that this inequality is true whenever the size of the externality is larger than a threshold $\widehat{\gamma} < 2$ that increases with the cost of preventive actions.[29]

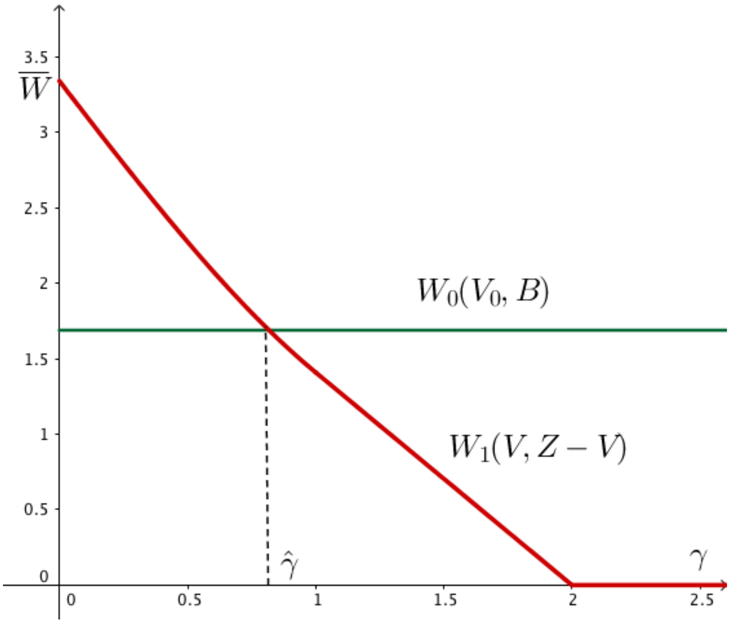

Figure 5: Standard *vs* Modulated patent policy

Proposition 8 *The modulated patent policy is strictly more efficient than the standard policy whenever the innovation generates a negative externality that is sufficiently large.*

When the negative externality is small, the *standard* and the *modulated* patent policy are equally efficient. In that case, indeed, even though the public agency

[29]To see this, note that $W_0(V_0, B)$ is continuous and decreasing with respect to γ. If $\gamma \geq 2$, then $W_0(V_0, B) = 0$. If $\gamma \to 0$ then $W_0(\widehat{V}, B) \to \overline{W} = \frac{2}{27}\frac{s^4}{r^2 c}$. Besides, $W_1(V, Z-V)$ is constant with respect to γ since the externality does not arise. Also, it decreases with the cost of preventive actions (i.e. it increases with a). If $a \to 0$, then $W_1(V, Z-V) \to 0$ whereas $W_1(V, Z-V) \to \overline{W}$ when $a \to 1$. Therefore, by the intermediary value theorem, we conclude that there exists a single value $\widehat{\gamma}$ that decreases with a and such that $W_1(V, Z-V) \geq W_0(V_0, B) \Leftrightarrow \gamma \geq \widehat{\gamma}$. Unfortunately, $\widehat{\gamma}$ cannot be computed analytically.

is able to reward preventive actions by giving the firm stronger protection, he is better off not to. The reason is that the (small) damage that can be prevented is outweighed by the cost of preventive actions. Clearly, the more those actions undermine the chances that R&D succeeds, the wider the *range* of externalities for which the standard and the modulated policy are equally efficient When the negative externality is large, the optimal modulated policy induces the firm to undertake preventive actions so the public agency actually takes advantage of the flexibility allowed by the modulated policy. In that case, the inability of the standard policy to reward preventive actions is detrimental and induces welfare losses. In particular, the efficiency gap that separates the two types of policies grows as the size of the external effect goes up.

4 Concluding remarks

In this chapter, we investigated the extent to which the introduction of a negative externality influences the design of the optimal patent policy. First, we computed the optimal length and breadth of a patent whose value was given. That is, we determined the socially efficient way to allocate a pre-specified reward over time. Our main point was that the trade-off between length and breadth should not only be driven by the impact of industrial property on the deadweight loss but also by its influence on the damage generated by the negative externality. In particular, although broader patents may generate additional deadweight loss, we argued that they can also reduce the damage by decreasing the equilibrium output or by making non-infringing imitations convey a smaller share of the external effect. Therefore, breadth should not only be seen as a cost for society but also as an instrument with the potential to address the externality. In particular, an increase in breadth may be welfare improving if the additional deadweight loss is outweighed by the reduction in damage. In the presence of a negative external effect, we showed that unlike Gilbert & Shapiro's findings, narrow and infinitely-lived patents may not be optimal even though breadth is increasingly costly in terms of deadweight loss. Instead, we provided an *extended condition* which indicates that the public agency should prefer length to breadth whenever the total socials costs (i.e. the usual deadweight loss and the damage generated by the externality) are convex in breadth.

