Licensing "Weak" Patents*

David Encaoua[†] and Yassine Lefouili[‡]

This Version: March 2009

Abstract

In this paper, we revisit the issue of licensing "weak" patents under the shadow of litigation. Departing from the seminal paper by Farrell and Shapiro (2008), we consider innovations of any size and not only "small" innovations, and we allow the number of licensees to be less than the number of firms in the downstream industry. It is shown that the optimal two-part tariff from the patent holder's perspective may either deter or trigger litigation and conditions under which each case arises are provided. We also reexamine the claim that the licensing revenues from a "weak" patents overcompensate the patent holder relative to what a natural benchmark would command. We establish that this overcompensation result does not always hold, especially if the cost reduction magnitude is sufficiently large. Finally we discuss some policy levers that may alleviate the harm raised by the licensing of "weak" patents.

Keywords: Licensing Schemes, Probabilistic Rights, Patent Litigation.
JEL classification: D45, L10, O32, O34.

---

*We thank two anonymous referees and Yannis Katsoulakos for useful comments and suggestions. We are grateful to Rabah Amir, Claude d'Aspremont, Claude Crampes, Vincenzo Denicolò, Georg von Graevenitz, Abraham Hollander, Robert Hunt, Patrick Rey, David Ulph, Patrick Waelbroeck and Bertrand Wigniolle for helpful discussions. We would also like to thank participants at the CRESSE 2008 in Athens and the AEA 2008 in Tokyo. The usual disclaimer applies.

[†]Paris School of Economics, Université Paris-1 Panthéon Sorbonne, CES, 106-112, Bd de l'Hôpital 75647 cedex 13, Paris, France. E-mail: encaoua@univ-paris1.fr

[‡]CORE, Université Catholique de Louvain, and Paris School of Economics, Université Paris-1 Panthéon Sorbonne. Address: CORE, Voie du Roman Pays 34, B-1348, Louvain-la-Neuve, Belgium. E-mail: yassine.lefouili@m4x.org

1 Introduction

The literature on licensing patents has extensively examined the case where patents give their owners perfect protection, the so-called iron-clad patents. Based largely on previous works by Arrow (1962), Katz and Shapiro (1985, 1987), Kamien and Tauman (1984, 1986), Kamien *et al.* (1992), the survey by Kamien (1992) summarizes some results, especially by comparing the patent holder's profits under different licensing schemes. The per-unit royalty scheme and the up-front fee mechanism have been set against each other. While the earlier literature claimed that a per unit royalty always generates lower profits than a fixed fee, regardless of the industry size and the magnitude of the innovation (Kamien and Tauman,1984 and 1986), a more recent work (Sen, 2005) has shown that when the number of firms in the industry is sufficiently high, the innovator's payoff is higher with royalty licensing than with a fixed fee or an auction. Moreover, the number of licensees depends on the licensing method and the magnitude of the cost reduction. Sen and Tauman (2007) generalize these findings by allowing the optimal combination of an auction and a per-unit royalty in situations where the innovator may be either an outsider or an insider in the industry.[1]

The economic literature on licensing uncertain patents is more scarce despite the empirical evidence on the issuance and enforcement aspects showing that patents do not give their owners perfect protection. It is now widely recognized that the quality of many granted patents raises serious doubts. In a recent seminal work, Farrell and Shapiro (2008) - FS hereafter - analyze the licensing properties of a cost-reducing technology covered by a patent whose validity is uncertain. They consider a situation where an upstream agent holding a "probabilistic patent"[2] uses a two-part tariff licensing scheme to sell licenses to a set of symmetric downstream firms in an oligopolistic industry. The patent validity being uncertain, licensing occurs under the "shadow of litigation". FS assume that if a downstream firm rejects the license offer and infringes the patent, the patentee sues the potential violator in court, and the defendant counter-sues by claiming first that the patent is invalid, and second that her product or technology does not infringe the patent.[3] If the patent is held invalid by the court,

[1] Another burgeoning literature explores the consequences of informational asymmetries on licensing. Aoki and Hu (1999) examines how the choice between strategic licensing and litigating is affected by the levels of the litigation costs and their allocation between the plaintiff and the defendant. Brocas (2006) identifies two informational asymmetries: the moral hazard due to the inobservability of the innovator's R&D effort, and the adverse selection due to the private value of holding a license. Macho-Stadler *et al.* (1996) introduces know-how transfer and shows that the patentholder prefers contracts based on per-unit royalties rather than fixed fee payments. Other contributions, emphasizing either risk aversion (Bousquet *et al.*, 1998), strategic delegation (Saracho, 2002), strategic complementarity (Muto, 1993, Poddar and Sinha, 2004), or the size of the oligopoly market (Sen, 2006) reach the same conclusion stating the superiority of the royalty licensing scheme.

[2] The term has been coined by Ayres and Klemperer (1999) and used by Lemley and Shapiro (2005, 2007). The uncertainty creates several effects, summarized in the recent survey by Rockett (2008).

[3] Another possibility, more or less equivalent to the previous, is that a downstream firm which refuses the licensing contract acts as a plaintiff against the patent holder, claiming that the patent is not valid in order to benefit freely from the new technology. One example is given by the Institut Curie's who introduced

all downstream firms use the cost-reducing technology free of charge, whereas if the patent is ruled valid, any licensing contracts already signed remain in force, and the unsuccessful challenger is constrained to use the backstop technology. A litigation may be avoided if the patent holder licenses the patent at a tariff that no potential licensee refuses. The main result in FS is that "weak" patents - that is, patents that have a high probability to be invalidated by a court if challenged - are "overcompensated" relative to their true strength: (1) The per-unit royalty rate accepted by all firms is higher than the expected royalty when patent validity is determined prior to licensing; (2) The licensing revenue is also higher than the revenue expected by the applicant if he or she were granted a patent by the patent office with a probability equal to the patent strength. In other words, "weak" patents *punch above their weight* (Rockett, 2008). FS show that this result prevails in the class of two-part tariffs with an unconstrained up-front fee as well as for an up-front fee constrained to be non-negative. At one extreme, when negative up-front fees are allowed, a weak patent is licensed as if it were an iron-clad patent since the corresponding maximal per-unit royalty rate is compensated by a lump-sum transfer from the licensor to the licensee. At the other extreme, where negative fixed fees are not allowed, the overcompensation result still holds, but in a weaker form: the per-unit royalty rate for a "weak" patent is still larger than the expected royalty but is lower than the corresponding value for an iron-clad patent.

FS identify two mechanisms that boost per-unit royalties above the expected royalty for a "weak" patent. Consider first the situation where the sign of the up-front fee is not constrained. In this case, the price set by the licensor is the price that maximizes the joint profit of the licensor and the set of licensees. The reason is two-fold: i/ The monopoly outcome in the downstream industry can be mimicked with the right choice of a royalty since a higher royalty leads to a larger downstream marginal cost that affects downstream competition; ii/ A high per-unit royalty is compensated by a negative fixed fee, i.e. a reverse payment from the licensor to the licensee. The second mechanism relies on the fact that a downstream firm's decision to challenge the patent validity benefits all other downstream firms (Farrell and Merges, 2004, Lemley and Shapiro, 2005). Therefore, the positive externality lowers the individual incentive to challenge a patent, and this explains why downstream firms will accept larger per-unit royalties than the patent strength would command.[4]

These two mechanisms - joint profit incentive and free riding - explain why the holder of a "weak" patent can obtain a licensing revenue higher than the expected revenue when the patent validity is determined prior to licensing or when the initial review process at the patent office is sufficiently thorough to weed out "weak" patents before licensing takes place. Even if

an opposition procedure in october 2001 against the patent granted to Myriad Genetics for a "method for diagnosing a predisposition for breast and ovarian cancer associated with the BRCA1 gene".

[4]Following the *Blonder-Tongue* decision (1971), it became clear that "the attacker is not able to exclude others from appropriating the benefit of its successful patent attack", Blonder-Tongue Labs., Inc. v. Univ. of Illinois Found, 402, U.S. 313, 350 (1971).

these two mechanisms are central in licensing "weak" patents, we argue in what follows that two other unjustified restrictions in FS play an important role in explaining the overcompensation result.

First, their analysis is restricted to innovations of "small size". In the case of an innovative process, this means that the cost reduction magnitude is small relative to the market size. Making this restriction, FS implicitly identify the notion of a "weak" or a "bad" patent - i.e. a patent that has a high probability to be invalidated by a court if it is challenged - with a patent granted to a very limited application, either because it corresponds to a non-novel invention or an insufficiently inventive one. We argue below that this identification is quite restrictive: a patent may also be "bad" or "weak" for reasons different than these two legal standards.[5]

Second, they consider only equilibria in which the optimal license offer made by the patent holder is such that the whole set of firms in the downstream industry accept it. In other words, they implicitly assume that it is never in the interest of the patent holder to set a licensing contract that induces a number of licensees less than the number of downstream firms. We show that this does not characterize the set of subgame-perfect equilibria of the multi-stage licensing game.

Why are these two assumptions excessively restrictive?

a. *Are "bad" patents necessarily equivalent to innovations of "small" size* ? Two reasons justify a negative answer and explain why the restriction to "small" size process innovations is unjustified when dealing with the consequences of licensing "bad" patents.

First, it is well known that for many reasons the patent office may improperly grant a patent that will likely be found invalid by a court, in case of litigation. The invalidation may be made either on the grounds of the usual patentability standards of novelty and nonobviousness or by using other standards such as utility (industrial application in Europe), and patentable subject matter. Whatever the adopted criterion, it is abusive to assimilate the "strength" of the patent with the "size" of the invention which it is supposed to protect. Take first the novelty and inventivity criteria. The application made to the patent office may have derived from a previous unpatented industrial practice that the office may have failed to identify as a prior art. The examination process may indeed be inadequate, either because the patent office is overloaded (its resources are insufficient) or because finding the relevant prior art is difficult, particularly in some new patentable subject matters in which the prior art is to be found in industry practice rather than in the patent databases usually consulted by the

[5] Guellec and van Pottelsberghe (2007) define a "bad" patent or a patent of "bad quality" with regards to the economic objectives of the patent system (encourage innovation and the diffusion of technology). A "bad "patent refers to an intellectual property right such that *its negative social costs outweigh its positive benefits* (Guellec and van Pottelsberghe, p. 115). It has been argued that the gap between the massive growth of patent applications and the insufficient resources at the patent office creates a "*vicious circle*": incentives to file "bad" applications increase the patent office overload, and a larger overload leads to further deterioration of the examination process (Caillaud and Duchêne, 2005).

examiners.[6] The know-how which constitutes the relevant prior art may nevertheless have been a "big step" for the industry. Similarly, the patent office may determine its assessment on secondary factors, such as evidence of commercial success, to decide that an invention was not nonobvious.[7] Moreover, while the claims granted by the patent office are supposed to delineate the patent scope, their *ex post* validation depends on the judicial doctrine adopted by the court. For instance, an application of the "doctrine of equivalents" may lead to a large scope in view of market conditions. But this fact can be recognized by a court in case of litigation and lead to a substantial reduction of the effective patent scope.[8]

Another important argument in differentiating a "bad" patent from a patent covering a "small" invention is that uncertainty is even more pervasive concerning the extent of the patentable subject matter.[9] The extension to areas such as software, research tools or business methods, in which defining the prior art is a very difficult task, has been made possible by a decline of the usual criteria of the patentability (novelty, inventiveness, industrial application). For instance, software is now considered as a patent subject matter in the US since the Diamond v. Diehr 1981 decision[10], but in Europe, patentability of a computer program is still explicitly excluded by the European Patent Convention, even if the exclusion applies only to computer programs "as such".[11] In biotechnology, despite the fact that the genes or proteins are apparently already present in nature, the novelty criterion has been solved by denying their natural character because of their isolation and connecting the genomics with the chemical engineering. Many other examples exist which convey the same conclusion that an uncertain patent must not be confused with a patent covering a "small" invention, and since the sources of patent uncertainty are very numerous, it would be restrictive to assimilate a "weak" patent - i.e. a patent having a high probability to be invalidated in case of litigation - with the extent of the cost improvement it allows. Many empirical studies[12] show that the small proportion of granted patents that are litigated, and the still fewer which continue

[6] The average time spent by an examiner on each patent is about 15-20 hours in the USPTO (Jaffe & Lerner, 2004) and around 30 hours in the EPO.

[7] As an example, take the patent examined by the Supreme Court decision in *Graham v. Deere*. It claimed a combined sprayer and cap used on bottles of household chemicals. The main elements had been developed by others, but they had never been assembled in the particular way which explains the commercial siuccess of the combination. On the difficulties in assessing the nonobviousness criterion, see Merges and Nelson (1990) and Barton (2003).

[8] Returning to the Myriad Genetics case, we note that following the appeal procedure of November 2008, the European Patent Office has partly gone back on its previous decision, by granting the University of Utah, now owner and sole license holder of the patents of Myriad Genetics. However, the two patents re-examined (BRCA1 and BRCA2) have been largely reduced in scope.

[9] We thank an anonymous referee for suggesting a discussion of this point. .

[10] At that time, the patent office rejected the patent application and the Supreme Court disagreed. *The rejection was based on the grounds that the only new aspect of the invention was the computer program, supposed to be a non-patentable subject matter, while the Supreme Court argued that the invention was an improved process of making rubber goods that happened to use a computer program* (Hunt, 2001)

[11] See Guellec and van Pottelsberghe (2007, pp. 125-129).

[12] See Scotchmer (2004, ch.7), Lanjouw and Schankerman (2001, 2004) for statistics in the US.

litigation until a trial, *appear to be the high value patents and those drawn from a subset of particularly litigious technology areas* (Rockett, 2008).

Second, it is clear that the consequences of a new technology depend not only on its characteristics but also on other economic features, such as the intensity of competition, the price elasticity of the demand function, the number of firms in the downstream industry, etc.[13] One cannot exclude the possibility that even a very small step in terms of performance allowed by the new technology may have a substantial consequence for an inefficient firm. This occurs for instance when competition is very tough, the extreme case being Bertrand competition with homogeneous products where an inefficient firm cannot survive whatever the size of its inefficiency. What matters in fact is whether a firm remains viable or not in the market when it has no access to the new technology while the rivals do. An unsuccessful attacker may be more or less in jeopardy or even evicted from the market once deprived from the new technology if the cost reduction is sufficiently large or if the competition in the market is very intense. It is also possible that an infringer be condemned after receiving an injunction to shut down.[14] For all these reasons, the incentive to challenge a patent validity is affected by the risk to be evicted from the market when the challenge is unsuccessful. The consequence is that the per-unit royalty that a downstream firm will accept to pay for the new technology is not the same, depending on whether the technology is "essential" or not for the firm. As we will show, this argument, absent in FS, affects in a substantive way the overcompensation result.

