

“Essential” Patents, FRAND Royalties and Technological Standards*

Mathias Dewatripont and Patrick Legros[†]

March 17, 2008

Abstract

In this paper we abandon the usual assumption that patents bring known benefits to the industry or that their benefits are known to all parties. When royalty payments are increasing in one’s patent portfolio, private information about the quality of patents leads to a variety of distortions, in particular the incentives of firms to “pad” by contributing weak patents. Three main results that emerge from the analysis are that: (i) the threat of court disputes reduces incentives to pad but at the cost of lower production of strong patents; (ii) mitigating this undesirable side-effect calls for a simultaneous increase in the cost of padding, that is, a better filtering of patent applications; (iii) upstream firms have more incentives to pad than vertically-integrated firms which internalize the fact that patent proliferation raises the share of profits going to the upstream segment of the industry but at the expense of its downstream segment. This seems consistent with recent evidence concerning padding.

1 Introduction

While firms try to differentiate themselves from competitors, they also benefit from having standards established: this facilitates in particular the access of consumers to other providers’ consumers, allows economies of scale in the production of various goods (e.g. chipsets, other electronic parts, etc). At the same time, once a standard is established in the industry, it becomes very costly for a single firm to produce a non-standard good. For this reason, it is also well-known that firms are vulnerable to ex-post opportunism by innovators who detain the IP rights to some of the technologies that are essential to the use of the standard. This problem is particularly important when the creation of a standard requires the use of many different technologies, or innovations (in the case of mobile telephony for instance, a handset can require

*We thank Mathias Roende, Pekka Saaskilahti and Jean Tirole for useful comments on earlier drafts. We are also grateful to the Nokia corporation for financial support and for useful discussions. The views expressed in this paper are however solely our own.

[†]ECARES, Université Libre de Bruxelles and CEPR.

the use of more than a thousand technologies protected by patents). This is what has led to the FRAND concept (“fair, reasonable and non-discriminatory” royalties).

There is by now a significant literature on both the role of standards and of patents and the licensing thereof. As far as licensing agreements are concerned, the literature has in particular considered the role of fixed versus two-part tariffs; the licensing to multiple producers and the benefits from licensing term discrimination among producers; the use of licensing as a way to foreclose the downstream market; and the use of royalties negotiated ex-post as a way to “hold-up” producers.

However the theoretical literature seems relatively silent on the issue of licensing when different patent family holders must *cooperate* for the creation of a standard.¹ The literature on patent pools is a noted exception. It analyzes the efficiency of such patent pools, in particular with respect to antitrust concerns, as well as incentives to join patent pools or to price patents outside the pool.² There is less of a theoretical literature on SSO’s that are not organized as patent pools.³ While patent pools presume that firms agree on a well defined set of patents to share (and verify carefully the essentiality of the patents in the pool) many standards are dynamic in nature and firms define along the way which technologies are needed and seek potential patent holders for these technologies. The process of verification of essentiality of the patents is then more difficult than in patent pools since the technology against which essentiality is verified is evolving.

Moreover, there is one aspect that is missing in the literature (and this is true for both SSO’s and patent pools): *the process by which firms decide which patents to contribute to the standard*. Our focus in this paper is not on the choice of the standard per se, but on which patents end up being considered essential to this technology, with the implications of this process on investment in R&D (which may or may not lead to essential inventions).

There is indeed a fair amount of uncertainty as to which patents are “essential” for a given standard. This uncertainty is particularly acute at two levels:

1. There is a lot of uncertainty as to which patent applications will in fact be granted, due in

¹In the following we will use “patent” and “patent family” interchangeably and assume that patent families have full geographical coverage.

²See for example Shapiro (2001), Aoki and Nagaoka (2004), Lerner and Tirole (2004) and Layne-Farrar and Lerner (2006).

³See however the empirical analysis of Chiao et al. (2006).

particular to randomness of the work of patent offices: see the book by Jaffe and Lerner (2004) on the shortcomings of the US patent system for example. See also the work by Farrell and Shapiro (2005), Shapiro (2006) and Lemley and Shapiro (2005, 2006) on ‘weak patents’.

2. There seems to be a lot of ‘strategic claims’ about essentiality. Since each participant to the standard should expect to be compensated “in relation” to what it brings to the standard, there is an incentive to overemphasize one’s contribution. This is something that should be taken quite seriously for the mobile telephony industry. For instance, in a recent paper Goodman and Myers (2005) show, on the basis of evaluations by engineers, that up to 80% of the patents that firms claimed to be “essential” for a mobile telephone standard were not in fact essential (from a purely technical point of view, that is ignoring the possibility for these patents to be infringing other patents). There is therefore a fair amount of behavior which we can call “padding”.

This paper builds a model where firms invest in R&D and end up producing a variety of inventions, some of which may become, during the creation of a standard, essential for this standard, where essentiality means that the invention once patented is both legal (i.e., does not infringe on existing patents) and is also technically crucial for the standard. The firm may also decide to patent some of the other inventions, and bring these ‘weak’ patents to the standard. The sharing of profits however depends on the *total* number of patents submitted, unless these are knocked down in court. To model such profit sharing, we resort to the well-known Shapley value. Beyond its simplicity, we argue that the notion of symmetry embedded in this concept is naturally compatible with the principles behind FRAND. Moreover, we introduce a parameterization of the shares of profit going to the upstream and downstream segments of the industry (i.e., the patent holders and the makers of the final product), making the formulation pretty flexible.

Equipped with this bargaining solution, we analyze the impact of court disputes when an upstream firm provides patents. We derive a ‘limit padding’ condition, that is, a maximum level of patents submitted such that downstream firms, when correctly anticipating the patent mix of the upstream firm, prefer not to sue. When courts are not too expensive, that is, when the limit padding condition binds, a key result is that a reduction of the cost of going to court reduces

padding (the number of weak patents submitted) but also reduces the number of strong patents produced. This argues for an alternative, improved certification method to reduce padding. We show that an appropriate combination of lower court fees and better filtering of patents through higher ‘padding costs’ can simultaneously raise the number of strong patents while reducing (by the same amount) weak patenting.

We then turn to a situation where an upstream firm and a vertically-integrated firm are involved in patenting. To simplify we assume that patents cannot be disputed in court. Here, a key result is that the two firms, if equipped with the same patent production technology, will produce the same number of strong patents, but that the upstream firm will pad more than the vertically-integrated firm. The overall incentive to submit patents is indeed higher for the upstream firm because more patents mean more money for the upstream industry segment but less money for the downstream industry segment. Firms which are also active in the downstream market have therefore a lower incentive to submit patents. On the other hand, the choice of strong versus weak patents, for a given total number of patents submitted, will be driven by the patent-production technologies.

2 Model

2.1 Firms and Markets

There are $n + 1$ firms, denoted $0, 1, 2, \dots, n$. Firm 0 is specialized in producing patents and is not present in the downstream market; think of this firm as being “upstream”. Firms $1, 2, \dots, n$ are present in the downstream markets and have market shares $\alpha_i, i = 1, \dots, n$ decreasing in i .

There are M downstream markets. There is only one firm present on a given market but a firm can be present in more than one market: a firm’s market share is the ratio of the markets in which it is present to M . The profit on a market is π from the new product and is 0 otherwise.

We want to distinguish between “strong” and “weak” patents. Strong patents are patents that are truly essential to the standard and that will be found essential in court with probability one. Weak patents are patents that have successfully gone through the patent office but that are not essential to the standard: if there is a dispute, the court will find these patents to be

inessential with probability one. In a first best world, patent holders should present only their strong patents to the standard; however when there is no certification, patent holders may have incentives to engage in “padding” by presenting also weak patents.