Then, we successively examined several interpretations of the concept of breadth in order to establish closed-form conditions for the optimal length and breadth of a patent. Specifically, breadth was alternatively understood as the fraction of the optimal royalty fee that the patentee was allowed to charge by the public agency, as the (fixed) cost of *inventing around* the patent and as the fraction of a product or process improvement that could not be captured by imitators without infringing the patent. If the public agency were to overlook the impact of breadth on the externality, then we stressed that the rule according to which he would choose to allocate the value of the patent over time might be inefficient. Generally speaking, we showed that the optimal structure of the patent is closely related to the size of the external effect and to the extent to which breadth can make non-infringing imitations less harmful.

Finally, we turned to the question of how much innovators should be rewarded when the size of the negative external effect is determined by their R&D strategy. We first examined a *standard* policy in which the value of the patent was the same regardless of the firm's R&D strategy. Then, we investigated a *modulated policy* that allowed the public agency to reward safe R&D strategies by giving stronger protection to the innovator. Our conclusions indicate that the latter type of policy is strictly more efficient than the former whenever the size of the externality is sufficiently large so the public agency actually takes advantage of the flexibility allowed by the modulated policy.

5 References

1. Barnett, A. H. (1980). The Pigouvian tax rule under monopoly. *The American Economic Review*, 70(5), 1037-1041.

2. Buchanan, J. M. (1969). External diseconomies, corrective taxes, and market structure. *The American Economic Review*, 59(1), 174-177.

3. Coase, R. H. (1960). The problem of social cost. In *Classic Papers in Natural Resource Economics* (pp. 87-137). Palgrave Macmillan UK.

4. Coast, J., Smith, R. D., & Millar, M. R. (1998). An economic perspective on policy to reduce antimicrobial resistance. *Social science & medicine*, 46(1), 29-38.

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

6. Denicolò, V. (1996). Patent races and optimal patent breadth and length. *The Journal of Industrial Economics*, 249-265.

7. Denicolò, V. (1999). The optimal life of a patent when the timing of innovation is stochastic. *International Journal of Industrial Organization*, 17(6), 827-846.

8. Denicolò, V., & Franzoni, L. A. (2010). On the winner-take-all principle in innovation races. *Journal of the European Economic Association*, 8(5), 1133-1158.

9. Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty.* Princeton university press.

10. Gallini, N. T. (1992). Patent policy and costly imitation. *The RAND Journal of Economics*, 52-63.

11. Gilbert, R., & Shapiro, C. (1990). Optimal patent length and breadth. *The RAND Journal of Economics*, 106-112.

12. Horowitz, J. B., & Moehring, H. B. (2004). How property rights and patents affect antibiotic resistance. *Health economics*, 13(6), 575-583.

13. Klemperer, P. (1990). How broad should the scope of patent protection be?. *The RAND Journal of Economics*, 113-130.

14. Kolitch, S. (2006). Environmental and Public Health Impacts of US Patent Law: Making the Case for Incorporating a Precautionary Principle, *Envtl. L.*, 36, 221.

15. Laxminarayan, R. (2002). How broad should the scope of antibiotics patents be?. *American Journal of Agricultural Economics*, 84(5), 1287-1292.

16. Laxminarayan, R. (2006). Fighting Antibiotic Resistance Can Economic Incentives Play. *The RFF Reader in Environmental and Resource Policy*, 289.

17. Mitchell, M. F., & Hopenhayn, H. A. (2000). Innovation Variety and Patent Breadth. *Rand Journal of Economics*, 32(1).

18. Nordhaus, W. D. (1967). *The optimal life of a patent* (No. 241). Cowles Foundation for Research in Economics, Yale University.

19. Nordhaus, W. D. (1972). The optimum life of a patent: Reply. *The American economic review*, 62(3), 428-431.

20. Oates, W., & Baumol, W. (1975). The instruments for environmental policy. In *Economic Analysis of Environmental Problems* (pp. 95-132). NBER.

21. Scherer, F. M. (1972). Nordhaus' theory of optimal patent life: A geometric reinterpretation. *The American Economic Review*, 62(3), 422-427.

22. Scotchmer, S. (2004). *Innovation and incentives.* MIT press.

23. Tandon, P. (1982). Optimal patents with compulsory licensing. *The Journal of Political Economy*, 470-486.

24. Wright, D. J. (1999). Optimal patent breadth and length with costly imitation. *International Journal of Industrial Organization*, 17(3), 419-436.