Finally, a last point is related to the so-called "pass-through" argument. Licensees are induced to accept a high per-unit royalty rate insofar as they are able to pass-on the royalty to their customers.[15] This possibility depends on how downstream profits are affected by the cost structure since increasing the per-unit royalty is equivalent to increasing the unit cost.

b. *Does the whole set of firms necessarily buy a license at equilibrium*? FS exclusively consider the case where at the subgame perfect equilibrium, the patent holder proposes a licensing contract such that the whole set of firms in the downstream industry accept it. This is restrictive for many reasons. First, one has to begin by defining a demand function for licenses. This demand function, which appears in the literature devoted to licensing iron-clad patents, is absent in FS and we devote some effort to define it in the framework of uncertain patents. Second, one has to derive from the demand function the two-part tariff that maximizes the licensing revenues. The case where the level of the per-unit royalty acceptable by all firms is so low that the patent holder prefers to choose a licensing contract that induces litigation

[13] See Boone, 2001.

[14] As a recent illustration, Research in Motion (RIM), the developer of the Blackberry was sued by a patent holding company NTP who won the case and was eventually granted an injunction that would shut down RIM's main activity. This induced RIM to settle the litigation for about $ 600 million (Hunt, 2008).

[15] The question of the extent to which royalties are passed-on to consumers is a difficult one. *A priori*, the pass-through will be the more effective the more competitive the product market (Farrell and Merges, 2004).

can not be excluded *a priori*. FS explain that *even if the patent holder may choose to exclude some firms in an iron-clad framework, restricting the number of licenses offered does not work as a licensing strategy for a probabilistic patent since firms that do not receive licenses will infringe the patent*[16], and the patent holder will sue them. But, one cannot exclude the situation where the patent holder prefers the outcome of a trial to an outcome in which the licensing contract must include a low royalty to be accepted by the whole set of firms. We show that such outcome is possible, and it depends on the size of the innovation and other factors such as the intensity of competition in the downstream industry. The two questions of how many firms will accept a licensing contract for a probabilistic patent and what is the best offer of the patent holder are far from being trivial questions and we devote a large attention to them in the paper.

We follow the literature on licensing iron-clad patents by using its methodology to analyze the licensing of an uncertain patent. We develop a three-stage game in which the patent holder, acting as a Stackelberg leader, determines a two-part tariff at the first stage. At the second stage, each downstream firm independently decides whether to accept the licensing contract or not. If a firm does not, it challenges the patent validity. If the patent is found valid, the unsuccessful challenger is bound to use the old technology. If the patent is found invalid, all the firms in the oligopolistic industry have free access to the technology. This allows us to determine a demand function for licenses. In the last stage, licensees and non-licensees compete in the product market with marginal costs inherited from the result of the previous stage.

Our paper departs from FS in several ways. The first is minor: in our model, the plaintiff is a potential licensee that refuses the licensing offer and decides to challenge the patent validity. The second is more substantial: insofar as we do not assimilate a "weak" patent with a small step innovation, we investigate the consequences of licensing "weak" patents whatever their size. The third is that we relax the assumption that the patent holder licenses every firm in the industry, which makes it possible to endogeneize the number of licensees. Finally we discuss the role of some policy levers that may alleviate the harm caused by licensing "weak" patents.

Our main results are as follows. First, we show that if we restrict the analysis to a pure per-unit royalty, we obtain only two possible equilibria: in the first, the whole set of n firms of the downstream industry accept the royalty, and in the second only one of these firms does not accept. Second, we show that among the class of two-part tariffs (r, F) with $F \geq 0$, the optimal scheme for "weak" patents is a pure per-unit royalty that avoids litigation, while the optimal scheme for "stronger" patents is a two-part tariff involving a per-unit royalty component that decreases in the patent strength. Third, we establish necessary and sufficient conditions under which litigation occurs. More specifically, we show that the optimal licensing contract

[16] See footnote 16 in FS.

depends on the level of the per-unit royalty that deters litigation. If this level is "high", i.e. above some defined threshold, the optimal licensing scheme is the per-unit royalty that deters litigation. But, if this level is "low", i.e. below the defined threshold, the patent owner prefers to trigger litigation by selling the license to a subset of firms. The comparison of the per-unit royalty that deters litigation with the threshold allows thus a complete characterization of the optimal licensing contract. Fourth, we give a necessary and sufficient condition distinguishing whether "weak" patents are overcompensated or not relative to their strength. If the elasticity of profits with respect to cost reduction is lower than one, the holder of a "weak" patent is overcompensated, but if the elasticity is higher than one, the holder of a "weak" patent cannot impose a overweighted royalty without triggering litigation. Thus, whenever the downstream market is such that a firm's profit is highly affected by a uniform cost reduction, the per-unit royalty that deters any challenge is below the benchmark level. We check in two standard oligopoly models that the latter condition is fulfilled when the cost reduction magnitude is sufficiently large. Finally, we explore some economic policy levers that could alleviate the harm raised by licensing "weak" patents.

The remainder of the paper is organized as follows. Section 2 presents the model and describes the timing of the 3-stage licensing game. Section 3 presents two simple examples illustrating the role of the cost reduction magnitude. Section 4 derives the demand function for licenses for any two-part tariff and any patent strength. Section 5 is devoted to the determination of the optimal two-part tariff by the patent holder for licensing a "weak" patent. In section 6, we discuss some policy levers that reduce the range of the overcompensation result. Section 7 concludes.

2 The model

We consider an industry consisting of $n \geq 2$ symmetric risk-neutral firms producing at a marginal cost c (fixed production costs are assumed to be zero). A firm P outside the industry holds a patent covering a technology that allows each firm to reduce its marginal cost from c to $c - \epsilon$. The patent is uncertain in the sense that it could be invalidated by a court if litigated: it is only with a probability θ that the patent is upheld. The parameter θ measures the patent's strength. We examine the following three-stage game:

First stage: The patent holder P proposes a two-part tariff licensing contract (r, F) whereby a licensee can use the patented technology against the payment of a per-unit royalty rate r and a fixed fee F.[17]

Second stage: The n firms simultaneously and independently decide whether to purchase a license (r, F). If a firm does not accept the license offer, it can challenge the patent validity

[17] We assume as in FS, and consistently with intellectual property law in the US, that the royalty rate r cannot exceed the innovation size ϵ.

before a court.[18] If the patent is upheld then a firm that does not purchase the license uses the old technology,[19] thus producing at marginal cost c, whereas those who accepted the license offer use the new technology and pay the royalty rate r to the patent holder, having thus an effective marginal cost equal to $c - \epsilon + r$ and an effective fixed cost equal to F. If the patent is invalidated, all the firms, including those who accepted the offer, can use for free the new technology and their common marginal cost is $c - \epsilon$. [20]

Third stage: The n firms produce under the cost structure inherited from stage 2. The kind of competition that occurs is not specified. It is only assumed that there exists a unique Nash equilibrium in the competition game between the members of the oligopoly for any cost structure of the firms. Considering an industry of n firms out of which k firms ($k < n$) - called "efficient" firms - produce at the marginal cost $x < c$ and the remaining $n - k$ firms - called "inefficient" firms - produce at the marginal cost c, we denote by $\pi^e(k, x)$ (respectively $\pi^i(k, x)$) the equilibrium profit function - gross of a potential fixed cost corresponding to the up-front fee - of an efficient firm (respectively an inefficient firm). In the case where all firms produce at the same marginal cost $x \leq c$, we denote a firm's profit indifferently by $\pi^e(n, x)$ or $\pi^i(n, x)$ since all firms are equally efficient.

We set the following general assumptions that hold for a large class of economic environments including Cournot competition and differentiated Bertrand competition both with linear demand:

A0. If all the downstream firms produce with the old technology, they make positive profits: $\pi^e(n, c) > 0$.

A1. An efficient (respectively inefficient) firm's equilibrium profit $\pi^e(k, x)$ (respectively $\pi^i(k, x)$) is continuous in x over $[0, c]$ and twice differentiable in x over the subset of $[0, c]$ in which $\pi^i(x, k) > 0$.

A2. An inefficient firm's equilibrium profit is increasing in the efficient firms' marginal cost: If $\pi^i(k, x) > 0$ then $\pi^i_2(k, x) \equiv \frac{\partial \pi^i(k,x)}{\partial x} > 0$ and if $\pi^i(k, x) = 0$ then $\pi^i(k, x') = 0$ for any $x' > x$.

[18] In the US, a firm can seek a declaratory judgement against the validity of a patent if it has a "reasonable apprehension" of being sued for infringement by the patentholder. A firm that is planning to use a patented technology, or is currently using it, without a license can reasonably fear to be sued for infringement.

[19] This assumption may seem quite strong but recall that IP laws do not compel patentholders to license others, particularly those who challenge the validity of a patent or sue the patentholder for infringement of their own patents. To illustrate, when Intergraph (a company producing graphic work stations) sued Intel (microprocessors) for infringement of its Central Processing Unit patent, Intel countered by removing Intergraph from its list of customers and threatening to discontinue the sale of Intel microprocessors to Intergraph (See Encaoua and Hollander, 2002). We relax this assumption in Section 6 by introducing renegotiation between the unsuccessful challenger and the patentholder.

[20] Note that, in a setting without litigation costs, as in our model and FS, who the plaintiff/defendant is does not matter. What matters in both models is that a trial in which patent validity will be examined by the court, will occur whenever at least one firm does not accept the licensing contract. In FS, a patent holder always finds it optimal to sue a firm that uses its technology without a license and the alleged infringer challenges the patent validity as a defense strategy. In our model a firm that refuses the licensing contract always finds it optimal to challenge the patent validity.

A3. In a symmetric oligopoly, an identical drop in all firms' costs raises each firm's equilibrium profit: $\pi_2^e(n,x) \equiv \frac{\partial \pi^e(n,x)}{\partial x} < 0$.
A4. A firm's profit is decreasing in the number of efficient firms in the industry: $\pi^e(k,x) < \pi^e(k+1,x)$ and $\pi^i(k,x) \leq \pi^i(k+1,x)$ for any $x < c$ and any $k < n$.
A5. The incremental profit from getting efficient decreases with the number of efficient firms: for any $x < c$, $\pi^e(k,x) - \pi^i(k-1,x)$ is decreasing in k.
Note that assumption A3 holds if own cost effects dominate rival's cost effects. This assumption, while being fulfilled in a wide range of competitive settings may not be satisfied under Cournot competition when the demand is "very convex" (see Kimmel 1992, Février and Linnemer 2004). The other assumptions are quite usual in oligopoly theory (see for instance Amir and Wooders, 2000).

3 "Weak" patents are not always overcompensated: Two examples.

Before solving for the subgame perfect Nash equilibria of the three-stage game, we start with the following two examples which show directly that the main result in FS stating that "weak" patents are always overcompensated may not hold in some usual economic environments, as soon as a crucial assumption - namely the "small" size of the innovation - is not fulfilled anymore. To simplify the analysis in these examples, we use a pure per-unit royalty licensing scheme. We will show in section 5 that this restriction does not entail any loss of generality relative to two-part tariff licenses for "weak" patents as we will prove that optimal two-part tariff licensing schemes do not to involve an up-front fee when the patent strength is sufficiently low.
A royalty rate r such that all firms accepting the licensing contract is a Nash equilibrium of the second stage if and only if no firm has an incentive to deviate by refusing to buy a license at this rate and challenging the patent validity.[21] The expected profit from such a unilateral deviation when the patent strength is θ is given by $\theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\pi^e(n, c-\epsilon)$. Thus, all firms accepting a pure per-unit royalty contract r is a Nash equilibrium if and only if:

$$\pi^e(n, c-\epsilon+r) \geq \theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\pi^e(n, c-\epsilon)$$

[21] Since there are no litigation costs, a firm that refuses the licensing contract always finds it optimal to challenge the patent validity. For an analysis of the effects of litigation costs on licensing under the shadow of litigation, see Aoki & Hu (1999). See also Bebchuk (1984), Reiganum and Wilde (1986), Meurer (1989) and Crampes and Langinier (2002) for an analysis of the effects of litigation costs on enforcement issues.

3.1 Example 1: Cournot with homogeneous product and linear demand

Suppose that the downstream market is an homogeneous oligopoly market under Cournot competition with linear demand $Q = \max(0, a - p)$. We want to focus on the case where an unsucessful challenger is not viable when the other firms buy a license, i.e. $\pi^i(n-1, c-\epsilon+r) = 0$. It is straightforward to show that the latter condition holds if and only if $r \leq \hat{r} = \epsilon - \frac{a-c}{n-1}$. For this condition to hold for a non-empty range of royalty rates $r \in [0, \epsilon]$, we need to assume that $\epsilon \geq \frac{a-c}{n-1}$.

For a patent of strength θ, a licensing contract based on a royalty rate $r \leq \hat{r}$ is accepted by all firms if and only if $r \leq r(\theta)$ where $r(\theta)$ is the unique solution in $r \in [0, \epsilon]$ to the following equation:

$$[a - c + \epsilon - r]^2 = (1 - \theta)[a - c + \epsilon]^2$$

The positive solution to this equation is given by $r(\theta) = [1-\sqrt{1-\theta}](a-c+\epsilon)$. This expression can be used to determine the patent strength threshold $\hat{\theta} = 1 - [\frac{n(a-c)}{(n-1)(a-c+\epsilon)}]^2$ such that $r(\hat{\theta}) = \hat{r}$.