In the “real world”, firms typically do not “aim” to produce weak patents. However, their R&D expenditures turn out to generate inventions that may or may not be essential for the standard which ‘emerges in the end’ (we leave in reduced form the question of which standard does emerge, a process which takes place largely after R&D expenditures have been incurred). It is then, in a second step, that firms may “disguise” some inessential inventions and successfully pretend they are essential (by the way, it is quite possible that these inventions might in fact have proved essential in case another standard would have emerged). Specifically, assume that the R&D cost of obtaining and patenting S strong patents is $\varphi(S)$, where $\varphi(S)$ is a strictly increasing convex function of S , while the cost of patenting a weak patent is c .⁴

Since strong patents embody essential new research, their effect on the profit of the industry should be greater than that of weak patents. To simplify, let us assume that weak patents have no effect on profits while strong patents have. Because of market uncertainty, if S strong patents are contributed to the standard, the realization of market profit is a random variable with distribution parametrized by S : $\pi \sim F(\pi, S)$ and we assume to simplify that the support of F is independent of S and that the mean profit

$$\pi(S) = \int \pi dF(\pi, S)$$

is increasing and concave in S .

Note that the technological standard that emerges is treated as a *reduced form* here: we simply say that firms invest in R&D, knowing that it takes $\varphi(S)$ to produce an expected number S of strong patents that can be included in the standard. There can of course be a lot of uncertainty concerning the ex-post value of S .

Concerning market structure, we will consider two cases. In the next section we consider the leading case when firms are specialized: firm 0 is the only producer of patents (we call this firm “upstream”) and the other firms are producers of final goods on the downstream market only.

⁴We implicitly assume there are enough inessential inventions around so that the marginal cost of producing weak patents is only the make-up cost.

Firm 0 could be also construed as a syndicate of patent holders. Our goal will be to analyze whether the upstream firm wants to pad and how changes in legal costs or the cost of creating weak patents affect the level of padding and welfare.

In the reality of markets however, producers of patents can be vertically integrated, that is be present both as producers of patents and of final goods. We consider therefore in section 4 the case where firms 0 and 1 can produce patents; firm 1 will be called “vertically integrated.” Firms 0 and 1 can both engage in padding and our objective here is to understand if there are different incentives for padding by an upstream or a vertically integrated firm. As we will see, everything else being equal in terms of patent production or stock of “strong” patents, upstream firms pad significantly more than vertically integrated firms in equilibrium.

Before turning to the analysis of these two cases, we discuss the royalty rates that participants to the standard will choose for a given number of patents that are deemed essential.

2.2 Fair Payoffs

We postulate that, when there are M markets and P patent families, each patent earns its owner a profit of $\phi_p(P, M)$ while each individual market earns its sole supplier a profit of $\phi_m(P, M)$. We assume that there is a unique firm on a market and for the purpose of modelling bargaining, we consider each firm as an integrator of the technologies available in its supplier network (see for instance Kranton and Minehart 2000). There are K suppliers for each firm, and the “firm” can achieve the profit π only if all the managers of its suppliers have the patents needed for their technology. The profit levels $\phi_p(P, M)$ and $\phi_m(P, M)$ are defined as:

$$\phi_p(P, M) = \frac{M}{P+K}\pi \quad \text{and} \quad \phi_m(P, M) = \frac{K}{P+K}\pi. \quad (1)$$

The appendix provides cooperative foundations for these expressions, based on the well-known Shapley value. While the Shapley value is a cooperative game theoretical concept, we view it as a convenient shortcut for modeling the outcome of a potentially complex multilateral noncooperative bargaining (e.g., Gul 1989). The reader should interpret the payoffs in (1) as the anticipation that the players have about the outcome of future negotiations rather than as an explicit pricing rule. Alternatively, these expressions can also be taken as a reduced form. In

this perspective, beyond the convenient linear formulation, note that they imply that:

1. The ‘downstream segment’ is assumed to receive a total amount $MK\pi/(P + K)$ of profits while the ‘upstream segment’ receives a total amount $MP\pi/(P + K)$ of profits (total profits being $M\pi$). The ratio P/K is therefore a measure of the relative bargaining power between the upstream and downstream industry segments. While P represents the number of patents involved in the standard, K is a parameter that could be ‘calibrated’ to replicate the upstream and downstream profit shares in a particular market.
2. Every patent family holder receives a royalty of $P/(P + K)$ per unit of profit, and independently of the level of profit, and each downstream market contributes to patent family holder revenues to the same extent. This symmetry can in fact be seen as being in accordance with FRAND.
3. Beyond this, note that a rise in the number of patents (e.g. because of padding) will reduce the profit of both preexisting patent owners and downstream suppliers.

As we have already noted, padding is limited by the cost of generating patents, the competition from other patent holders in the standard and by the incentives of other parties to dispute the essentiality of patents in the standard. We turn to this now, and we analyze two situations in turn.

3 Padding in the Shadow of Disputes

A necessary condition for padding is that the set of strong patents is privately known to its owner. Since “fair” royalties are computed on the basis of the set of patents that are deemed essential to the standard, a lower production of strong patents can be compensated by a larger production of weak patents without detection possibility by the other participants. This is no longer true in the case of certification or court disputes if the patent holder bears the cost of disputes on its weak patents.

To highlight the role of disputes, we will focus here on the leading case where only firm 0 can produce patents and only firm 1 can dispute these patents in court. Hence, we assume

implicitly that the other firms $2, \dots, n$ as “small” or as facing large costs of going to court. See also footnote 9.⁵

Because there is no ambiguity, we denote by S the number of strong patents of firm 0, by P the total number of patents contributed to the standard, and by d the proportion of patents that firm 1 decides to dispute (or the patents over which firm 1 refuses to pay the royalty).

The cost of going to court is f per patent disputed and we assume that this cost is borne by the party who loses the dispute in court.⁶ Hence if firm 1 disputes the essentiality of a patent and the court agrees, it is firm 0 which pays f , otherwise it is firm 1. As explained before, we assume for simplicity that only strong patents can be found essential in court.

It is convenient to consider a linear expected profit function: $\pi(S) = S\pi$ and a quadratic R&D cost $\varphi(S) = \mu S^2/2$. Note that the cost of one strong patent is $\mu/2$ while the industry profit per component when there is only one patent is $M\pi/K$. We assume throughout that the industry is *high profit* :

$$\frac{\mu}{2} < \frac{M\pi}{K}. \quad (2)$$

The marginal cost of weak patents is constant and equal to c .

The timing is the following:

- Firm 0 chooses S and W .
- Firm 1 observes the total number of patents $P = S + W$ and decides on the proportion d of patents to dispute.
- If after a dispute there are P' patents that have not been found inessential, royalty payments are decided on the basis of condition (1) with $P = P'$ patents.

An equilibrium is a pair (S, W) for firm 0 and a binary decision $d \in \{0, 1\}$ by firm 1 to dispute or not the patents of firm 0.

We proceed as follows:

⁵Alternatively, we could consider out-of-court settlements. In this case, discoveries that some patents are weak do not become known to the other firms: firm 0 therefore bears a lower cost in the case of dispute since it can still ask the other firms to pay the high royalty rate corresponding to the total number of patents submitted. This suggests that firm 0 will pad more. However, if in the negotiated settlement firm 1 is able to extract some of the royalty gains of firm 0 from non-disclosure, this should induce firm 1 to dispute more often. The net effect will depend on the way negotiation is modelled and on the distribution of bargaining power between firm 0 and firm 1.