Hence, whenever $\epsilon \geq \frac{a-c}{n-1}$, the maximal royalty rate the patent holder can make all firms accept is given by:

$$r(\theta) = [1 - \sqrt{1-\theta}](a - c + \epsilon) \quad \text{if} \quad 0 \leq \theta \leq \hat{\theta}$$

We compare this rate to a "natural" benchmark: the expected value of the maximal royalty rate in case of litigation, that we denote by $r^e(\theta)$. This benchmark can be easily computed: with probability θ the patent is upheld by the court, becoming thus an iron-clad right that can be licensed at a maximal per-unit royalty ϵ, and with probability $1 - \theta$ the patent is invalidated and the firms can use it for free, leaving the patent holder with zero royalty. Thus, the expected value of the maximal royalty rate in case of litigation is equal to $r^e(\theta) = \theta\epsilon$. In FS, this benchmark is interpreted as the *ex ante* value of the per-unit royalty rate that the owner of a process innovation reducing the cost by ϵ can expect when the patent has a probability θ to be granted by the patent office.

A direct comparison of $r(\theta)$ with the benchmark $\theta\epsilon$ leads to two cases depending on the magnitude of ϵ:

- If $\epsilon \in \left[\frac{a-c}{n-1}, a - c\right]$ then $r(\theta) > \theta\epsilon$ for any $\theta \in \left]0, \hat{\theta}\right]$.

- If $\epsilon \geq a-c$ then $r(\theta) < \theta\epsilon$ for any $\theta < \min(\check{\theta}, \hat{\theta})$ where $\check{\theta}$ is defined by $\frac{\epsilon}{a-c} = \frac{1-\sqrt{1-\check{\theta}}}{\check{\theta}-[1-\sqrt{1-\check{\theta}}]}$.

We see through this example that "weak" patents, i.e. for which θ is low - can punch above as well as below their strength. While FS show that punching above always occurs for "small" innovations, the homogeneous linear Cournot case shows that punching below occurs for "large" innovations:[22] the per-unit royalty for a "weak" patent covering a sufficiently

[22] The fact that the threshold $a - c$ for the cost reduction magnitude is the same as the the threshold defining

large cost reduction technology is below the benchmark. In this case, the overcompensation result is not satisfied.

3.2 Example 2 : Differentiated Bertrand duopoly with linear demand

Consider now a market with $n = 2$ firms producing differentiated goods. Assume that the inverse demand function for product $i = 1, 2$ is given by $p_i = a - (q_i + \gamma q_j)$ where $j \neq i$ and $\gamma \in]0, 1[$ is an inverse measure of differentiation. Here again we focus on the case where $\pi^i(n-1, c-\epsilon+r) = \pi^i(1, c-\epsilon+r) = 0$. Some calculations show that the latter is satisfied if and only if $r \leq \epsilon - \bar{\epsilon}(\gamma)$ where $\bar{\epsilon}(\gamma) = \frac{(2+\gamma)(1-\gamma)^2}{\gamma}(a-c)$. For this condition to hold for a non-empty range of royalty rates $r \in [0, \epsilon]$, we need to assume that $\epsilon \geq \bar{\epsilon}(\gamma)$. Note that the threshold $\bar{\epsilon}(\gamma)$ is, intuitively, decreasing in γ. Furthermore, the equilibrium profit when both firms produce at marginal cost $c - \epsilon$ is $\pi^e(2, c-\epsilon) = \frac{1-\gamma}{(2-\gamma)^2(1+\gamma)}(a-c+\epsilon)^2$. After some computations, we obtain for $\epsilon \geq \bar{\epsilon}(\gamma)$ the same solution to the equation $\pi^e(2, c-\epsilon+r) = (1-\theta)\pi^e(2, c-\epsilon)$ as in the previous case, i.e. $r(\theta) = [1 - \sqrt{1-\theta}](a-c+\epsilon)$. Since $\bar{\epsilon}(\gamma) \underset{\gamma \downarrow 0}{\longrightarrow} +\infty$, $\bar{\epsilon}(\gamma) \underset{\gamma \uparrow 1}{\longrightarrow} 0$ and $\bar{\epsilon}(\gamma)$ is continuous and strictly decreasing, there exists $\bar{\gamma} \in]0, 1[$ such that $\bar{\epsilon}(\bar{\gamma}) = (a-c)$. Moreover, $\bar{\epsilon}(\gamma) < a - c$ if and only if $\gamma < \bar{\gamma}$. To compare $r(\theta)$ with the benchmark $\theta\epsilon$, we must therefore distinguish two cases:

1. If the differentiation between the two products is low, i.e. $\gamma \geq \bar{\gamma}$, two sub-cases are possible:

1.a. If the cost reduction magnitude ϵ is small, i.e. $\epsilon \in [\bar{\epsilon}(\gamma), a-c]$, then $r(\theta) > \theta\epsilon$ for θ sufficiently small.

1.b. If the cost reduction magnitude ϵ is large, i.e. $\epsilon > a - c$,then $r(\theta) < \theta\epsilon$ for θ sufficiently small.

2. If the differentiation between the two products is high enough, i.e. $\gamma < \bar{\gamma}$, then for any cost reduction magnitude $\epsilon \geq \bar{\epsilon}(\gamma)$, it holds that $r(\theta) < \theta\epsilon$ for θ sufficiently small.

In this second example, it appears again that the result that the per-unit royalty for a "weak" patent is "overweighted" occurs only when the cost reduction is sufficiently small. When the cost reduction is large enough, proposing a royalty rate below the benchmark is necessary to induce all firms to accept the license.

These examples constitute special cases of a more general result that will be presented in section 5.

---

a drastic innovation results only from the specific framework in this example, that is, Cournot competition with a linear demand.

4 The demand function for licenses

We now return to the general case of two-part tariff licenses (r, F) .To determine the demand function for licenses, we start with a preliminary observation: when only $k < n$ firms accept the license offer (r, F), a situation where none of the remaining $n - k$ firms challenges the patent validity cannot be a Nash equilibrium of stage 2 whenever $\theta < 1$.[23]

The following proposition fully characterizes the equilibria of stage 2 according to the two-part tariff offer (r, F)[24]:

Proposition 1 *Denote* $F_k(r) = \pi^e(k, c - \epsilon + r) - \pi^i(k - 1, c - \epsilon + r)$, $k = 1, ..., n$, *and* $\Psi_n(r, \theta) = F_n(r) - (1 - \theta)\left[\pi^e(n, c - \epsilon) - \pi^i(n - 1, c - \epsilon + r)\right]$

- If $F \leq \Psi_n(r, \theta)$ *then all the firms purchasing a license is the unique equilibrium of stage 2.*

- If $F_n(r, \theta) < F \leq F_{n-1}(r)$ *then the Nash equilibria of stage 2 are the situations where only* $n - 1$ *firms buy a license.*

- If $F_k(r) < F \leq F_{k-1}(r)$ *where* $2 \leq k \leq n - 1$ *then the Nash equilibria of stage 2 are the situations where only* $k - 1$ *firms buy a license*

- If $F > F_1(r)$ *then the unique equilibrium of stage 2 is the situation where all the firms refuse the license offer.*

Proof. See Appendix A. ■

This proposition shows that for any pair (r, F) there exists an integer $k(r, F)$ such that all the equilibria of stage 2 involve the number $k(r, F)$ of licenses. This allows to interpret $k(r, F)$ as the demand function for licenses. The intuition behind the proposition follows from two conditions that must be satisfied at a Nash equilibrium: i/ a licensee has no incentive to deviate unilaterally by refusing the contract; ii/ a non-licensee has no incentive to deviate and become a licensee. These two conditions respectively define an upper bound and a lower bound for F.

Note that no restriction has been put on the fixed fee F up to now. In particular, in proposition 1, we allow F to be negative, that is, we do not discard the possibility of a transfer from the patent holder to the licensee. Note also that a necessary and sufficient

[23] Indeed, if one of these firms challenges the patent validity it gets an expected profit of $\theta\pi^i(k, c - \epsilon + r) + (1 - \theta)\pi^e(n, c - \epsilon)$ whereas it gets a profit equal to $\pi^i(k, c - \epsilon + r)$ if no firm challenges the patent validity. From A3 and A4, it follows that:

$$\pi^i(k, c - \epsilon + r) < \pi^i(n, c - \epsilon + r) = \pi^e(n, c - \epsilon + r) \leq \pi^e(n, c - \epsilon)$$

which yields:

$$\theta\pi^i(k, c - \epsilon + r) + (1 - \theta)\pi^e(n, c - \epsilon) > \pi^i(k, c - \epsilon + r)$$

whenever $\theta < 1$.This means that if not all firms accept the license offer, there is necessarily litigation in equilibrium.

[24] We assume that a firm which is indifferent between accepting the license offer and not, purchases a license.

condition to avoid any litigation when the two-part tariff for a patent of strength θ is (r, F) is that $F \leq \Psi_n(r, \theta)$. If $\Psi_n(r, \theta) < 0$, then the contract must involve a reverse payment from the licensor to the licensee at least equal to $\mid \Psi_n(r, \theta) \mid$to induce every firm to accept it.

It is easy to derive from proposition 1 a demand function for pure per-unit royalty licenses as this merely amounts to imposing the restriction $F = 0$. We do so because we get a quite remarkable result on the number of licensees in this case. Moreover, the pure per-unit royalty licensing scheme will turn to be optimal in the class of constrained two-part tariff licenses for "weak" patents as we will see later.

Corollary 1 *Consider the class of licenses involving a pure per-unit royalty $r \leq \epsilon$. Only two possibilities arise at Nash equilibrium:*

- If $\Psi_n(r, \theta) \geq 0$ there exists a unique equilibrium of stage 2: the n firms purchase a license;

- If $\Psi_n(r, \theta) < 0$ then the Nash equilibria of stage 2 are the situations where $n-1$ firms buy a license.

Proof. See Appendix A. ■

Licensing an uncertain patent under a pure per-unit royalty scheme may lead to only two types of equilibria: either each firm accepts the licensing contract or all firms but one accept the contract. Note that the latter case occurs if and only if $\Psi_n(r, \theta) < 0$ which is equivalent to:

$$\pi^e(n, c - \epsilon + r) < \theta \pi^i(n-1, c - \epsilon + r) + (1-\theta)\pi^e(n, c - \epsilon)$$

This inequality means that when confronted to $n-1$ firms that accept the license at a royalty r, the remaining firm prefers to challenge the patent's validity rather than accept the license. The intuition behind corollary 1 is that when the licensing scheme does not involve any fixed fee, a firm is always better off accepting to pay a royalty rate $r \leq \epsilon$ if it anticipates that litigation will be initiated by one of its rivals, which rules the possibility of a Nash equilibrium with less than $n-1$ firms.

We now return to the class of two-part tariffs and we assume that negative fixed fees are not allowed, i.e. $F \geq 0$. Under this assumption, all firms accept the licensing contract (r, F) if and only if:

$$\Psi_n(r, \theta) \geq 0 \tag{1}$$

and

$$0 \leq F \leq \Psi_n(r, \theta) \tag{2}$$

We have already seen that inequality 1[25] can be rewritten as:

[25] The role of inequality 1 here is to provide a necessary condition on r for inequality 2 to hold over a non-empty range of fixed fee values F.

$$\pi^e(n, c-\epsilon+r) \geq \theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\pi^e(n, c-\epsilon) \quad (3)$$

It is important to note that the royalty rate r affects both sides of inequality (3). Due to assumption A3, the LHS, which represents a firm's gross profit when all firms accept the license is decreasing in r. Due to assumption A2, the RHS, which represents the expected profit of a challenger when all other firms accept the license offer, is (weakly) increasing in r. Thus, for a potential licensee, a lower royalty rate r makes the license option more attractive than the outside option, namely the challenge option, for two reasons:

- It increases the payoff from the license option: $\pi^e(n, c-\epsilon+r)$ increases with r (direct effect)

- It decreases the payoff from the outside option: $\theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\pi^e(n, c-\epsilon)$ decreases with r (indirect effect).

Note that the indirect effect arises only if $\pi^i(n-1, c-\epsilon+r) > 0$. However, it may happen that the extent of the cost asymmetry between the licensees and an unsuccessful challenger result in zero profit for the latter, that is, $\pi^i(n-1, c-\epsilon+r) = 0$. In this case the indirect effect does not appear. We therefore distinguish between two cases according to whether such royalty rate values exist or not.

Case 1: $\pi^i(n-1, c-\epsilon) = 0$

This case, absent from the analysis in FS, may occur for a sufficiently large innovation (high value of ϵ) or a sufficiently intense competition (e.g. large number n of firms, price competition with high substitutability between the products, etc.).

Using assumptions A0 and A2, one easily shows that there exists a threshold $\hat{r} \in [0, \epsilon]$ such that $\pi^i(n-1, c-\epsilon+r) = 0$ if $r \leq \hat{r}$ and $\pi^i(n-1, c-\epsilon+r) > 0$ if $r > \hat{r}$. In other words, an unsuccessful challenger will not be viable if the royalty rate is below some threshold $\hat{r}$, and will make positive profit if the royalty rate is above the threshold $\hat{r}$.

The next two lemmas define a threshold function, in each of the subcases $r \leq \hat{r}$ and $r > \hat{r}$, that will be shown to be the maximal per-unit royalty acceptable by all firms.

Consider first a two-part tariff (r, F) involving a royalty rate $r \leq \hat{r}$. In this case, condition (3) can be rewritten as:

$$\pi^e(n, c-\epsilon+r) \geq (1-\theta)\pi^e(n, c-\epsilon) \quad (4)$$

Let $\hat{\theta} \in [0, 1]$ be the unique solution in θ to the equation $\pi^e(n, c-\epsilon+\hat{r}) = (1-\theta)\pi^e(n, c-\epsilon)$.

Lemma 1 *Assume that $\pi^i(n-1, c-\epsilon) = 0$. The equation $\pi^e(n, c-\epsilon+r) = (1-\theta)\pi^e(n, c-\epsilon)$ has a unique solution in r over $[0, \hat{r}]$ for any $\theta \in \left[0, \hat{\theta}\right]$. This solution, denoted $r_1(\theta)$, satisfies the* *following properties: i/ $r_1(\theta)$ is differentiable*[26] *and increasing in θ over $\left[0,\hat{\theta}\right]$, ii/ $r_1(0)=0$ and $r_1(\hat{\theta})=\hat{r}$.*

Proof. See Appendix A. ■

Consider now a two-part tariff (r,F) involving a royalty rate $r>\hat{r}$. It will be accepted by all firms if and only if the conditions (1) and (2) hold.

Lemma 2 *Assume that $\pi^i(n-1,c-\epsilon)=0$. The equation $\pi^e(n,c-\epsilon+r)=\theta\pi^i(n-1,c-\epsilon+r)+(1-\theta)\pi^e(n,c-\epsilon)$ has a unique solution in r over $[\hat{r},\epsilon]$ for any $\theta\in\left[\hat{\theta},1\right]$. This solution, denoted $r_2(\theta)$, satisfies the following properties: i/ $r_2(\theta)$ is differentiable and increasing in θ over $\left[\hat{\theta},1\right]$, ii/ $r_2(\hat{\theta})=\hat{r}$ and $r_2(1)=\epsilon$.*

Proof. See Appendix A. ■

We can now characterize the set of two-part tariff licenses (r,F) that are accepted by all firms whenever $\pi^i(n-1,c-\epsilon)=0$.

Proposition 2 *If $\pi^i(n-1,c-\epsilon)=0$ then all firms accepting the two-part tariff license (r,F) is a Nash equilibrium if and only the following conditions hold:*

i/ $r\leq r(\theta)$ where:

$$r(\theta)=\begin{cases} r_1(\theta) & \textit{if} \quad \theta\in\left[0,\hat{\theta}\right] \\ r_2(\theta) & \textit{if} \quad \theta\in\left]\hat{\theta},1\right] \end{cases}$$

ii/ $0\leq F\leq\Psi_n(r,\theta)$

Proof. See Appendix A. ■

To sum-up, when the innovation size is sufficiently large or the intensity of competition sufficiently high, proposition 2 shows that the firms' incentives to accept a given licensing contract crucially depend on whether the patent is relatively "weak" (i.e. $\theta\leq\hat{\theta}$) or relatively "strong" (i.e. $\theta>\hat{\theta}$). When the patent is "strong", the positive effect of a higher royalty rate on the outside option profit (i.e. a challenger's profit) plays a role in constraining the royalty rates acceptable by all firms: $\pi^i(n-1,c-\epsilon+r(\theta))>0$ because $r(\theta)>\hat{r}$ for all $\theta>\hat{\theta}$. However, when the patent is "weak", this indirect effect does not play a role since $\pi^i(n-1,c-\epsilon+r(\theta))=0$, due to $r(\theta)\leq\hat{r}$ for all $\theta<\hat{\theta}$. In this sense, a firm has an additional incentive not to accept a licensing contract when the patent is strong enough.[27]

[26] Throughout this paper, a function f will be said to be differentiable over a closed interval $[a,b]$ if it is differentiable at any point of the open interval $]a,b[$, right-differentiable at a and left-differentiable at b.

[27] One can get to the same interpretation using a more formal argument: defining the threshold $r_1(\theta)$ not only for $\theta\in\left[0,\hat{\theta}\right]$ but for all $\theta\in[0,1[$ as the unique solution to the equality derived from inequality (4), we can show that $r_1(\theta)<r_2(\theta)$ for all $\theta\in\left]\hat{\theta},1\right[$.

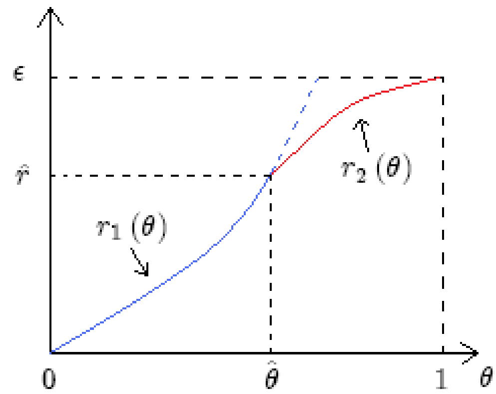

Fig 1: Shape of $r(\theta)$ if $\pi^i(n-1, c-\epsilon) = 0$

Remark **:** From lemmas 1 and 2, it is clear that the maximal royalty rate $r(\theta)$ acceptable by all firms is increasing and continuous over $[0,1]$. Moreover, it is differentiable over $\left[0, \hat{\theta}\right]$ and $\left[\hat{\theta}, 1\right]$ but its left-sided derivative is different from its right-sided derivative at point $\theta = \hat{\theta}$. One can show that the former is greater than the latter (see figure 1) which is in line with our previous observation that an extra force (stemming from the indirect effect we pointed out) constrains the royalty rates acceptable by all firms when $\theta > \hat{\theta}$.

Case 2: $\boldsymbol{\pi^i(n-1, c-\epsilon) > 0}$

In this case, whatever the royalty rate $r \geq 0$ proposed by the patent holder, the profit of an unsuccessful challenger remains positive even when all other firms purchase a license: $\pi^i(n-1, c-\epsilon+r) \geq \pi^i(n-1, c-\epsilon) > 0$. Therefore, in this case, we use the same notation $r_2(\theta)$ for the unique solution in r to the equation $\pi^e(n, c-\epsilon+r) = \theta\pi^i(n-1, c-\epsilon+r)+(1-\theta)\pi^e(n, c-\epsilon)$ for all $\theta \in [0,1]$.[28] The existence, uniqueness and properties of $r_2(\theta)$ can be established as under case 1. These are stated in the following lemma:

Lemma 3 *Assume that* $\pi^i(n-1, c-\epsilon) > 0$. *The equation* $\pi^e(n, c-\epsilon+r) = \theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\pi^e(n, c-\epsilon)$ *has a unique solution in* r *over* $[0, \epsilon]$ *for any* $\theta \in [0,1]$ *. This solution, denoted* $r_2(\theta)$*, satisfies the following properties: i/* $r_2(\theta)$ *is differentiable and increasing in* θ *over* $[0,1]$*, ii/* $r_2(0) = 0$ *and* $r_2(1) = \epsilon$.

Proof. See Appendix A. ■

The next proposition characterizes the set of licenses accepted by all firms whenever $\pi^i(n-1, c-\epsilon) > 0$.

[28] The threshold $r_2(\theta)$ that could be denoted $r_2(\theta, \epsilon)$ to explicitly display its dependence upon ϵ, has been previously defined for the values of ϵ such that $\pi^i(n-1, c-\epsilon) = 0$, and for patent strength values $\theta \in \left[\hat{\theta}, 1\right]$ (see lemma 5). Here, this threshold is defined for the values of ϵ that satisfy $\pi^i(n-1, c-\epsilon) > 0$ and for all patent strength values $\theta \in [0,1]$.

Proposition 3 *If* $\pi^i(n-1, c-\epsilon) > 0$ *then, for any* $\theta \in [0,1]$, *all firms accepting the two-part tariff license* (r, F) *is a Nash equilibrium if and only if the following two conditions hold:*

i/ $r \leq r(\theta) = r_2(\theta)$.

ii/ $0 \leq F \leq \Psi_n(r, \theta)$

Proof. See Appendix A. ∎

Note that the indirect effect that captures the positive externality of a higher royalty rate on a challenger's expected profit is always at work in constraining the royalty rates acceptable by all firms when the innovation size is sufficiently small or/and the competition intensity is sufficiently low to permit an unsuccessful challenger to be maintained in the market. This is the case on which FS focus their analysis.

5 The patent holder's optimal license offer

The question is to know whether the optimal two-part tariff from the patent holder's perspective induces the whole set of firms or only a subset of them to become licensees. For a patent strength θ, we consider four classes of two-part tariffs:

- $L = \{(r, F)\,/r \in [0, \epsilon], F \geq 0\}$ is the general set of two-part tariffs with a non-negative fixed fee;

- $L_n = \{(r, F)\,/r \in [0, r(\theta)], 0 \leq F \leq \Psi_n(r, \theta)\}$ is the subset of two-part tariffs that induce the whole set of n firms to accept the licensing contract, hence deterring litigation;

- $L_{-n} = \{(r, F)\,/r \in]r(\theta), \epsilon]$ or $F > \Psi_n(r, \theta)\}$ is the subset of two-part tariffs that induce less than n firms to accept the licensing contract, hence triggering litigation;

- $\bar{L}_n = \{(r, F)\,/r \in [0, \epsilon], F \leq \Psi_n(r, \theta)\}$ is the set of two-part tariffs, with no constraint on the sign of the fixed fee, that induce the whole set of firms to accept the licensing contract.

The optimal two-part tariffs in each of these classes are respectively denoted $(r^*(\theta), F^*(\theta))$, $(r_n^*(\theta), F_n^*(\theta))$, (r_{-n}^*, F_{-n}^*) and $(\bar{r}_n(\theta), \overline{F}_n(\theta))$. Note first that in the class L_{-n}, the optimal license (r_{-n}^*, F_{-n}^*) does not depend on the patent strength θ as can be easily derived from the demand function for licenses.

Define now $P^*(\theta)$, $P_n^*(\theta)$ and $P_{-n}^*(\theta) = \theta P_{-n}^*(1)$ as the patent holder's (expected) licensing revenues corresponding to the first three optimal two-part tariffs.

We need an additional assumption to ensure the existence and uniqueness of those licenses. Defining $q^e(n, x)$ as a firm's equilibrium output when all firms produce at the marginal cost x, we introduce the following technical assumption that holds for instance in the case of Cournot competition with linear demand

A6. The licensing revenue (per license) function $rq^e(n, c-\epsilon+r) + \Psi_n(r, \theta)$ is strictly concave in r.[29]

[29] This assumption is quite reasonable since a higher royalty rate increases the revenue per unit of ouput but

We first examine the optimal two-part tariff which deters litigation.

5.1 The optimal two-part tariff deterring litigation

Under the restriction $F \geq 0$, the optimal two-part tariff $(r_n^*(\theta), F_n^*(\theta))$ deterring litigation for a patent of strength θ is such that:

$$\begin{aligned} r_n^*(\theta) &= \arg\max_{0\leq r\leq r(\theta)} \left[rq^e(n, c-\epsilon+r) + \Psi_n(r,\theta)\right] \\ &= \arg\max_{0\leq r\leq r(\theta)} \left[\underbrace{rq^e(n,c-\epsilon+r)+F_n(r)}_{\text{objective function under } \theta=1} + \underbrace{(1-\theta)\left[\pi^i(n-1,c-\epsilon+r) - \pi^e(n,c-\epsilon)\right]}_{\text{increasing in } r \text{ and decreasing in } \theta}\right] \end{aligned}$$

and

$$F_n^*(\theta) = \Psi_n(r_n^*(\theta), \theta)$$

The fact that $r_n^*(\theta) \leq r(\theta)$ ensures that the optimal fixed fee $F_n^*(\theta) = \Psi_n(r_n^*(\theta), \theta)$ is indeed non-negative. If we were not restricting to licenses involving a non-negative fixed fee, the optimal royalty rate would be given by the maximum of the same objective function over the larger set of royalty rates $[0, \epsilon]$:

$$\bar{r}_n(\theta) = \arg\max_{0\leq r\leq \epsilon} \left[rq^e(n, c-\epsilon+r) + \Psi_n(r,\theta)\right]$$

In order to know how the optimal two-part tariff $(r_n^*(\theta), F_n^*(\theta))$ is affected by the patent strength θ, we first need to know how $\bar{r}_n(\theta)$ varies with θ. We get the following result which is in line with figure 3 in FS.[30]

Lemma 4 *In the class of two-part tariff licenses $\bar{L}_n$, the optimal royalty rate $\bar{r}_n(\theta)$ that induces n licensees is (weakly) decreasing in θ*

Proof. See Appendix A. ■

Using the previous lemma, we can now characterize the optimal license in the class L_n of licenses that deter litigation under the restriction $F \geq 0$:

Proposition 4 *There exists a threshold $\bar{\theta} \in]0,1]$, such that*

$$(r_n^*(\theta), F_n^*(\theta)) = \begin{cases} (r(\theta), 0) & if \quad \theta \leq \bar{\theta} \\ (\bar{r}_n(\theta), \Psi_n(\bar{r}_n(\theta), \theta)) & if \quad \theta > \bar{\theta} \end{cases}$$

is likely to have a negative effect on the demand adressed to each licensee, which would make the licensing revenues subject to two opposite effects, possibly resulting in a concave shape for those revenues.

[30] Note however, that in the general setting we consider, there is no reason that $\bar{r}_n(0) = \epsilon$ as in FS. Moreover, FS show that $\bar{r}_n(\theta) = \epsilon$ for θ sufficiently small but do not formally establish that $\bar{r}_n(\theta)$ is (weakly) decreasing over the interval $[0, 1]$ as we do.

Proof. See Appendix A. ■

This proposition states that in the set of two-part tariffs L_n that are accepted by all firms, the optimal licensing scheme for "weak" patents, i.e. $\theta \leq \bar{\theta}$, is a pure per-unit royalty scheme. This result hinges on the constraint put on the up-front fee ($F \geq 0$). For "strong" patents, i.e. $\theta > \bar{\theta}$, the optimal licensing scheme is the unconstrained two-part tariff that maximizes the licensing revenue.[31]

5.2 The optimal two-part tariff

An important question, absent from the analysis in FS, is whether the optimal two-part tariff induces the whole set or only a subset of the downstream firms to become licensees. To address this question, we now compare the revenues from the license $(r_n^*(\theta), F_n^*(\theta))$ that deters litigation to the revenues $P_{-n}^*(\theta) = \theta P_{-n}^*(1)$ derived from the optimal license offer $\left(r_{-n}^*, F_{-n}^*\right)$ inducing less than n licenses and hence triggering litigation. To make this comparison in the case of "weak" patents, i.e. $\theta \leq \hat{\theta}$, we consider the following equation:

$$nrq^e(n, c - \epsilon + r) = \theta P_{-n}^*(1) \tag{5}$$

The LHS of this equation corresponds to the licensing revenues from a license $(r, 0)$ accepted by all firms and the RHS is the highest expected licensing revenues the patent holder can get from a license not accepted by all firms. Denote $s(\theta)$ the solution to this equation in r over the interval $[0, \tilde{r}]$ where $\tilde{r} = \underset{0 \leq r \leq \epsilon}{\arg\max}\,[rq^e(n, c - \epsilon + r)]$. We show in the appendix that this solution exists and is unique. Further, denote $\tilde{\theta} = \min\left(\hat{\theta}, \frac{\tilde{r}q^e(n, c-\epsilon+\tilde{r})}{P_{-n}^*(1)}\right)$. The following proposition characterizes the optimal license offer for sufficiently weak patents:

Proposition 5 *For sufficiently weak patents, i.e. $\theta \leq \tilde{\theta}$, the optimal license offer is:*

$$(r^*(\theta), F^*(\theta)) = \begin{cases} (r(\theta), 0) & if \quad r(\theta) \geq s(\theta) \\ \left(r_{-n}^*, F_{-n}^*\right) & if \quad r(\theta) < s(\theta) \end{cases}$$

where $s(\theta)$ is the unique solution in r to the equation $nrq^e(n, c - \epsilon + r) = \theta P_{-n}^*(1)$.