⁶This assumption simplifies the algebra without affecting the essence of the results.

- We first characterize the optimal choice of strong patents for a given total number of patents P assuming that firm 1 will not dispute the patents. We show that there is padding whenever the total number of patents is greater than a cutoff level P^{eq} ;
- As P increases, the number of strong patents and of weak patents increases and there exists a cutoff value $P^{lim} > P^{eq}$ at which firm 1 is indifferent between not disputing and disputing given the optimal choice of firm 0.
- A revealed preference argument shows that there cannot be an equilibrium where firm 1 disputes patents with probability one.
- We then derive the optimum choice of firm 0, that is the number of patents it will effectively produce for the standard and show that there is “limit padding” in the sense that firm 0 produces the number of patents that make firm 1 indifferent between disputing and not the patents.⁷

Choice of Strong Patents when Firm 1 does not dispute

Assuming that firm 1 does not dispute any patent when there are P patents put forward by firm 0, the optimal choice of strong patents by firm 0 solves

$$\max_{S \leq P} \frac{P}{P+K} M\pi S - \mu \frac{S^2}{2} - c(P-S).$$

Ignoring the constraint, the first order condition is

$$\frac{P}{P+K} M\pi - \mu S + c = 0.$$

and the unconstrained maximum is achieved at

$$\sigma(P) = \frac{1}{\mu} \left\{ \frac{P}{P+K} M\pi + c \right\}. \quad (3)$$

⁷Moreover, firm 1 does not randomize and uses the pure strategy of not disputing. This result is due to our assumption that disputes happen before the profit is known. If disputes can arise after firm 1 gets information about the market profit, there will be a dispute for high levels of profit and no dispute for low levels of profit: this is apparent by inspecting (5) below and interpreting π as the realized profit. We chose our timing because it is empirically reasonable (market profits are realized well after the royalty agreements are made) and because it captures in a simple way the role of court fees on padding behavior.

The constrained maximum is $S(P) = \min\{\sigma(P), P\}$. A first observation is that there is no padding - that is $S(P) = P$ - if and only if P is smaller than a cutoff value.

Proposition 1. ⁸ (i) *There exists $P^{eq} > 0$ such that the optimal number of strong patents at P assuming that firm 1 does not dispute is*

$$S(P) = \begin{cases} P & \text{if } P \leq P^{eq} \\ \sigma(P) & \text{if } P \geq P^{eq}. \end{cases}$$

(ii) $S(P)$ is increasing and concave in P ,

(iii) $P - S(P)$ is increasing in P , $(P - S(P))/P$ is increasing and concave in P .

In other words, without disputes, there is padding only for $P > P^{eq}$ and the level of padding is increasing in P , both in absolute number but also relative to the number of patents. Everything being equal, a larger number of patents indicates a higher proportion of weak patents.

Limit Padding

Consider now the behavior of firm 1. If at P , firm 1 has beliefs S , and disputes a proportion d of patents, while firm 0 actually chooses a number S' of strong patents, the payoffs for firms 0 and 1 are as follows

$$u_0(S', d; P) = \frac{S' + (1-d)(P-S')}{S' + (1-d)(P-S') + K} MS'\pi - fd(P-S') - \frac{\mu S'^2}{2} - c(P-S') \quad (4)$$

$$u_1(S, d; P) = \frac{\alpha_1 K}{S + (1-d)(P-S) + K} MS\pi - fdS.$$

Note that if firm 1 disputes the patents, it creates an industry wide externality since it becomes known which patents are truly essential. This externality translates into lower royalty payments for all firms.⁹

⁸All proofs are in the appendix.

⁹Hence, disputing the patents of firm 0 has the flavor of a public good. If more than one firm can dispute patents, there is a free rider problem since each firm will prefer another firm to dispute in order to benefit from the externality without having to bear the court costs. We should expect that in equilibrium, the level of patents for which a given firm is indifferent between disputing or not will increase: in other words, the process of ex-post certification by disputes will be even less efficient in preventing padding than under our assumption.

Since $u_1(S, d; P)$ is convex in d , the best response of firm 1 is either to dispute all patents ($d = 1$) or dispute no patent ($d = 0$). It is best for firm 1 not to dispute if and only if

$$\frac{\alpha_1 K}{P + K} M S \pi \geq \frac{\alpha_1 K}{S + K} M S \pi - f S$$

or,

$$\frac{\alpha_1 K}{S + K} M \pi - \frac{\alpha_1 K}{P + K} M \pi - f \leq 0 \quad (5)$$

In order to avoid disputes, firm 0 must choose a minimum number of strong patents. However because firm 1 observes only P and not S , the condition must be met at the optimal choice for firm 0.

Can we have an equilibrium in which firm 1 disputes all P patents with probability one? In such a case, firm 0 will still choose S in order to maximize the payoff given by (4). However, if $S^d(P)$ is the optimum, it must be the case that $P = S^d(P)$, that is, firm 0 should abstain from any padding: It would rationally expect that such padding would be successfully undone by the dispute and would only imply a cost of $f(P - S^d(P))$ for firm 0. But if, in equilibrium, there are no weak patents, it will be anticipated by firm 1, which will find it optimal not to dispute the patents, which is a contradiction. This means we have the following result:

Lemma 1. *In equilibrium, the padding constraint (5) is satisfied.*

When $P > P^{eq}$, by substituting $\sigma(P)$ for S in (5) we need,

$$\Delta(P) \leq f$$

where

$$\Delta(P) = \frac{\alpha_1 K}{\sigma(P) + K} M \pi - \frac{\alpha_1 K}{P + K} M \pi$$

has the following properties.

Lemma 2. *Let $\Delta(\infty) = \mu \frac{\alpha_1 K M \pi}{M \pi + c + \mu K}$.*

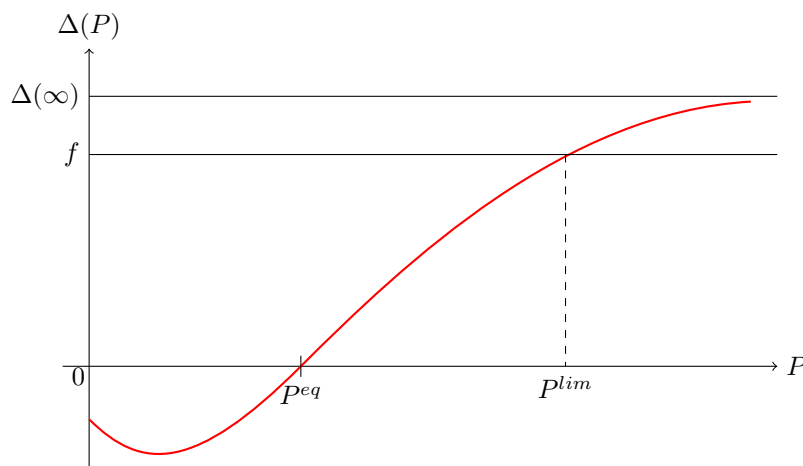
(i) *If $\Delta(\infty) \leq f$, then (5) is satisfied for all values of P .*

(ii) *If $\Delta(\infty) > f$, then there exists a unique value $P^{lim} > P^{eq}$ such that (5) binds.*

(iii) *$\Delta'(P) > 0$ for all $P > P^{eq}$.*

As we show in the proof of the lemma, the function $\Delta(P)$ has typically

¹⁰ the following graph. We have also illustrated the case (ii) of the lemma where the limit padding constraint binds.¹¹



The condition $\Delta(\infty) > f$ holds when f is small enough; moreover from (iii), the value of P^{lim} for which (5) binds is an increasing function of f . This is indeed intuitive: decreasing the cost of going to court will make firm 1 more aggressive in disputing patents when P is greater than P^{eq} .

We consider the two regimes of high and low court costs in turn. While the case of high court fees is of lesser interest here, it provides the benchmark to evaluate the role of disputes as a disciplining device for firm 0. Moreover, the technical analysis is broadly similar, so that the high-court-fee case will be helpful to derive the main result of this Section, namely Proposition 5.

Note that in both cases, despite the fact that firm 0 could decide to limit the number of patents and produce only strong patents, its desire to extract higher royalty payments will lead the upstream firm to increase the number of patents offered in the standard and in the process will produce weak patents.

¹⁰It may not be decreasing at zero, but is always increasing beyond P^{eq}

¹¹Obviously the economically relevant part of the graph is when $P \geq P^{eq}$.

3.1 Two Regimes

3.1.1 High Court Fees: $\Delta(\infty) \leq f$

In this case, there is no dispute for any value of P . To derive our result, it will be helpful to split the space of P 's, into $P \leq P^{eq}$ and $P > P^{eq}$.

When $P \leq P^{eq}$, there is no padding and firm 0 chooses P to maximize $u_{0-}(P) = \frac{P^2}{P+K}M\pi - \mu\frac{P^2}{2}$. The derivative of this function is :

$$u'_{0-}(P) = \frac{P^2 + 2PK}{(P+K)^2}M\pi - \mu P. \quad (6)$$

and the sign of the derivative is the same as the sign of $\frac{P+2K}{(P+K)^2}M\pi - \mu$. This function is decreasing in P and has value $\frac{2M\pi}{K} - \mu$ at $P = 0$, and is therefore positive by (2). As $P \rightarrow \infty$, the function has limit equal to $-\mu$. It follows that:¹²

Lemma 3. *The function $u_{0-}(P) = \frac{P^2}{P+K}M\pi - \mu P$ is single peaked.*

This optimal value of P can be smaller or greater than P^{eq} depending on the value of c . If c is “small” however, the optimum is greater than P^{eq} and therefore it is optimal for firm 0 to choose $P = P^{eq}$.

Corollary 1. *There exists $\hat{c} > 0$ such that for all $c < \hat{c}$, the maximum payoff to firm 0 over $P \leq P^{eq}$ is attained at P^{eq} .*

Consider now the case $P > P^{eq}$. There is padding and firm 0 chooses P to maximize

$$u_{0+}(P) = \frac{P}{P+K}M\pi\sigma(P) - \mu\frac{\sigma(P)^2}{2} - c(P - \sigma(P)).$$

By the envelop theorem,

$$u'_{0+}(P) = \frac{K}{(P+K)^2}M\pi\sigma(P) - c.$$

From proposition 1, the ratio $\sigma(P)/(P+K)^2$ is decreasing in P . Hence the sign of $u'_{0+}(P)$ is “decreasing ” and $u_{0+}(P)$ is single peaked. While there is a change of regime at P^{eq} , it

¹²The proof of this lemma and the following corollary are in the appendix.

is possible to show that the payoff is differentiable at P^{eq} ,¹³ and therefore that for $c < \hat{c}$, $u'_{0+}(P^{eq}) = u'_{0-}(P^{eq})$ is positive. It follows that there is an interior solution $u'_{0+}(P) = 0$. Because $P = P^{eq}$ is feasible, this interior solution is also a global maximum.

Proposition 2. *Suppose that court fees are high and that the cost of weak patents is small ($c < \hat{c}$). Firm 0 chooses optimally to produce weak patents: the number of patents P^{no} , $P^{no} > P^{eq}$ solves*

$$\frac{K}{(P + K)^2} M\pi\sigma(P) - c = 0.$$

Remark 1. When the cost of producing weak patents, c , is instead greater than the cutoff \hat{c} , firm 0 will not pad: the optimum when $P \leq P^{eq}$ is strictly lower than P^{eq} , implying $u'_0(P^{eq}) < 0$, while the optimum when $P \geq P^{eq}$ is at P^{eq} , implying a global maximum less than P^{eq} . This is rather intuitive: weak patents increase the revenue only through the level of royalties but not through the level of market profits: if the cost of weak patents is high enough, they become a poor substitute for strong patents in raising total profits since strong patents increase both the royalty rate and the market profit π .

3.1.2 Low Court Fees: $\Delta(\infty) > f$

The analysis in the regime $P \leq P^{eq}$ is the same as before and P^{eq} is the optimal choice when c is small.

In the regime $P \geq P^{eq}$, consider P^{no} defined in Proposition 2, the optimum when firm 0 does not face disputes, and P^{lim} the level of patents at which firm 1 is indifferent between disputing and not when firm 0 chooses $\sigma(P)$ strong patents. Since when there is no dispute the payoff to firm 0 is single peaked, it optimally chooses $P^d = \min(P^{lim}, P^{no})$.

If $P^d = P^{no}$, we are done.

If $P^d = P^{lim}$, firm 1 is indifferent between disputing and not. Could we have equilibria in which firm 1 disputes with probability $\gamma > 0$? No, because firm 0 strictly prefers $\gamma = 0$ to $\gamma > 0$, and would infinitesimally reduce its total number of patents to make it strictly optimal for firm 1 to stop disputing altogether. Consequently, the unique equilibrium behavior at P^{lim} is for firm

¹³To see this, use $P^{eq} = \frac{1}{\mu} \{ \frac{P^{eq}}{P^{eq}+K} M\pi + c \}$ and substitute $\mu = \frac{M\pi}{P^{eq}+K} + \frac{c}{P^{eq}}$ in u'_{0-} .

1 not to dispute the patents.

Proposition 3. *Suppose that court fees are low and that the cost of weak patents is low ($c < \hat{c}$). Firm 0 produces $\min(P^{lim}, P^{no})$ patents and there is no dispute in equilibrium.*

3.2 Comparative Statics and Policy Implications

Within our model, we can analyze the role played by the courts (f) and certification or other ways to make padding more difficult (c could be the cost of certifying patents, either via the patent system or via the standard). It is useful to summarize our previous findings and illustrate the roles of c and f in the different regimes.

A key result is Proposition 1, which establishes the covariation between strong and weak patents, or between the level of R&D and padding. This is an implication of FRAND and of the fact that the share of industry profits accruing to the upstream firm is increasing in the number of patents it brings to the standard. Indeed, when there are fewer weak patents, the share of profits accruing to the upstream firm decreases, the marginal revenue from strong patents is also lower and the upstream firm will invest less in R&D.

There are three relevant cutoff values for the total number of patents P , and we now make explicit their dependence on the parameters c and f .

$P^{eq}(c)$ is the cutoff value of P where firm 0 is indifferent between starting to pad or not. This value is increasing in c : intuitively, as the cost of padding increases, the upstream firm is more willing to invest in R&D rather than go for weak patents in order to increase its revenues.

$P^{no}(c)$ is the optimal number of patents that the upstream firm will bring to the standard, assuming that it does not face the possibility of dispute (for instance if f is large). As we know, this value is greater than $P^{eq}(c)$ – and there is padding – if and only if c is smaller than \hat{c} . As c increases, it is more costly to pad, and the upstream firm should pad less. However, it can be shown that is also the case that the number of strong patents decreases (note that this last result does not follow immediately from proposition 1 however, which assumes c to be constant).