Proof. See Appendix A. ■

This result states that the optimal license offer depends on the level of maximal the per-unit royalty $r(\theta)$ that deters litigation. If this level is high, i.e. above the defined threshold $s(\theta)$, the optimal licensing scheme is the pure royalty rate $r(\theta)$, and it is accepted by all firms. But if this level is low, i.e. under the threshold $s(\theta)$, the patent owner prefers to sell

[31] The result in Proposition 4 is present in FS. However FS state that: "*If licenses cannot use such negative fixed fees, we **assume** that they will consist simply of a per-unit royalty rate" (p.1350).* This assumption is justified later in FS through a "heuristic" argument, but is not rigorously proven as in our paper.

its license to a subset of firms, at the optimal two-part tariff that triggers litigation. Recall that the latter tariff does not depend on the patent strength θ.

Proposition 5 calls for a comparison of $r(\theta)$ and $s(\theta)$ for sufficiently "weak" patents. This comparison, which is rather technical, is presented in Appendix B.

5.3 Are "weak" patents always overcompensated?

Now that we have characterized the optimal license offer for "weak" patents, we can now address one of the main questions raised in this paper: are "weak" patents always overcompensated?

5.3.1 Comparison of the equilibrium royalty with the benchmark

We compare first the optimal royalty rate when litigation is deterred at equilibrium, i.e. $r^*(\theta) = r(\theta)$, with the expected value of the maximal royalty rate in case of litigation, i.e. $r^e(\theta) = \theta\epsilon$. The comparison of $r^*(\theta)$ and $\theta\epsilon$ when litigation is deterred is made when an unsuccessful challenger is not viable (e.g. because the innovation is large or the competitive environment is tough). Defining $\eta(\epsilon) = \frac{\varepsilon\left|\pi_2^e(n,c-\epsilon)\right|}{\pi^e(n,c-\epsilon)}$ as the elasticity of a firm's profit with respect to cost reduction when all firms benefit from this reduction, we get the following result:

Proposition 6 *Assume that the parameters of the model are such that:*

i/ the innovation size is sufficiently large or the competition intensity is high enough, i.e. $\pi^i(n-1, c-\epsilon) = 0$, *and*

ii/ sufficiently "weak" patents are licensed to all firms, i.e. $r(\theta) \geq s(\theta)$.

The following statements hold:

If $\eta(\epsilon) < 1$ *then* $r^*(\theta) = r(\theta) \geq \theta\epsilon$ *for sufficiently "weak" patents*

If $\eta(\epsilon) > 1$ *then* $r^*(\theta) = r(\theta) \leq \theta\epsilon$ *for sufficiently "weak" patents*

Proof. See Appendix A. ■

This proposition states that the elasticity of a firm's profit (in a symmetric oligopoly) with respect to cost reduction $\eta(\epsilon)$ plays a crucial role in the comparison of the optimal royalty rate $r^*(\theta)$ with the "fair" benchmark $r^e(\theta) = \theta\epsilon$. The intuition behind the result is that a low value of this elasticity entails a low (negative) effect of an increase in the royalty rate on the firms' profit when they all purchase a license. Under such conditions, the patent holder may be able to impose a high royalty rate. In particular, the level of the royalty rate may be greater than the benchmark level $r^e(\theta)$. However, if the elasticity of the profits with respect to cost reduction is high, the patent holder may not be able to overcharge the license with a royalty higher than $\theta\epsilon$ without triggering a challenge: such royalty could result in a relatively weak profit for the licensees hence making the challenge option more attractive for them. Thus, in a situation where the patent holder prefers to deter a challenge, the per-unit royalty must be less than $\theta\epsilon$ when the elasticity of a firm's profit with respect to cost reduction is greater than one. One can easily check that the comparisons between $r(\theta)$ and $\theta\epsilon$ obtained directly in the two examples analyzed in section 2 lead to the same outcomes if we use the more general proposition 6.[32]

5.3.2 Comparison of $P^*(\theta)$ with $\theta P^*(1)$

We intend now to compare the patent holder' licensing revenues $P^*(\theta)$ under the shadow of litigation and the (fair) benchmark $\theta P^*(1)$ which corresponds to the expected licensing revenues if the patent validity were determined before the licensing agreement takes place. To make this comparison we need to define, for any $\theta \leq \tilde{\theta}$, a threshold $v(\theta)$ as the unique solution to the equation $nrq^e(n, c - \epsilon + r) = \theta P^*(1)$.[33] Note that $v(\theta) = s(\theta)$ if and only $P^*(1) = P^*_{-n}(1)$, that is, if and only if the optimal license offer in the case of an iron-clad patent $(\theta = 1)$ is not accepted by all firms in the industry. Otherwise, it holds that $P^*(1) > P^*_{-n}(1)$ which entails that $v(\theta) > s(\theta)$. The comparison of $P^*(\theta)$ and $\theta P^*(1)$ is stated in the next proposition:

Proposition 7 *Consider sufficiently "weak" patents, i.e. $\theta \leq \tilde{\theta}$.*

1. When $P^*(1) = P^*_{-n}(1)$, i.e. it is optimal for an iron-clad patent holder to offer a contract which is not accepted by all the firms, the following statements hold:

1.a. If $r(\theta) \leq s(\theta)$ then $P^*(\theta) = \theta P^*(1)$

1.b. If $r(\theta) > s(\theta)$ then $P^*(\theta) > \theta P^*(1)$

2. When $P^*(1) > P^*_{-n}(1)$, i.e. it is optimal for an iron-clad patent holder to offer a contract accepted by the whole set of firms, the following statements hold:

2.a. If $r(\theta) < v(\theta)$ then $P^*(\theta) < \theta P^*(1)$

2.b. If $r(\theta) = v(\theta)$ then $P^*(\theta) = \theta P^*(1)$

2.c. If $r(\theta) > v(\theta)$ then $P^*(\theta) > \theta P^*(1)$

Proof. See Appendix A. ■

---

[32] Under the specification of example 1, we obtain $\eta(\epsilon) = \frac{2\epsilon}{a-c+\epsilon}$ which leads to $\eta(\epsilon) \leq 1 \Leftrightarrow \epsilon \leq a - c$. Therefore:
- If $\epsilon \in \left[\frac{a-c}{n-1}, a - c\right]$ then $\eta(\epsilon) \leq 1$ which yields $r_2(\theta) > \theta\epsilon$ for θ sufficiently small.
- If $\epsilon \geq a - c$ then $\eta(\epsilon) > 1$ which yields $r_2(\theta) < \theta\epsilon$ for θ sufficiently small. These results are consistent with those obtained by direct comparison in example 1.

Under the specification of example 2, the equilibrium profit when both firms produce at marginal cost $c - \epsilon$ is $\pi^e(2, c - \epsilon) = \frac{1-\gamma}{(2-\gamma)^2(1+\gamma)}(a - c + \epsilon)$ which yields $\eta(\epsilon) = \frac{2\epsilon}{a-c+\epsilon}$ Thus, the results obtained through direct comparison hold when we use the previous proposition:
- If $\gamma \geq \bar{\gamma}$, then $\eta(\epsilon) \leq 1$ whenever $\bar{\epsilon}(\gamma) < \epsilon \leq a - c$, and $\eta(\epsilon) > 1$ whenever $\epsilon > a - c$. Thus, under the former case, $r_2(\theta) > \theta\epsilon$ for θ sufficiently small, while under the latter $r_2(\theta) < \theta\epsilon$ for θ sufficiently small.
- If $\gamma < \bar{\gamma}$, then $\eta(\epsilon) > 1$ for any $\epsilon \geq a - c$. Here, it always true that $r_2(\theta) < \theta\epsilon$ for θ sufficiently small.

[33] The existence and uniqueness of $v(\theta)$ can be established in the same fashion as $s(\theta)$

This proposition shows that licensing "weak" patents may lead to overcompensation as well as undercompensation relative to the expected licensing revenues $\theta P^{*}(1)$ in case validity is determined prior to licensing . Two questions matter to determine the outcome of such a comparison: i/ Would it be optimal for the patent holder to license every firm or only a subset of them, if the patent were iron-clad? ii/ What is the location of the per-unit royalty $r(\theta)$ that deters litigation relative to the thresholds $s(\theta)$ and $v(\theta)$ derived from the situations where the patent holder is indifferent between deterring any litigation, and respectively licensing to less than n firms or obtaining the expected revenue under litigation? The answer to the first question has been investigated in the literature on licensing iron-clad patents (Sen and Tauman, 2007). The answer to the second question is the result of a comparison: it is only when $r(\theta)$ is above the thresholds that the licensing revenue from a weak patent punches above the benchmark revenue corresponding to the patent's weight. Note that undercompensation can occur at equilibrium even when litigation is deterred. This happens when $s(\theta) \leq r(\theta) < v(\theta)$, that is, when the licensing revenues from the optimal license deterring litigation is greater than the expected licensing revenues from litigation $\theta P^{*}_{-n}(1)$ but less than the benchmark licensing revenues $\theta P^{*}(1)$.

6 Economic policy levers

In this section, we discuss some policy levers that can be used to alleviate the concerns raised by licensing "weak" patents. First it must be clear that the patent quality problem has several dimensions related to the processes that occur in different phases, running from application by innovators, prosecution by patent examiners, private settlements between the concerned agents to solve litigation issues and avoid trials, until legal enforcement by different courts to reach decisions on the private suits brought by different agents . All these phases are complex, evolve in time, differ among countries and largely depend on the intellectual property law adopted by legislative bodies. Thus, it is very difficult if not impossible to discuss all the aspects raised by the "patent quality" problem or in other words the "weakness" notion. But remember what has been said in the introduction: "bad patents" or patents of "weak" quality are not only patents that cover non-novel or obvious inventions. They concern also inventions that may be invalidated by a court for other reasons such as the patentable subject matter, the utility criterion or any other ambiguities that exist in the patent law. Therefore, the modest objective of this section is to derive some policy suggestions to improve the performance of the licensing process, provided that the "weakness" of some patents and the low individual incentives to challenge them are recognized.

The first suggestion is to prevent a patent holder from refusing to license its right to an agent who would have tempted unsuccessfully to dispute the validity. We argue in the first sub-section that the effect of such prevention would be to reduce the level of the per-unit royalty acceptable by every firm, anticipating that such royalty should be renegotiated in case of an unsuccessful challenge. The second suggestion is to encourage a set of agents to dispute collectively the validity of a patent rather than to restrict this possibility to each of them individually. The point we develop in the second sub-section is that, by ruling out the positive externality that an agent offers to competitors when he or she disputes alone the patent validity, a collective challenge rules out the possibility that a "weak" patent could impose a royalty rate higher than the benchmark.

6.1 Preventing license refusal to an unsuccessful challenger

So far we have assumed that in case of litigation, an unsuccessful challenger produces with marginal cost c because the patent holder refuses to sell him or her a license. Whether such a commitment to refuse a license to an unsuccessful challenger is credible or not must be discussed. From the challenger's perspective this commitment is equivalent to an offer of a new licensing contract involving a royalty rate $\bar{r} = \epsilon$. However, from the patent holder's perspective, this equivalence does not hold. Moreover a situation where an unsuccessful challenger is offered a new licensing contract involving a royalty rate $\bar{r} < \epsilon$ may be preferred by the patent holder to a situation where it is offered a contract based on $\bar{r} = \epsilon$. Such an issue is important since a potential challenger will take the decision whether to accept the license or contest the patent validity, anticipating what would happen if the patent is validated. If patent law prevents the patent holder from refusing to license an unsuccessful challenger, then the commitment of the former not to renegotiate with the latter after the challenge is undermined.

Formally if we allow for renegotiation when $(n-1)$ firms accept a licensing contract based on a royalty rate r and the remaining firm challenges the patent unsuccessfully, then the patent holder will offer to the challenger a contract involving a royalty rate $\bar{r} \in [0, \epsilon]$ that maximizes its licensing revenues denoted $P(r, \bar{r})$ and given by:

$$P(r, \bar{r}) = (n-1)\, r q^{L}\left(c-\epsilon+r, c-\epsilon+\bar{r}\right) + \bar{r} q^{NL}\left(c-\epsilon+r, c-\epsilon+\bar{r}\right)$$

where[34] $q^{L}\left(c-\epsilon+r, c-\epsilon+\bar{r}\right)$ denotes the equilibrium quantity produced by each of the $(n-1)$ firms that accepted initially the license offer r and $q^{NL}\left(c-\epsilon+r, c-\epsilon+\bar{r}\right)$ is the equilibrium quantity produced by the unsuccessful challenger who produces at marginal cost $c-\epsilon+\bar{r}$. If $\bar{r}(r)$ is the royalty rate that maximizes $P(r, \bar{r})$ with respect to $\bar{r}$, a licensing contract involving a royalty rate r will be accepted by all the firms if and only if:

$$\pi(c-\epsilon+r, c-\epsilon+r) \geq \theta\pi(c-\epsilon+\bar{r}\left(r\right), c-\epsilon+r) + (1-\theta)\pi(c-\epsilon, c-\epsilon) \tag{6}$$

[34] A notation different from the one used in previous sections is needed here since an unsuccessful challenger produces now at marginal cost $c-\epsilon+\bar{r}$ and not at marginal cost c.