Proposition 4. *Suppose that $c < \hat{c}$ and that the equilibrium number of patents is $P^{no}(c)$. Then, locally, as c increases, the number of strong patents and the number of weak patents decrease.*

Hence, when disputes are not effective, it is not possible to simultaneously decrease the level of padding without depressing the level of R&D: increasing c will limit the number of weak patents but at the cost of reducing the total industry profit, and the value of the standard itself.

Consider now the case where disputes can be effective. $P^{lim}(c, f)$ is the number of patents for which the downstream firm is indifferent between disputing and not the patents of the upstream firm. The consequences of a variation of an increase in c are more subtle than in the previous case, because there are two effects at play. First as c increases, there is the usual substitution effect: for a given number of patents, the upstream firm will have relatively more strong patents than weak patents. Second, there is the effect on limit padding: because of the substitution effect, the downstream firm disputes less aggressively, hence the limit padding constraints of the upstream firm is weakened and it will *increase* the number of patents it brings to the standard.

The substitution effect goes in the direction of a decrease in padding; however the other effect goes, via the covariation result, in the direction of an increase in padding. The net effect is a priori ambiguous. However, we prove below that the net effect is positive, thus overturning the results of the previous proposition.

As f decreases, the downstream firm finds it less costly to dispute patents and the limit padding constraint is strengthened, leading to a decrease in the number of weak patents that are submitted. However, by the covariation result, it is also the case that the number of strong patents (or R&D investment) decreases. Hence, like in the previous situation, the use of f only cannot limit padding without adverse effects on R&D incentives.

Proposition 5. *Suppose that $c < \hat{c}$ and that the equilibrium number of patents is $P^{lim}(c, f)$.*

(i) As f decreases locally, the total number of patents, the number of strong patents and the number of weak patents decrease.

(ii) As c increases locally, the total number of patents, the number of strong patents and the number of weak patents increase.

An increase in the cost of weak patents *does not limit* padding, but has the desirable effect to increase the quality of the standard. At the same time, Proposition 5 illustrates how using f only will have the undesirable consequence of reducing the investment in R&D, and the quality of the standard. In particular, this illustrates a limitation of courts for disciplining firms in standards:

while lower court fees will reduce padding, they also reduce the value of the standard since a lower number of strong patents is produced. This suggests that non-court based certification processes may be better able to correct for padding without destroying incentives to produce strong patents. One possibility would be for the participants to the standard to share the costs of an ex-ante certification process. There are indications that the industry is experimenting in this direction (see <http://www.3glicensing.com/> for certification by neutral third parties in 3G licensing pools).

Interestingly, combining a decrease in f with an increase in c can also achieve the goal of reducing padding while providing incentives to increase the quality of the standard. Because $P^{lim}(c, f)$ is increasing in c and in f , it is indeed possible to find a positive variation dc of c and a negative variation df of f such that the total variation of the number of patents is zero. Since f does not affect the substitution effect, the increase in c will then lead to an increase in the proportion of strong patents and a decrease in padding.

Proposition 6. *Suppose that $c < \hat{c}$ and that the equilibrium number of patents is $P^{lim}(c, f) < P^{no}(c)$. There exist $dc > 0$ and $df < 0$ such that the number of patents stays constant, but the number of weak patents decreases.*

The policy implications of the model are nontrivial because of the covariation between strong and weak patents. This covariation further suggests that it may be possible for the upstream firm to *over-invest* in R&D in order to increase its share of the industry profits. Indeed, while in standard moral hazard situations, there is a tendency for underinvestment because the agent does not get the full marginal return of its investment, in our model there is an additional effect at play, due to the dependence of the share on the contribution of the agent.

In our model, the first best investment solves $\max_S M\pi S - \varphi(S)$, yielding $S^* = M\pi/\mu$. From (3), the number of strong patents in the second best is bounded above by $\sigma(\infty) = (M\pi + c)/\mu$. Hence it is a priori possible for the second-best investment to be greater than the first-best investment, but only for P very large, which is not likely to happen if for instance the limit padding constraint holds. If there is no possibility of padding at all, then there is necessarily underinvestment.

However, this underinvestment result is largely due to the linearity of the industry profit

function. Consider instead a concave profit function $\pi(S)$, and assume to simplify that padding is not possible. The first best solves:

$$M\pi'(S^*) = \varphi'(S^*). \quad (7)$$

The second best is obtained at the maximum of $\frac{S}{S+K}M\pi(S) - \varphi(S)$, that is when:

$$\underbrace{\frac{S}{S+K}M\pi'(S)}_{\text{usual underinvestment effect}} + \underbrace{\frac{K}{(S+K)^2}M\pi(S)}_{\text{endogenous share effect}} = \varphi'(S). \quad (8)$$

There is overinvestment if, evaluated at S^* , the left-hand side of (8) is greater than $M\pi'(S^*)$.

This condition reduces to:

$$\pi'(S^*) < \frac{\pi(S^*)}{S^* + K}.$$

Clearly this condition fails if π is linear but can hold if π is sufficiently concave (for instance, if $\pi(S) = \log(S+a)$ where $a \geq K$). As far as we know this observation is new to the literature on moral hazard: in traditional settings, the issue is to provide incentives to exert effort by shares of the final output; in our setting, this is true but it is also the case that the share itself is a function of an observable signal contingent on the agent's effort.

4 Competitive Padding

The previous section has highlighted that padding is likely to be an equilibrium outcome of the patent application process: royalty payments go up and as long as the cost of producing weak patents is small enough, a patent producer will find it beneficial to pad. One specificity of the previous section is that it has treated all patent contributors symmetrically, by assuming that patent contributors are not active on the downstream market. Since in the reality of markets, patent contributors are also often producers of final goods, it is natural to understand whether the presence of a patent producer active on the downstream market will lead to a change in the equilibrium choice of padding by the upstream firm.

To simplify, we will ignore the possibility of disputes here and assume that only firms 0, 1

can produce patents.¹⁴ Since there is no possibility of dispute, the analysis will highlight the disciplinary role of competition for padding among heterogenous firms.

Let S_0, S_1 be the numbers of strong patents of firms 0, 1: only i knows S_i . In addition to these strong patents, these firms can present weak patents (or “pad”) and we denote by W_i these numbers. The sum of the strong and weak patents is denoted by $P_i = S_i + W_i$ for firm i and the total number of patents is denoted by $P = P_0 + P_1$.

Since there is no possibility of dispute or of certification, for a given realization of π , the payoff to patent holders is obtained by multiplying the Shapley value of a patent by the number of patents he contributes and the payoff to a downstream firm by multiplying the Shapley value of a market by the market share $\alpha_i M$ of this firm. Since firm 1 is vertically integrated, its expected payoff is the sum of these two values, where the expectation is taken with respect to $F(\pi, S)$. Hence,

$$\begin{aligned} u_0(S, W) &= \frac{S_0 + W_0}{S + W + K} M\pi(S) - cW_0 - \varphi(S_0) \\ u_1(S, W) &= \frac{S_1 + W_1 + \alpha_1 K}{S + W + K} M\pi(S) - cW_1 - \varphi(S_1) \\ u_i(s, w; M) &= \frac{\alpha_i K M \pi(S)}{S + W + K}, \quad i \geq 2 \end{aligned}$$

Let us now take derivatives for firm 0:

$$\frac{\partial u_0}{\partial W_0} = \frac{P_1 + K}{(S_0 + W_0 + P_1 + K)^2} M\pi(S_0 + S_1) - c \quad (9)$$

and:

$$\begin{aligned} \frac{\partial u_0}{\partial S_0} &= \frac{P_1 + K}{(S_0 + W_0 + P_1 + K)^2} M\pi(S_0 + S_1) \\ &\quad + \frac{S_0 + W_0}{S_0 + W_0 + P_1 + K} M\pi'(S_0 + S_1) - \varphi'(S_0). \end{aligned}$$

¹⁴The extension to the case where all downstream firms can contribute is straightforward. Proposition 7 generalizes simply: all firms produce the same number of strong patents but the number of weak patents is *decreasing* in the market share.