Since $\bar{r}(r) \leq \epsilon$ we have $\pi(c-\epsilon+\bar{r}(r), c-\epsilon+r) \geq \pi(c, c-\epsilon+r)$ which entails that constraint (6) is (weakly) more stringent than (3). More specifically, a royalty rate r could be accepted if the patent holder commits to refuse a license to a challenger or license him at $\bar{r}=\epsilon$, but not accepted if he cannot commit. This implies that the maximal royalty rate the patent holder can make the n firms pay is (weakly) smaller when renegotiation of a licensing contract (after patent validation) is introduced.[35]

6.2 Encouraging collective challenges

Suppose that at stage 2 the firms cooperatively agree on whether to buy the license or refuse it and challenge all together the patent validity.[36] In this case, the firms will cooperatively accept a licensing contract involving a royalty rate r if and only if:

$$\pi^e(n, c-\epsilon+r) \geq \theta\pi^e(n, c) + (1-\theta)\pi^e(n, c-\epsilon)$$

The function w defined by $w(r) = \pi^e(n, c-\epsilon+r) - \theta\pi^e(n, c) - (1-\theta)\pi^e(n, c-\epsilon)$ is continuous, strictly decreasing (by A3) and satisfies the conditions $w(0) \geq 0$ and $w(\epsilon) \leq 0$. Hence there exists a unique solution $r^c(\theta) \in [0, \epsilon]$ to the equation $w(r) = 0$, and the inequality $w(r) \geq 0$ is equivalent to $r \leq r^c(\theta)$. This means that all firms cooperatively accept to buy a license at a royalty rate r if and only if $r \leq r^c(\theta)$.

We establish in the next proposition that the maximal royalty deterring a collective challenge is lower than the maximal royalty that deters individual challenge, which is not surprising because the free-riding problem that arises when the decision to challenge is made non-cooperatively disappears when challenges are conducted collectively. The proposition gives also a condition under which the royalty rate deterring a collective challenge $r^c(\theta)$ is lower than the expected royalty rate in case of litigation $\theta\epsilon$.

Proposition 8 *The maximal royalty rate deterring a collective challenge is lower than the non-cooperatively royalty rate accepted by all firms :* $r^c(\theta) \leq r(\theta)$ *for all* $\theta \in [0, 1]$. *Moreover, the function* $r^c(\theta)$ *satisfies the following properties:*

i/ $r^c(\theta)$ *is increasing over* $[0, 1]$ *and* $r^c(0) = 0$, $r^c(1) = \epsilon$,

ii/ $r^c(\theta)$ *is convex over* $[0, 1]$ *if (and only if) the function* $x \to \pi^e(n, x)$ *is convex over* $[c-\epsilon, c]$ *and in this case* $r^c(\theta) \leq r^e(\theta) = \theta\epsilon$

[35] We can show that under Cournot competition with linear demand, it is possible that the maximal royalty rate accepted by all firms if renegotiation is possible is below the benchmark $\theta\epsilon$ whereas the maximal royalty rate if renegotiation is not possible is above the benchmark $\theta\epsilon$ under the same conditions (computations available upon request).

[36] Firms are allowed to challenge collectively the validity of a patent, at least in the US. An example is the PanIP Group Defense Fund which is a coalition of fifteen e-retailers that has been created to invalidate a patent covering some key aspects of electronic commerce, hold by Pangea Intellectual Properties (US patent number 5.576.951).

Proof. See Appendix A. ■

Note that the convexity of $x \to \pi^e(n, x)$ holds in a wide range of competitive environments including Cournot competition with linear or iso-elastic demand as well as differentiated Bertrand oligopoly with linear demand. Hence the fact that the equilibrium royalty rate may exceed the benchmark $\theta\epsilon$ is mainly due the free-riding problem. Getting rid of the latter by encouraging collective challenges may then be a solution to reduce the potentially high market power of "weak" patent holders.

7 Conclusion

The consequences of licensing "weak" patents have been examined in this paper by addressing the following question: to what extent licensing a patent that has a high probability to be invalidated by a court (the so-called "weak" or "bad" patent) if it is challenged may favor the patent holder when the license occurs prior to the patent validity determination? This question was addressed by FS under the heading "How strong are weak patents?" We have shown in this paper that the answer given by FS is very sensitive to two implicit restrictions made in their analysis: i/ a weak patent covers necessarily a small size innovation; ii/ the patent holder always chooses a licensing contract that deters any litigation, i.e. sells the license at a price accepted by every firm. The first restriction is unjustified insofar as the weakness of a patent may arise from other reasons than the standard novelty and non-obviousness criteria. The second restriction is no more justified as it rests on conditions that are not always satisfied. Removing these two restrictions and keeping the same structure of a two-part tariff for the licensing contract as in FS, we have developed a framework that allows a more complete characterization for licensing a weak patent, whatever the size of the invention it protects.

We have shown that the optimal structure for licensing a weak patent depends on the level of the per-unit royalty that deters litigation. When this level is above a defined threshold, the optimal licensing scheme is a pure per-unit royalty that deters litigation, and this confirms FS result. But, when this level is below the defined threshold, the patent owner prefers to sell its license to a subset of firms, at the optimal two-part tariff that triggers litigation. It is precisely when the royalty rate acceptable by all the firms in the downstream industry is too low that the holder of a weak patent may prefer to sell at a higher royalty rate. Second, we provided a necessary and sufficient condition to check whether "weak" patents are overcompensated or not. This condition gives a key role to the cost reduction magnitude and the intensity of competition in the downstream market. It states that the per-unit royalty used in licensing a "weak" patent punches above the patent strength if and only if the elasticity of the profit function with respect to the cost reduction magnitude is lower than one. We have shown that in two standard oligopoly frameworks, this condition is not satisfied for a sufficiently large cost reduction. Thus, the overcompensation result in FS crucially depends on the size of the innovation. Finally, we analyze the effects of two policy levers that could reduce the potential harm raised by licensing "weak" patents. One policy suggestion explores the idea that, due to the positive externality that benefits to competitors, individual incentives to challenge a patent validity are small. Therefore, allowing and favouring collective challenge may help to solve the problem. The other suggestion explores the idea that, when the cost reduction magnitude is high or when the competition in the downstream market is tough, an unsuccessful challenger is threatened to be evicted from the market if deprived from the new technology, exactly in the same way as he or she would be forced to shut down after an injunction on the alleged infringement. One way to solve the problem is to prevent a refusal to sell a license to an unsuccessful challenger.

These policy suggestions do not exhaust the different ways to alleviate the problem raised by licensing "weak" patents. As we explained above, the patent quality problem is indeed much more complex than suggested by the previous analysis. Since the patent system involves a two-tier process combining patent office examination and possible challenge of the granted patent validity before a court, one has to explore the two possible approaches to the problem raised by the existence of "weak" patents and their possible bad consequences. The first approach relies on the FS view that "weak" patents are mainly patents that cover non novel or obvious innovations. Therefore, according to this view, one way to alleviate the patent quality problem is to improve the screening process inside the patent office itself through the strengthening of the novelty and inventivity standards, turning back the Lemley's "rational ignorant patent office principle" (Lemley, 2001). This approach (IDEI, 2006) could be interesting, particularly when the patent strength is no more a common knowledge parameter. The patent office could thus propose to any applicant a menu involving the choice between paying an extra fee to obtain a thorough examination process, signalling thus a high quality of the examination process recognizable by a court, or paying a lower fee to simply obtain a "standard" examination process that may signal the potential weakness of the patent. Designing an efficient mechanism to implement such a procedure is an important task for future work (see Chiou, 2008 for a first step in this direction).

The second approach, privileged in this paper, starts from the observation that the weakness of a patent may arise from different sources, more or less related to the ambiguities of the patent law, the evolution of its content and the institutions (like the CAFC in the US) and not only because the overloading charges of a patent office prevent patent examiners to weed out applications that lack novelty or inventivity. Therefore, according to this approach, the solution is to find some ways to encourage third parties to bring to a court pieces of evidence in order to challenge the validity of these presumably "weak" patents. For instance, the post-grant opposition in Europe seems to play this role in a more appropriate way than the post-grant reexamination in the United States (see Graham *et al.*, 2003). Similarly, injunction against infringement seems to be a less desirable remedy than damages (Hylton, 2006). Therefore, giving more incentives to potential licensees to challenge a patent validity is necessary in this perspective. We have suggested two policy levers: the renegotiation of the licensing contract with an unsuccessful challenger and the cooperative approach among potential licensees to collectively accept or refuse a licensing contract.

We think that these different approaches to conceptualize the notion of patent weakness are more complementary than conflicting and the policy recommendations should combine the levers that result from them. It is certainly not sufficient to concentrate effort on refining the novelty and inventivity standards but it would be a mistake not to do it.

8 References

Amir, R. and J. Wooders, 2000, One-way spillovers, endogenous innovator/imitator roles, and research joint ventures, *Games and Economic Behavior*, 31, 1-25.

Aoki, R. and J-L. Hu, 1999, Licensing versus litigation: the effect of the legal system on incentives to innovate, *The Journal of Economics and Management Strategy*, 8, 133-160.

Arrow, K., 1962, Economic welfare and the allocation of resources for invention, in R. Nelson (Ed.), *The rate and direction of incentive activity*, Princeton, Princeton University Press.

Ayres, J. and P. Klemperer, 1999, Limiting patentees market power without reducing innovation incentives: the perverse benefits of uncertainty and non-injunctive remedies, *Michigan Law Review,* Vol. 97, 985-1033.

Barton, J., 2003, "Nonobviousness", *IDEA, The Journal of Law and Technology*, 43, 475-508

Bebchuk, L., 1984, Litigation and settlement under imperfect information, *The Rand Journal of Economics*, 15,3, 404-415.

Boone, J., 2001, Intensity of competition and the incentives to innovate, *International Journal of Industrial Organization,* 19, 705-726.

Bousquet, A., H. Cremer, M. Ivaldi and M. M. Wolkowicz, 1998, Technology transfer with moral hazard, *International Journal of Industrial Organization*, 16, 535-554.

Brocas, I., 2006, Designing auctions in R&D: optimal licensing of an innovation, *Topics in Economic Analysis & Policy*, 6, Issue 1, Article 11.

Caillaud, B. and A. Duchêne, 2005, Patent office in innovation policy: No-body's perfect, mimeo available at http://ceco.polytechnique.fr/Colloques/ETAPE/duchene.pdf.

Chiou, J-Y., 2008, The patent quality control process: Can we afford an (rationally) ignorant patent office?, mimeo.

Crampes, C. and C. Langinier, 2002, Litigation and settlement in patent infringement cases, *Rand Journal of Economics*, 33(2), 258-274.

Encaoua, D. and A. Hollander, 2002, Competition policy and innovation, *Oxford Review of Economic Policy,* 18, 1, 63-79.

Farrell, J. and R. Merges, 2004, Incentives to challenge and defend patents: why litigation won't reliably fix patent office errors and why administrative patent review might help, *Berkeley Technology Law Journal*, 19, 2, 943-970.

Farrell, J. and C. Shapiro, 2008, How strong are weak patents?, *American Economic Review,*98:4, 1347-1369

Février, P and L. Linnemer, 2004, Idiosyncratic shocks in a asymetric Cournot oligopoly, *International Journal of Industrial Organization*, 22, 6, 835-848.

Graham, S., B. Hall, D. Harhoff and D. Mowery, 2003, Post-issue patent "quality control": a comparative study of US patent re-examinations and European patent oppositions, in W. Cohen and S. Merrill, eds.: *Patents in the Knowledge-Based Economy,* The National academic Press, Washington, D.C.

Guellec, D. and B. von Pottelsberghe de la Potterie, 2007, *The Economics of the European Patent System*, Oxford University Press, Oxford.

Hunt, R., 2001, You can patent that? Are patents to computer programs and business methods good for the New Economy?, *Federal Reserve Bank of Philadelphia Business Review,* Quarter 1, 5-15

Hunt, R., 2008, Business method patents and US financial services, WP n° 08-10, Research Department, Federal Reserve Bank of Philadelphia

Hylton, K., 2006, Property rules and liability rules, once again, *Review of Law and Economics, 2,*137-191.

IDEI, 2006, *Objectives and Incentives at the European Patent Office*, Institut d'Economie Industrielle, Toulouse

Jaffe, A. and J. Lerner, 2006, *Innovation and its discontents: How our broken patent system is endangering innovation and progress and what to do about it*, Princeton University Press: Princeton

Kamien, M., 1992, Patent licensing, In: Aumann, R.J., S. Hart (Eds.) *Handbook of Game Theory with Economic Applications*, Elsevier Science, North Holland, ch. 11, 331-354.

Kamien, M., S. Oren and Y. Tauman, 1992, Optimal licensing of cost-reducing innovation, *Journal of Mathematical Economics*, 21, 483-508.

Kamien, M., and Y. Tauman, 1984, The private value of a patent: a game-theoretic analysis, *Z. Nationalőkon*, *Journal of Economics,* 4 (Supplement), 93-118.

Kamien, M. and Y. Tauman, 1986, Fees versus royalties and the private value of a patent, *Quarterly Journal of Economics*, 101, 471-491.

Katz, M. and C. Shapiro, 1985, On the licensing of innovations, *Rand Journal of Economics*, 16, 504-520.

Katz, M. and C. Shapiro, 1987, R&D rivalry with licensing or imitation, *American Economic Review,* 77, 3, 402-420

Kimmel, S., 1992, Effects of cost changes on oligopolists' profits, *Journal of Industrial Economics*, 40, 4, 441-449.

Lanjouw, J. and M. Schankerman, 2001, Characteristics of patent litigation: A window on competition, *Rand Journal of Economics,* 32, 1, 129-151

Lanjouw, J. and M. Schankerman, 2004, Protecting intellectual property rights: Are small firms handicapped?, *Journal of Law and Economics*, 47, 1, 45-74

Lemley, M. and C. Shapiro, 2005, Probabilistic patents, *Journal of Economic Perspectives*, 19, 2, 75-98.

Lemley, M. and C. Shapiro, 2007, Patent hold-up and royalty stacking, *Texas Law Review*, 85, 1990-2049.

Macho-Stadler, I., X. Martinez-Giralt and D. Perez-Castrillo, 1996, The role of information in licensing contract design, *Research Policy*, 25, 25-41.

Merges, R. and R. Nelson, 1990, On the complex economics of patent scope, *Columbia Law Review*, 90, 839-916.

Meurer, M., 1989, The settlement of patent litigation, *The Rand Journal of Economics*, 20, 1, 77-91.

Muto, S., 1993, On licensing policies in Bertrand competition, *Games and Economic Behavior*, 5, 257-267.

Poddar, S. and U. Sinha, 2004, On patent licensing and spatial competition, *Economic Record*, 80, 208-218.