Setting both of these derivatives to zero, we have:

$$\frac{S_0 + W_0}{S_0 + W_0 + P_1 + K} M\pi'(S_0 + S_1) = \varphi'(S_0) - c.$$

Similarly, for firm 1, we have:

$$\frac{\partial u_1}{\partial W_1} = \frac{P_0 + K(1 - \alpha_1)}{(P_0 + S_1 + W_1 + K)^2} M\pi(S_0 + S_1) - c$$

and:

$$\begin{aligned} \frac{\partial u_1}{\partial S_1} &= \frac{P_0 + K(1 - \alpha_1)}{(P_0 + S_1 + W_1 + K)^2} M\pi(S_0 + S_1) \\ &\quad + \frac{S_1 + W_1 + \alpha_1 K}{P_0 + S_1 + W_1 + K} M\pi'(S_0 + S_1) - \varphi'(S_1). \end{aligned}$$

Setting these derivatives to zero, we have:

$$\frac{S_1 + W_1 + \alpha_1 K}{P_0 + S_1 + W_1 + K} M\pi'(S_0 + S_1) = \varphi'(S_1) - c.$$

The first-order conditions on weak patents for the two firms imply:

$$P_0 = P_1 + \alpha_1 K.$$

But having these convex costs of producing strong patents means that putting together the last condition for each of the two firms implies:

$$S_0 = S_1$$

Note that it would be socially optimal for firms not to produce weak patents and to produce strong patents in order to maximize $M\pi(S) - \varphi(S_0) - \varphi(S_1)$, yielding a first best optimum of

$$W^{FB} = 0, \quad S_0^{FB} = S_1^{FB} = S^{FB} \quad \text{s.t.} \quad M\pi'(2S^{FB}) = \varphi'(S^{FB}) \quad (10)$$

In equilibrium, we have instead:

$$\frac{P_0}{P_0 - \alpha_1 K + K} M \pi'(2S_0) + c = \varphi'(S_0). \quad (11)$$

Proposition 7. (i) *In equilibrium, firms 0 and 1 produce the same number of strong patents*
(ii) *Firm 0 produces $\alpha_1 K$ more weak patents than firm 1*

Firm 0 thus contributes more patents to the standard but these are weak ones, not strong ones, which come equally from both firms. The incentive to submit patents is higher for firm 0 because more patents mean more money for the upstream industry segment but less money for the downstream industry segment. Firms which are also active in the downstream market have therefore a lower incentive to submit weak patents. And as far as strong versus weak patents are concerned, since this is done for a given total number of patents submitted, two firms with identical technologies will make the same choices, namely they will equate the marginal cost of producing strong patents with c , the assumed constant marginal padding cost.

5 Conclusion

By abandoning the usual assumption that patents bring known benefits to the industry or that their benefits are known to all parties, we have been able to derive several results. In particular, we have shown that the threat of court disputes reduces incentives to pad but at the cost of lower production of strong patents. Second, we have shown that upstream firms have more incentives to pad than vertically-integrated firms, which internalize the fact that patent proliferation raise the share of profits going to the upstream segment of the industry but at the expense of its downstream segment.

These results seem consistent with empirical results obtained by Goodman and Myers (2005) for the case of mobile telephone standards. Indeed, they show that: (i) all major patent producers seem to exaggerate claims of essentiality, and (ii) the extent of exaggeration seems to be much more significant in the case of a firm like Qualcomm, which is ‘more upstream’ than its main rivals.¹⁵

¹⁵Geradin et al. (2006) classify in their analysis Qualcomm as an upstream firm and Nokia, Ericsson and Motorola as vertically integrated.

While this deserves some further investigation, this is evidence consistent with our analysis.

Our results on court disputes have potentially significant policy implications. They show that easier court access can have an undesirable side-effect in terms of strong patent production. This calls for combining easier court access with an increase in the cost of padding, that is, better filtering of patent applications. Looking for an ‘operational’ way of limiting padding while simultaneously encouraging innovation constitutes an interesting avenue for further research. While this model is a useful starting point in this respect, it would benefit from extensions. For example, it would be interesting to allow for multiple patent jurisdictions, with differences in the granting and enforcement of patent rights. The patent dispute process could also be generalized, to allow for the optimality of ‘partial disputes’, i.e., on a subset of the patent portfolio. And, very importantly, we have kept the choice of standard in ‘reduced form’, while it would be very interesting to explicitly link this choice to the outcome of the R&D investment process. This would probably require us to relax our assumption of a deterministic link between R&D expenses and the number of strong patents, something that allowed us to avoid adverse selection issues in the padding and dispute phases. These extensions are beyond the scope of this paper but could constitute interesting avenues for future research.

6 References

1. Aoki, R. and S. Nagaoka (2004), “The consortium standard and patent pools,” Hitotsubashi University Hi-Stat WP 32.
2. Chiao, B., J. Lerner and J. Tirole (2006), “The rules for standard-setting organizations: an empirical analysis,” mimeo.
3. Farrell, J. and C. Shapiro (2005), “How strong are weak patents?” University of California at Berkeley, Competition Policy Center WP 05-054.
4. Geradin, D., A. Layne-Farrar and A.J. Padilla (2006), “Royalty stacking in high-tech industries: separating myth from reality,” mimeo.
5. Goodman, D.G. and R.A. Myers (2005), “3G cellular standards and patents,” IEEE WirelessCom 2005, June 13.

6. Gul, F. (1989), "Noncooperative Foundations for the Shapley Value," *Econometrica*, 57(1): 81-95.
7. Hart, O. and J. Moore (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6):1119-1158.
8. Jaffe, A.B. and J. Lerner (2004), *Innovation and its discontents: How our broken patent system is endangering innovation and progress, and what to do about it*, Princeton University Press.
9. Kranton, R. E. and D. F. Minehart (2000), "Networks versus vertical integration," *The Rand Journal of Economics*, 31(3):570-601.
10. Layne-Farrar, A. and J. Lerner (2006), "To join or not to join: examining patent pool participation and rent sharing rules," mimeo.
11. Layne-Farrar, A., J. Padilla and R. Schmalensee (2006), "Pricing patents for licensing in standard setting organizations: making sense of FRAND commitments," mimeo.
12. Lemley, M. and C. Shapiro (2005), "Probabilistic patents," *Journal of Economic Perspectives*.
13. Lemley, M. and C. Shapiro (2006), "Patent hold-up and royalty stacking," mimeo.
14. Lerner, J. and J. Tirole (2004), "Efficient Patent Pools," *American Economic Review* 94: 691-711.
15. Myerson, R. (1977), "Graphs and Cooperation in Games," *Mathematics of Operations Research*, 2: 225-229.
16. Shapiro, C. (2001), "Navigating the patent thicket: cross licenses, patent pools, and standard-setting," in A. Jaffe, J. Lerner and S. Stern (eds), *Innovation policy and the economy, volume I*, MIT Press.
17. Shapiro (2006), "Injunctions, hold-up, and patent royalties," mimeo.
18. Swanson, D.G. and W.J. Baumol (2005), "Reasonable and nondiscriminatory (RAND) royalties, standard selection, and control of market power," *Antitrust Law Journal*.