Rockett, K., 2008, Property rights and invention, DP n° 663, Department of Economics, University of Essex, to appear in *The Handbook of Economics and Technical Change*, Bronwyn Hall and Nathan Rosenberg, Editors.

Reinganum, J. and L. Wilde, 1986, Settlement, litigation and the allocation of litigation costs, *The Rand Journal of Economics*, 17, 4, 557-566.

Saracho, A., 2002, Patent licensing under strategic delegation, *Journal of Economics and Management Strategy*, 11, 225-251.

Sen, D., 2005, Fee versus royalty reconsidered, *Games and Economic Behavior*, 53, 141-147.

Sen, D. and Y. Tauman, 2007, General licensing schemes for a cost-reducing innovation, *Games and Economic Behavior*, 59, 163-186.

Scotchmer, S., 2004, *Innovation and Incentives*, The MIT Press, Cambridge, Ma.

9 Appendix

9.1 Appendix A: Proofs

Proof of Proposition 1

The situation where the n firms accept the licensing contract F is a Nash equilibrium if and only if:

$$\pi^e(n, c-\epsilon+r) - F \geq \theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\pi^e(n, c-\epsilon)$$

which can be rewritten as:

$$F \leq \pi^e(n, c-\epsilon+r) - \theta\pi^i(n-1, c-\epsilon+r) - (1-\theta)\pi^e(n, c-\epsilon)$$

that is

$$F \leq \Psi_n(r, \theta)$$

A situation where $n-1$ firms accept the licensing contract and one firm does not is a Nash equilibrium (of stage 2) if and only if:

$$\theta\pi^i(n-1, c-\epsilon+r) + (1-\theta)\,\pi^e(n, c-\epsilon) \geq \pi^e(n, c-\epsilon+r) - F \tag{7}$$

and

$$\theta[\pi^e(n-1, c-\epsilon+r) - F] + (1-\theta)\,\pi^e(n, c-\epsilon) \geq \theta\pi^i(n-2, c-\epsilon+r) + (1-\theta)\,\pi^e(n, c-\epsilon) \tag{8}$$

Condition (7) means that the one firm that does not accept the licensing contract and challenges the patent validity does not find it optimal to unilaterally deviate by accepting the licensing contract. Condition (8) means that none of the $n-1$ firms which accept the licensing contract find it optimal to unilaterally deviate by refusing the contract. When the number of firms accepting the contract is strictly less than n, litigation occurs which entails that the firms accepting the contract pay the fixed fee F and the running royalties only if the patent validity is upheld, which happens with probability θ. With the complementary probability $1-\theta$, the patent is invalidated and all the firms get the same profit namely $\pi^e(n, c-\epsilon)$. It is straightforward to show that conditions (7) and (8) are equivalent to the following double inequality:

$$\pi^e(n, c-\epsilon+r) - \theta\pi^i(n-1, c-\epsilon+r) - (1-\theta)\pi^e(n, c-\epsilon) \leq F \leq \pi^e(n-1, c-\epsilon+r) - \pi^i(n-2, c-\epsilon+r)$$

that is:

$$\Psi_n(r, \theta) \leq F \leq F_{n-1}(r)$$

Note that for all $\theta \in [0, 1]$, we have $\Psi_n(r, \theta) \leq \Psi_n(r, 1) = F_n\,(r) \leq F_{n-1}(r)$ due to assumption

A5, which ensures that the interval $[\Psi_n(r,\theta), F_{n-1}(r)]$ is not empty.
A situation where only $k \in \{1,2,...,n-2\}$ firms accept the licensing contract is a Nash equilibrium of stage 2 if and only if:

$$\theta\left(\pi^e(k,c-\epsilon+r)-F\right)+(1-\theta)\,\pi^e(n,c-\epsilon)\geq\theta\pi^i(k-1,c-\epsilon+r)+(1-\theta)\,\pi^e(n,c-\epsilon) \quad (9)$$

and

$$\theta\pi^i(k,c-\epsilon+r)+(1-\theta)\,\pi^e(n,c-\epsilon)\geq\theta\left(\pi^e(k+1,c-\epsilon+r)-F\right)+(1-\theta)\,\pi^e(n,c-\epsilon) \quad (10)$$

Condition (9) means that none of the k firms accepting the licensing contract finds it optimal to unilaterally deviate by refusing the contract and condition (10) means that none of the $n-k$ firms refusing the licensing contract finds it optimal to unilaterally deviate by accepting the contract. It is easy to see that conditions (9) and (10) can be combined into the following double inequality that does not depend on θ:

$$\pi^e(k+1,c-\epsilon+r,)-\pi^i(k,c-\epsilon+r)\leq F\leq\pi^e(k,c-\epsilon+r)-\pi^i(k-1,c-\epsilon+r)$$

that is:

$$F_{k+1}(r)\leq F\leq F_k(r)$$

Finally, a situation where no firm accepts the licensing contract is a Nash equilibrium if and only if:

$$\theta\pi^i(0,c-\epsilon+r)+(1-\theta)\,\pi^e(n,c-\epsilon)\geq\theta\left(\pi^e(1,c-\epsilon+r)-F\right)+(1-\theta)\,\pi^e(n,c-\epsilon)$$

which can be rewritten as:

$$F\geq\pi^e(1,c-\epsilon+r)-\pi^i(0,c-\epsilon+r)=F_1(r)$$

Proof of Corollary 1

Combining Proposition 1 under the restriction $F=0$ with the fact that $F_k(\epsilon)=0$ for all $1\leq k\leq n-1$ yields the result.

Proof of Lemma 1
Denote $h(r,\theta)=\pi^e(n,c-\epsilon+r)-(1-\theta)\,\pi^e(n,c-\epsilon)$ and consider, for a given $\theta\in\left[0,\hat{\theta}\right]$, the equation $h(r,\theta)=0$. Note that $h(0,\theta)=\theta\pi^e(n,c-\epsilon)\geq0$ and $h(\hat{r},\theta)=\left(\theta-\hat{\theta}\right)\pi^e(n,c-\epsilon)\leq0$ for any $\theta\in\left[0,\hat{\theta}\right]$. Since $h(.,\theta)$ is continuous and strictly decreasing over $[0,\hat{r}]$ (due to A1 and A3), we can use the intermediate value theorem to state that the equation $h(r,\theta)=0$ has a unique solution in r, which we denote $r_1(\theta)$, over $[0,\hat{r}]$. Moreover, assumption A1 implies

that $h(.,.)$ is continuously differentiable over $[0, \hat{r}] \times \left[0, \hat{\theta}\right]$, which allows to state (using the implicit function theorem for instance) that $r_1(\theta)$ is differentiable over $\left[0, \hat{\theta}\right]$ and

$$r_1'(\theta) = \frac{-\pi^e(n, c-\epsilon)}{\pi_2^e(n, c-\epsilon+r_1(\theta))}$$

This implies that $r_1'(\theta) > 0$ since $\pi_2^e(n, c-\epsilon+r_2(\theta)) < 0$ by A3. Therefore $r_1(\theta)$ increases in the patent strength θ over $\left[0, \hat{\theta}\right]$. Furthermore, it is obvious that $r_1(0) = 0$ and we derive from the definition of $\hat{\theta}$ that $r_1(\hat{\theta}) = \hat{r}$.

Proof of Lemma 2

Consider, for a given $\theta \in \left[\hat{\theta}, 1\right]$, the equation $\Psi_n(r, \theta) = 0$ where $\Psi_n(r, \theta) = \pi^e(n, c-\epsilon+r) - \theta\pi^i\left(n-1, c-\epsilon+r\right) - (1-\theta)\,\pi^e(n, c-\epsilon)$ is continuous and strictly decreasing in r over $[0, \epsilon]$ due to assumptions A1, A2 and A3. Moreover $\Psi_n(0, \theta) = \left(\theta - \hat{\theta}\right)\pi(c-\epsilon, c-\epsilon) \geq 0$ for any $\theta \in \left[\hat{\theta}, 1\right]$ and $\Psi_n(\epsilon, \theta) = (1-\theta)\left[\pi^e(n, c) - \pi^e(n, c-\epsilon)\right] \leq 0$. Using the intermediate value theorem, we can then state that the equation $\Psi_n(r, \theta) = 0$ has a unique solution in r over the interval $[\hat{r}, \epsilon]$, that we denote by $r_2\left(\theta\right)$. Further, $\Psi_n(r, \theta) = \pi(c-\epsilon+r, c-\epsilon+r) - \theta\pi(c, c-\epsilon+r) - (1-\theta)\pi(c-\epsilon, c-\epsilon)$. Note that $g(\theta, \epsilon) = (1-\theta)\left[\pi(c, c) - \pi(c-\epsilon, c-\epsilon)\right] \leq 0$ for any $\theta \in \left[\hat{\theta}, 1\right]$ (by A4). Moreover, the function $\Psi_n(., \theta)$ is continuous and strictly increasing over $[\hat{r}, \epsilon]$. Then, using the intermediate value theorem, we state that the equation $\Psi_n(r, \theta) = 0$ has a unique solution in r for any $\theta \in \left[\hat{\theta}, 1\right]$, which we denote by $r_1\left(\theta\right)$. Furthermore, assumption A1 ensures that $\Psi_n(.,.)$ is continuously differentiable over $[\hat{r}, \epsilon] \times \left[\hat{\theta}, 1\right]$, which allows to state that $r_2(\theta)$ is differentiable over $\left[\hat{\theta}, 1\right]$ and:

$$r_2'(\theta) = \frac{\pi^i\left(n-1, c-\epsilon+r_2(\theta)\right) - \pi^e(n, c-\epsilon)}{\pi_2^e(n, c-\epsilon+r_2(\theta)) - \theta\pi_2^i\left(n-1, c-\epsilon+r_2(\theta)\right)} \tag{11}$$

The denominator is negative due to A2 and A3. The numerator is negative as well because $\pi_2^i\left(n-1, c-\epsilon+r_2(\theta)\right) \leq \pi_2^i\left(n, c-\epsilon+r_2(\theta)\right) = \pi_2^e\left(n, c-\epsilon+r_2(\theta)\right) < \pi^e(n, c-\epsilon)$. The first inequality follows from $r_2(\theta) \leq \epsilon$ and the second one from A3. Thus, $r_2'(\theta) > 0$, that is $r_2(\theta)$ is strictly increasing in the patent strength θ over $\left[\hat{\theta}, 1\right]$. Furthermore, it is obvious that $r_2(\epsilon) = 1$ and we derive from the definition of $\hat{\theta}$ that $r_2(\hat{\theta}) = \hat{r}$.

Proof of Proposition 2

We distinguish two cases:

Case 1: $\theta \in \left[0, \hat{\theta}\right]$.

Consider a royalty rate $r \leq \hat{r}$. In this case, inequality (1) is equivalent to $h(r, \theta) \geq 0$ where h has been defined in the proof of lemma 1. Since $h(r, \theta)$ is decreasing in r, $h(r, \theta) \geq 0$ if and only if $r \leq r_1\left(\theta\right)$ where $r_1\left(\theta\right)$ is defined in lemma 1.

Consider now $r > \hat{r}$. Since $\Psi_n(r,\theta)$ and $h(r,\theta)$ are decreasing in r and $r_1(\theta) \leq \hat{r}$ for $\theta \in \left[0,\hat{\theta}\right]$ then for any $r > \hat{r}$, it holds that $\Psi_n(r,\theta) \leq \Psi_n(\hat{r},\theta) = h(\hat{r},\theta) < h(r_1(\theta),\theta) = 0$ which shows that such $r > \hat{r}$ will not be accepted by all firms. Hence, for any $\theta \in \left[0,\hat{\theta}\right]$, inequality (1) holds if and only if $r \leq \min(\hat{r}, r_1(\theta)) = r_1(\theta)$.

Case 2: $\theta \in \left[\hat{\theta},1\right]$.

Consider a royalty rate $r \leq \hat{r}$. In this case, inequality (1) is equivalent to $h(r,\theta) \geq 0$. Since $\Psi_n(r,\theta)$ and $h(r,\theta)$ are decreasing in r, it holds that $h(r,\theta) \geq h(\hat{r},\theta) = \Psi_n(\hat{r},\theta) \geq \Psi_n(r_1(\theta),\theta) = 0$.

Consider now a royalty rate $r > \hat{r}$. Since the function $\Psi_n(r,\theta)$ is decreasing in r, $\Psi_n(r,\theta) \geq 0$ if and only if $r \leq r_2(\theta)$.

Hence, for any $\theta \in \left[\hat{\theta},1\right]$, inequality (1) holds if and only if $r \leq \max(\hat{r}, r_1(\theta)) = r_1(\theta)$.

Proof of Lemma 3

The existence and unicity can be proven as in lemma 1. However, a difference with lemma 1 is that we do not need to restrict to $\theta \in \left[\hat{\theta},1\right]$ to get the differentiability property. Indeed, as $r \to \pi^i(n-1, c-\epsilon+r)$ remains strictly positive for any $r \geq 0$, it is differentiable over $[0,\epsilon]$ due to A1. This ensures the differentiability of $\Psi_n(r,\theta)$ over $[0,\epsilon]$ and allows to state that $r_2(\theta)$ is differentiable over $[0,1]$ and $r_2'(\theta)$ has the expression given by (11), which ensures the increasingness of $r_2(\theta)$. The equalities $r_2(0) = 0$ and $r_2(1) = \epsilon$ are straightforward.

Proof of Proposition 3

Let $\theta \in [0,1]$. A two-part tariff (r,F) is accepted by all firms if and only if:

1. $\Psi_n(r,\theta) \geq 0 = \Psi_n(r_2(\theta),\theta)$ which is equivalent to $r \leq r_2(\theta)$ because $\Psi_n(r,\theta)$ is decreasing in r.
2. $0 \leq F \leq \Psi_n(r,\theta)$.