7 Appendix

7.1 Fair payoffs: foundation

We show in this appendix that the “fair” payoffs expressed in (1) can be obtained using the Shapley value (see Myerson, 1977, for a general theoretical discussion, Hart and Moore 1990, for an application to incomplete contracting and Layne-Farrar et al., 2006, for an application to mobile-phone patents.) Specifically, assume that:

- There is a set \mathcal{P} of patents that are claimed to be essential to the production of a product. Since each patent is assumed to be controlled by one manager, they all have to have agreed to license their patent for production to go ahead.
- There are K managers in the supplier network in a given market, hence there are MK managers; all managers in one market are strict complements in the sense that their firm can produce the new product only if all managers have contracted with the patent holders.

By assuming extreme decentralization of both the patent decisions and the production decisions, we in fact parameterize the relative bargaining powers of the upstream and downstream segments of the industry.

The Shapley value is defined axiomatically and its “fairness” interpretation comes from its axiom of symmetry. The “nondiscriminatory” ingredient in FRAND can be interpreted as requesting that the royalty paid by firms is the same for each individual market, independently of market share, and that patent holders should receive the same per-patent royalty independently of their total portfolio. This is equivalent to assuming that each patent and each market are treated as a separate entity. Hence there are effectively $p + MK$ players: p patent holders and K managers per market. Because of symmetry, we know that the payoff to a manager on each market is the same, and that the payoffs to the patents are the same.

Remember that if for a given a set of players \mathcal{N} , the total payoff to a coalition S is described by a function $v(S)$, the Shapley value is defined as follows. Consider a random order on \mathcal{N} , let S_i be the set of all players preceding i in this random order. The marginal contribution of i is

$v(S_i \cup \{i\}) - v(S_i)$, and the Shapley value is given by

$$\phi_i = E [v(S_i \cup \{i\}) - v(S_i)]$$

where E is the expectation operator when all $|N|!$ orders over \mathcal{N} are assigned equal probability. It follows that if two players have the same marginal contributions if they occupy the same position in an order that their Shapley value are the same.

Consider the case of a unique market, hence when there are $P + K$ players. Since firms need all the patents in \mathcal{P} in order to use the standard, and since the profit of a firm is realized only if its K managers have acquired the patents, any coalition not containing all the patents or all the managers would not produce a profit.

Hence each patent and each manager has a positive marginal contribution of π if and only if it is “last” in the order on the $P + K$ players. It follows that each manager and each patent holder receives $\pi/(P + K)$. Since on a market the K managers belong to the same firm, the firm receives $K\pi/(P + K)$. We prove now that this value is the same independently of the number M of markets:

Lemma 4. *Consider P essential patents and M markets, the Shapley value for each patent and for a firm present on a market are respectively*

$$\phi_p(P, M) = \frac{M\pi}{P + K} \quad \text{and} \quad \phi_m(P, M) = \frac{K\pi}{P + K}. \quad (12)$$

Proof. We index patents by $i = 1, \dots, P$ and we index managers by $k(m, i), i = 1, \dots, K, m = 1, \dots, M$ where $k(m, i)$ is the i -th manager present on market m . By assumption all managers $k(m, i), i = 1, \dots, K$ belong to the same firm. Let \mathcal{K}_m be the set of managers.

By symmetry, we know that all patents have the same Shapley value ϕ_p and all managers in market m have the same Shapley value. Hence, if there are N players, it is enough to determine the Shapley value $\phi_p(\mathcal{N})$ common to all patent holders and the Shapley value $\phi_{j(m,k)}(\mathcal{N})$ of all the K managers in market m , where $m = 1, \dots, M$. The value of a firm present in a market is then $K\phi_{j(m,k)}(\mathcal{N})$

We use the following property of the Shapley value : balanced contribution (Myerson 1977) requires that what player i contributes to player j is equal to what player j contributes to player i . Or if \mathcal{N} is the set of players, that:

$$\phi_i(\mathcal{N}) - \phi_i(\mathcal{N} - \{j\}) = \phi_j(\mathcal{N}) - \phi_j(\mathcal{N} - \{i\})$$

In our case, $\mathcal{N} = \mathcal{P} \cup \mathcal{K}_{\mathcal{M}}$. Letting i be a patent and $j(m, k)$ be a manager, we know that $v(\mathcal{N} - \{i\}) = 0$ since no new product can be put to the market if one essential patent is missing and $v(\mathcal{N} - \{j(m, k)\}) = (M - 1)\pi$ since if one component is not present a product cannot be produced on market m .

In the game $\mathcal{N} - \{j(m, k)\}$ all players $j(m, k')$ have zero marginal contributions, hence their Shapley value is equal to zero: $\phi_m(\mathcal{N} - \{j(m, k)\}) = 0$. It follows that balanced contribution implies:

$$\phi_p(\mathcal{N}) - \phi_p(\mathcal{N} - \{j(m, k)\}) = \phi_{j(m, k)}(\mathcal{N}), \text{ for each } m, k \quad (13)$$

To show (12), we proceed by induction on the number of markets M . From the text, the result is true for $M = 1$. We suppose it is true for $M - 1$ and we show that the result is true for M .

Since for each coalition S , $v(S - \{j(m, k)\}) = v(S - \cup_{k=1}^K \{j(m, k)\})$, the marginal contributions of all players are the same when the set of players is $\mathcal{N} - \{j(m, k)\}$ and when it is $\mathcal{N} - \cup_{k=1}^K \{j(m, k)\}$. In the later case, the game is in fact the one with $M - 1$ markets and the Shapley values are given by (12).

In the initial game with \mathcal{N} , all managers are symmetric and therefore have the same value. Hence, by efficiency, we have

$$P\phi_p(\mathcal{N}) = M\pi - MK\phi_{j(m, k)}(\mathcal{N}) \quad (14)$$

Hence, (12), (13) and (14) imply

$$\begin{aligned} \phi_{j(m,k)}(\mathcal{N}) &= \frac{M\pi - MK\phi_{j(m,k)}(\mathcal{N})}{P} - \frac{(M-1)\pi}{P+K} \\ \Leftrightarrow \phi_{j(m,k)}(\mathcal{N}) &= \frac{\pi}{P+K} \end{aligned}$$

implying as claimed that $\phi_m(\mathcal{N}) = \frac{K\pi}{P+K}$ proving the induction hypothesis and the lemma. \square

7.2 Proof of Proposition 1

Part (i) follows from the fact that:

$$\sigma'(P) = \frac{M\pi K}{\mu(P+K)^2} = \frac{M\pi}{\mu(P+K)} \frac{K}{P+K} > 0$$

which is less than 1 whenever $\sigma(P) \leq P$, i.e. whenever:

$$\frac{M\pi}{\mu(P+K)} \leq \frac{P - c/\mu}{P} < 1.$$

This implies that $S(P)$ and $P - S(P)$ are increasing in P . Moreover, $\sigma''(P) < 0$, which implies (ii). Finally, $\sigma(P)/P$ is decreasing and convex in P , which implies (iii).