Proof of Lemma 4

$$\begin{aligned}
\bar{r}_n(\theta) &= \underset{0 \leq r \leq \epsilon}{\arg\max} P(r,\theta) = rq^e(n, c-\epsilon+r) + \Psi_n(r,\theta) \\
&= \underset{0 \leq r \leq \epsilon}{\arg\max} \left[\underbrace{rq^e(n,c-\epsilon+r) + F_n(r)}_{\text{objective function under } \theta=1} + \underbrace{(1-\theta)\left[\pi^i(n-1,c-\epsilon+r) - \pi^e(n,c-\epsilon)\right]}_{\text{increasing in } r \text{ and decreasing in } \theta} \right]
\end{aligned}$$

If $\bar{r}_n(\theta) \in]0,\epsilon[$ then the FOC $\frac{\partial P}{\partial r}(\bar{r}_n(\theta),\theta) = 0$ holds and differentiating it with respect to θ, we get that:

$$\begin{aligned}
\frac{d}{d\theta}\bar{r}_n(\theta) &= \frac{-\frac{\partial P}{\partial\theta\partial r}(\bar{r}_n(\theta),\theta)}{\frac{\partial^2 P}{\partial r^2}(\bar{r}_n(\theta),\theta)} \\
&= \frac{\frac{\partial}{\partial r}\pi^i(n-1,c-\epsilon+r)\,|_{r=\bar{r}_n(\theta)}}{\frac{\partial^2 P}{\partial r^2}(\bar{r}_n(\theta),\theta)}
\end{aligned}$$

Assumption A2 entails that the numerator $\frac{\partial}{\partial r}\pi^i(n-1,c-\epsilon+r)\mid_{r=\bar{r}_n(\theta)}$is non-negative. Combining this with the denominator being negative (since P is concave in r), we obtain that $\frac{d}{d\theta}\bar{r}_n(\theta) \leq 0$.

To rigorously conclude that $\bar{r}_n(\theta)$ is (weakly) decreasing in θ,it remains to show that if it happens that $\bar{r}_n(\theta)=0$ for some θ then $\bar{r}_n(\theta')=0$ for any $\theta' \geq \theta$. Assume that $\bar{r}_n(\theta)=0$ for some θ. Then, given the concavity of P in r, it must hold that $r \rightarrow P(r,\theta)$ is decreasing over $[0,\epsilon]$. Considering $\theta' \geq \theta$, we have: $P(r,\theta') = P(r,\theta) + (\theta-\theta')\pi^i(n-1,c-\epsilon+r)$. Using A2, we can then state that $(\theta-\theta')\pi^i(n-1,c-\epsilon+r)$ is (weakly) decreasing which yields that $P(r,\theta')$ is (weakly) decreasing and results in $\bar{r}_n(\theta')=0$.

We can now state that $\bar{r}_n(\theta)$ is (weakly) decreasing in θ.

Proof of Proposition 4

Since $rq^e(n,c-\epsilon+r)+\Psi_n(r,\theta)$ is strictly concave in r then $r_n^*(\theta) = \min(r(\theta),\bar{r}(\theta))$. Moreover we have already showed that $r(\theta)$ is strictly increasing over $[0,1]$ and $r(0)=0$, $r(1)=\epsilon$ while $\bar{r}(\theta)$ is (weakly) decreasing in θ and $\bar{r}(1) \leq \epsilon$. This allows us to state that there exists $\bar{\theta} \in]0,1]$ such that:

$$r_n^*(\theta) = \begin{cases} r(\theta) & if \quad \theta \leq \bar{\theta} \\ \bar{r}_n(\theta) & if \quad \theta > \bar{\theta} \end{cases}$$

This yields the result because:

$$F_n^*(\theta) = \Psi_n(r_n^*(\theta),\theta) = \begin{cases} 0 & if \quad \theta \leq \bar{\theta} \\ \Psi_n(\bar{r}_n(\theta)) & if \quad \theta > \bar{\theta} \end{cases}$$

Proof of the existence and unicity of $s(\theta)$

Consider $\theta \leq \tilde{\theta}$. The existence and unicity of $s(\theta)$ is derived from the following three points: i/ the function $g: r \longrightarrow nrq^e(n,c-\epsilon+r)$ is continuous and strictly increasing over $[0,\tilde{r}]$ because $r \rightarrow P(r,\theta) = n[rq^e(n,c-\epsilon+r)+\Psi_n(r,\theta)]$ is strictly increasing over $[0,r(\theta)]$ for any $\theta \leq \bar{\theta}$ and $\Psi_n(r,\theta)$ is decreasing in r, ii/ $g(0)=0 \leq \theta\tilde{P}(1)$, iii/ $g(\tilde{r}) \geq \theta\tilde{P}(1)$.

Proof of Proposition 5

The patent holder will prefer the licence $(r_n^*(\theta),F_n^*(\theta)) = (r(\theta),0)$ if and only if

$$nr(\theta)q^e(n,c-\epsilon+r(\theta)) \geq \theta\tilde{P}(1) = ns(\theta)q^e(n,c-\epsilon+s(\theta))$$

Note first that for any $\theta \leq \tilde{\theta}$, it holds that $r(\theta) \leq \tilde{r}$ because $r \rightarrow rq^e(n,c-\epsilon+r)$ is also strictly increasing over $[0,r(\theta)]$. Given that we have also $s(\theta) \leq \tilde{r}$ and $r \rightarrow rq^e(n,c-\epsilon+r)$ is strictly increasing over $[0,\tilde{r}]$, we can state that the inequality $nr(\theta)q^e(n,c-\epsilon+r(\theta)) \geq ns(\theta)q^e(n,c-\epsilon+s(\theta))$ holds if and only if $r(\theta) \geq s(\theta)$, which means that $(r^*(\theta),F^*(\theta)) =$

$(r_n^*(\theta), F_n^*(\theta)) = (r(\theta), 0)$ if and only if $r(\theta) \geq s(\theta)$. Otherwise, the patent holder prefers the license $\left(r_{-n}^*, F_{-n}^*\right)$ even though it triggers litigation.

Proof of Proposition 6

Since we tackle the case of sufficiently weak patents we can derive a comparison of $r^*(\theta) = r(\theta) = r_1(\theta)$ to $r^e(\theta)$ for θ small enough from the comparison of $r_1'(0)$ to ϵ. Indeed, if $r_1'(0) > \epsilon$ (resp. $r_2'(0) < \epsilon$) then for θ sufficiently small, but different from 0, we will have $r_1(\theta) > \theta\epsilon$ (resp. $r_1(\theta) < \theta\epsilon$).

Since $r_1(0) = 0$, we have:

$$r_1'(0) = \frac{-\pi^e(n, c-\epsilon)}{\pi_2^e(n, c-\epsilon)}$$

Therefore,

$$r_1'(0) > \epsilon \Longleftrightarrow \frac{-\pi^e(n, c-\epsilon)}{\pi_2^e(n, c-\epsilon)} > 1$$

Denoting $\lambda(\epsilon) = \pi^e(n, c-\epsilon)$, we obtain:

$$r_1'(0) > \epsilon \Longleftrightarrow \frac{\lambda(\epsilon)}{\epsilon\lambda'(\epsilon)} > 1 \Longleftrightarrow \eta(\epsilon) < 1$$

Proof of Proposition 7

We derive from proposition 5 that $P^*(\theta) = \theta P_{-n}^*(1)$ if $r(\theta) \leq s(\theta)$ and $P^*(\theta) > \theta P_{-n}^*(1)$ if $r(\theta) > s(\theta)$. This directly yields the result under case 1, i.e. $P^*(1) = P_{-n}^*(1)$. If the latter equality does not hold, which means $P^*(1) > P_{-n}^*(1)$ (as we always have $P^*(1) \geq P_{-n}^*(1)$) then: i/ $P^*(\theta) = \theta P_{-n}^*(1) < \theta P^*(1)$ if $r(\theta) < s(\theta)$, ii/ $P^*(\theta) = nr(\theta) q^e(n, c-\epsilon+r(\theta)) < nv(\theta) q^e(n, c-\epsilon+v(\theta)) = \theta P^*(1)$ if $s(\theta) \leq r(\theta) < v(\theta)$, iii/$P^*(\theta) = \theta P^*(1)$ if $r(\theta) = v(\theta)$, iv/ $P^*(\theta) = nr(\theta) q^e(n, c-\epsilon+r(\theta)) > nv(\theta) q^e(n, c-\epsilon+v(\theta)) = \theta P^*(1)$ if $r(\theta) > v(\theta)$.

Proof of Proposition 8

We have: $\pi^e(n, c) \geq \pi^i(n-1, c-\epsilon+r^c(\theta))$ because $r^c(\theta) \leq \epsilon$. Since $\pi^e(n, c-\epsilon+r^c(\theta)) = \theta\pi^e(n, c) + (1-\theta)\pi^e(n, c-\epsilon)$ we obtain that $\pi^e(n, c-\epsilon+r^c(\theta)) \geq \theta\pi^i(n-1, c-\epsilon+r^c(\theta)) + (1-\theta)\pi^e(n, c-\epsilon)$. The latter inequality implies that a royalty rate $r = r^c(\theta)$ will be non cooperatively accepted by all firms if proposed by the patent holder. Therefore $r^c(\theta) \leq r(\theta)$. Differentiating the equation $\pi^e(n, c-\epsilon+r^c(\theta)) = \theta\pi^e(n, c) + (1-\theta)\pi^e(n, c-\epsilon)$ with respect to θ, we get $\frac{dr^c(\theta)}{d\theta} = \frac{\pi^e(n,c,c)-\pi^e(n,c-\epsilon)}{\pi_2^e(n,c-\epsilon+r^c(\theta))}$. Both the numerator and the denominator are negative which implies that $r^c(\theta)$ is increasing.

Since $\pi^e(n, c) - \pi^e(n, c-\epsilon) < 0$ (due to A3), the derivative $\frac{dr^c(\theta)}{d\theta}$ is increasing in θ over $[0, 1]$ (i.e. $r^c(\theta)$ is convex) if and only if $\pi_2^e(n, c-\epsilon+r^c(\theta))$ is increasing in θ over $[0, 1]$. Since $r^c(\theta)$ is continuous and strictly increasing from $r^c(0) = 0$ to $r^c(1) = \epsilon$, the latter condition is equivalent to $\pi_2^e(n, x)$ is increasing in x over $[c-\epsilon, c]$, which means that $x \to \pi^e(n, x)$ is convex over $[c-\epsilon, c]$. In this case, $r^c(\theta) \geq \theta r^c(1) + (1-\theta) r^c(0) = \theta\epsilon$.

9.2 Appendix B: Comparison of $r(\theta)$ with the threshold $s(\theta)$ for "weak" patents

This comparison can be easily made for θ sufficiently small since we just need to compare $r'(0)$ and $s'(0)$. Given that $s(0)=0$, differentiating the equation

$$ns(\theta)\, q^{e}(n, c-\epsilon+s(\theta))=\theta P_{-n}^{*}(1)$$

with respect to θ at the point $\theta=0$, we get:

$$ns'(0)\, q^{e}(n, c-\epsilon)=P_{-n}^{*}(1)$$

which yields:

$$s'(0)=\frac{P_{-n}^{*}(1)}{nq^{e}(n, c-\epsilon)}$$

Differentiating the equation defining $r(\theta)$ and using the fact that $r(0)=0$, we get:

$$r'(0)=\begin{cases} r_1'(0)=\frac{\pi^{e}(n,c-\epsilon)}{\left|\pi_2^{e}(n,c-\epsilon)\right|} & \text{if} \quad \pi^{i}(n-1, c-\epsilon)=0 \\ r_2'(0)=\frac{\pi^{e}(n,c-\epsilon)-\pi^{i}(n-1,c-\epsilon)}{\left|\pi_2^{e}(n,c-\epsilon)\right|} & \text{if} \quad \pi^{i}(n-1, c-\epsilon)>0 \end{cases}$$

Note that in both cases, $r'(0)$ can be rewritten as:

$$r'(0)=\frac{\pi^{e}(n, c-\epsilon)-\pi^{i}(n-1, c-\epsilon)}{\left|\pi_2^{e}(n, c-\epsilon)\right|}=\frac{F_n(r=0)}{\left|\pi_2^{e}(n, c-\epsilon)\right|}$$

We get the following result:

Proposition 9 *If* $\frac{nF_n(r=0)}{P_{-n}^{*}(1)}>\frac{\left|\pi_2^{e}(n,c-\epsilon)\right|}{q^{e}(n,c-\epsilon)}$ *then for sufficiently weak patents, the patent holder's optimal license offer is* $(r(\theta), 0)$ *and litigation over the patent validity is deterred at equilibrium.*

If $\frac{nF_n(r=0)}{P_{-n}^{*}(1)}<\frac{\left|\pi_2^{e}(n,c-\epsilon)\right|}{q^{e}(n,c-\epsilon)}$ *then for sufficiently "weak" patents, the patent holder's optimal license offer is* $\left(r_{-n}^{*}, F_{-n}^{*}\right)$ *and litigation over the patent validity takes place at equilibrium.*

The interpretation of the inequality in proposition 9 is not evident. However, note that the LHS is a ratio of licensing revenues: the numerator $nF_n(r=0)$ is the highest licensing revenues that the patent holder can get from licensing an iron-clad patent to all firms through a pure fixed fee scheme $(r=0)$, while the denominator $P_{-n}^{*}(1)$ is the highest licensing revenues generated by a two-part tariff license offer inducing less than n licensees. The RHS of the inequality can be rewritten as the product of two terms $\eta(\epsilon)\left[\frac{p^{e}(n,c-\varepsilon)-(c-\varepsilon)}{\varepsilon}\right]$ where $\frac{p^{e}(n,c-\varepsilon)-(c-\varepsilon)}{\varepsilon}$ is the margin per unit of cost reduction the firms get if all of them use the new technology royalty-free and $\eta(\epsilon)=\frac{\varepsilon\left|\pi_2^{e}(n,c-\epsilon)\right|}{\pi^{e}(n,c-\epsilon)}$ is the elasticity of a firm's profit with respect to cost reduction when all firms benefit from this reduction. Therefore, the inequality in the proposition includes all the ingredients that play a role in licensing "weak" patents: the revenues generated by various licensing schemes on the one hand and the determinants of the downstream market competition on the other hand. One of the insights we can get from the latter proposition is that if the fixed fee mechanism is quite effective in extracting relatively high licensing revenues (as suggested by the literature on the comparison of licensing mechanisms in an iron-clad patent setting, see for instance Kamien 1992) then litigation deterrence at equilibrium is quite likely to hold. Furthermore, we show in proposition 6 that the elasticity $\eta(\epsilon)$ plays a crucial role in determining whether "weak" patents are overcompensated.