7.3 Proof of Lemma 2

Differentiating $\Delta(P)$, we have,

$$\begin{aligned} \Delta'(P) &= \alpha_1 KM\pi \left\{ -\frac{\sigma'(P)}{(\sigma(P)+K)^2} + \frac{1}{(P+K)^2} \right\} \\ &\propto -\frac{K}{(P+K)^2} \frac{M\pi}{\mu} \frac{1}{(\sigma(P)+K)^2} + \frac{1}{(P+K)^2} \\ &\propto \frac{1}{(P+K)^2} \left[1 - \frac{M\pi}{\mu} \frac{K}{(\sigma(P)+K)^2} \right]. \end{aligned}$$

The term in brackets is an increasing function of P since $\sigma(P)$ is an increasing function of P . Hence, the sign of $\Delta'(P)$ is “increasing” in P : if $\Delta'(P) > 0$, then $\Delta'(P')$ for all $P' > P$.

We note that $\Delta(0) < 0$ (since $\sigma(0) > 0$), $\Delta(P^{eq}) = 0$, and that $\Delta(\infty) = \mu \frac{\alpha_1 K M \pi}{M \pi + c + \mu K}$. Hence, there is a cutoff value of P such that Δ' is positive only when P is larger than this cutoff value.

Now, if $\Delta(\infty) < f$, $\Delta(P) < f$ for all P and (i) follows.

If $\Delta(\infty) > f$, there exists a unique value - strictly greater than P^{eq} - such that $\Delta(P) = f$, and at this value $\Delta(P)$ must be increasing in P , proving (ii) and (iii).

7.4 Proof of Lemma 3

Since the sign of the derivative is first positive and then negative, $u_{0-}(P)$ is single peaked. The zero of the derivative is attained at the positive root of $\mu P^2 + (4\mu K - M\pi)P + 4\mu K^2 - 2KM\pi = 0$. Simple algebra shows that the positive root is $\frac{1}{2\mu} \left(M\pi - 2K\mu + \sqrt{M\pi(M\pi + 4K\mu)} \right)$.

7.5 Proof of Corollary 1

From the definition of P^{eq} , we have

$$\mu = \frac{M\pi}{P^{eq} + K} + \frac{c}{P^{eq}}. \quad (15)$$

If $S(P) = P$, $u'_0(P) = 0$ at \hat{P} such that

$$\mu = \frac{M\pi}{\hat{P} + K} + \frac{2KM\pi}{(\hat{P} + K)^2} \quad (16)$$

Note that in (16), $\hat{P} > \frac{M\pi}{\mu} - K$ and that \hat{P} is not a function of c . When $c = 0$, $P^{eq} = \frac{M\pi}{\mu} - K$, and therefore $P^{eq} < \hat{P}$ for low enough values of c . Precisely, the result holds for all $c \leq \hat{c}$ such that $\mu \geq \frac{M\pi}{\hat{P} + K} + \frac{c}{\hat{P}}$ for then it is necessary to set $P^{eq} \leq \hat{P}$ in order to restore the equality in (15). The condition is equivalent to having c lower than a cutoff level \hat{c} :

$$\hat{c} = \hat{P} \left(\mu - \frac{M\pi}{\hat{P} + K} \right).$$

7.6 Proof of Proposition 4

The implicit function theorem applied to the expression in Proposition 4 implies that the sign of $dP^{no}(c)/dc$ is the same as the sign of:

$$\begin{aligned} & \frac{K}{(P^{no} + K)^2} M\pi \frac{d\sigma(P^{no}, c)}{dc} - 1 \\ &= \frac{K}{(P^{no} + K)^2} M\pi \frac{1}{\mu} - 1 \\ &= \frac{c}{\mu\sigma(P^{no}, c)} - 1 \\ &< 0 \end{aligned}$$

where the last equality follows the definition of $P^{no}(c)$ and the inequality the definition of $\sigma(P, c)$.

7.7 Proof of Proposition 5

We denote the partial derivative of $\sigma(P, c)$ with respect to P by $\sigma_P(P, c)$.

(i) Remember that $P^{lim}(c, f)$ solves $\Delta(P, c) = f$. As f increases, the left hand side must increase; because the left hand side is increasing in P (Lemma 2), it follows that $P^{lim}(c, f)$ is increasing in f . Since the number of strong patents $\sigma(P, c)$ is independent of f and is increasing in P , the number of strong patents also increases. Finally, the variation of padding is

$$\frac{\partial}{\partial f} (P^{lim} - \sigma(P^{lim}, c)) = \frac{\partial P^{lim}(c, f)}{\partial f} (1 - \sigma_P(P^{lim}, c))$$

which is positive if $\sigma_P(P^{lim}, c)$ is less than one. However, since $P^{lim} > P^{eq}(c)$ and since $P^{eq}(c)$ intersects the diagonal from above, the slope at P^{eq} is less than unity. By concavity of $\sigma(P, c)$ in P it is also the case that $\sigma_P(P^{eq}(c), c)$ is less than unity.

(ii) As c increases, $\sigma(P, c)$ increases by $1/\mu$; hence $\Delta(P, c)$ decreases. To restore the equality $\Delta(P, c) = f$ it is necessary to increase P , proving that $P^{lim}(c, f)$ increases with c . To facilitate the exposition we will write P^{lim} instead of $P^{lim}(c, f)$. For strong patents,

$$\frac{d\sigma(P^{lim}, c)}{dc} = \frac{\partial P^{lim}}{\partial c} \sigma_P(P^{lim}, c) + \frac{1}{\mu}$$

which is positive since $P^{lim}(c, f)$ is increasing in c . For weak patents,

$$\frac{d(P^{lim} - \sigma(P^{lim}, c))}{dc} = \frac{\partial P^{lim}}{\partial c} (1 - \sigma_P(P^{lim}, c)) - \frac{1}{\mu} \quad (17)$$

By the implicit function theorem,

$$\begin{aligned} \frac{\partial P^{lim}}{\partial c} &= -\frac{\partial \Delta / \partial c}{\partial \Delta / \partial P} \\ &= -\frac{\frac{-1}{\mu(\sigma+K)^2}}{\frac{-\sigma_P}{(\sigma+K)^2} + \frac{1}{(P+K)^2}} \\ &= \frac{1}{\mu} \frac{(P+K)^2}{(\sigma+K)^2 - \sigma_P(P+K)^2} \\ &> \frac{1}{\mu} \frac{1}{1 - \sigma_P} \end{aligned}$$

The last inequality follows the fact that when $P > P^{eq}$, $\sigma < P$. Substituting $\frac{\partial P^{lim}}{\partial c} > \frac{1}{\mu} \frac{1}{1 - \sigma_P}$ in (17) we get

$$\frac{d(P^{lim} - \sigma(P^{lim}, c))}{dc} > 0$$

Proving that when c increases, both strong and weak patents increase in numbers.

7.8 Proof of Proposition 6

Let $dP^{lim} = \frac{\partial P^{lim}}{\partial c} dc + \frac{\partial P^{lim}}{\partial f} df$. Since by (i) and (ii), the partial derivatives are positive, there exist $dc > 0$ and $df < 0$ such that $dP^{lim} = 0$. Now, we have

$$\begin{aligned} d\sigma(P^{lim}, c) &= \sigma_P(P^{lim}, c) dP^{lim} + \frac{1}{\mu} dc \\ &= \frac{1}{\mu} dc \\ &> 0. \end{aligned}$$

Obviously, because $dP^{lim} = 0$ and the number of strong patents increases, the number of weak patents must decrease.