The Inverse Cournot Effect in Royalty Negotiations with Complementary Patents*

Gerard Llobet
CEMFI and CEPR

Jorge Padilla
Compass Lexecon

First Draft: July 2016
This Version: July 7, 2017

Abstract

It has been argued that the licensing of complementary patents leads to excessively large royalties due to the well-known royalty-stacking effect. This paper shows that considering patent litigation and heterogeneity in portfolio size may lead to the opposite result due to a moderating force that we denote the Inverse Cournot effect. The lower the total royalty that a downstream producer pays, the lower the royalty that those patent holders restricted by the threat of litigation of downstream producers can charge. Interestingly, this effect is less relevant when all patent portfolios are weak, making royalty stacking more important.

JEL codes: L15, L24, O31, O34.
keywords: Intellectual Property, Standard Setting Organizations, Patent Licensing, R&D Investment, Patent Pools.

*We thank Heski Bar-Isaac, Guillermo Caruana, Yassine Lefouili, Bronwyn Hall, Louis Kaplow, Vilen Lipatov, Damien Neven, Georgios Petropoulos, Miguel Rato, Pierre Regibeau, Jan Philip Schain, Florian Schuett, Trevor Soames, and the audiences at the Hoover Institute, Toulouse School of Economics, Universitat Pompeu Fabra, University of Toronto, CRESSE 2017, the 10th SEARLE Center Conference on Innovation Economics, and WIPO for useful comments. The ideas and opinions in this paper, as well as any errors, are exclusively the authors'. Financial support from Qualcomm is gratefully acknowledged. The first author also acknowledges the support of the Spanish Ministry of Economics and Competitivity through grant ECO2014-57768 and the Regional Government of Madrid through grant S2015/HUM-3491. Comments should be sent to llobet@cemfi.es and jpadilla@compasslexecon.com.

1 Introduction

The fundamental nature of the patent system is under debate among claims on whether it fosters or hurts innovation. The main concerns focus on the impact of patent enforcement in the Information and Communications Technologies (ICT) industry. ICT products, such as laptops, tablets, or smartphones, use a variety of technologies covered by complementary patents. The royalties that must be paid for multiple patented technologies in a single product added together are said to form a harmful "royalty stack" (Lemley and Shapiro, 2007). This in turn is claimed to result in excessively high end-product prices and a reduction in the incentives to invest and innovate in product markets.

The arguments supporting royalty stacking and the need for a profound reform of the patent system rely on theoretical models which reformulate the well-known problem of Cournot-complements in a licensing framework. Cournot (1838) showed that consumers are better off when all products complementary from a demand viewpoint are produced and marketed by a single firm. In industries where each single product is covered by multiple patents, a patent holder may not fully take into account that an increase in the royalty rate is likely to result in a cumulative rate that may be too high according to other licensors, the licensees, and their customers. Since this negative externality (or Cournot effect) is ignored by all patent holders, the royalty stack may prove inefficiently high. For this reason, papers such as Lerner and Tirole (2004) have concluded that "patent pools", when they consolidate complementary patent rights into a single bundle, are generally welfare enhancing.

The Cournot effect also explains current concerns with the emergence of "patent privateers," firms that spin off patents for others to assert them. Lemley and Melamed (2013) argue that "patent reformers and antitrust authorities should worry less about aggregation of patent rights and more about disaggregation of those rights, sometimes accomplished by spinning them out to others." Similarly, "patent trolls" or "patent assertion entities" (PAEs)

- i.e. patent owners whose primary business is to enforce patents to collect royalties - are accused of imposing disproportionate litigation costs and extracting excessive patent royalties and damage awards because the existing patent system allows them to leverage even relatively small portfolios of "weak patents."[1] The America Invents Act (AIA) enacted by the US Congress in 2011 was designed in part to deal with the problems created by trolls.

The controversy about the empirical relevance of royalty stacking, or about the economic implications of the activity of patent trolls, is raging. It is, therefore, puzzling the absence of (clear-cut) evidence in support of royalty stacking given that the theoretical foundations of this hypothesis have remained unchallenged. In the Ericsson v D-Link case in front of the US Court of Appeals for the Federal Circuit, the defendants argued that in computing Ericsson's damages for the infringement of its patents, the effect of royalty stacking should be taken into account. Judge Davis considered they failed to provide evidence and rejected their claims stating that: "The best word to describe Defendant's royalty stacking argument is theoretical." In his final decision he stated that "If an accused infringer wants an instruction on patent hold-up and royalty stacking [to be given to the jury], it must provide evidence on the record of patent hold-up and royalty stacking."[2]

In this paper we develop a model of licensing complementary innovations under the threat of litigation that explains the circumstances under which royalty stacking is likely to be a problem in practice. This model departs from the extant literature in only one natural dimension; we assume that manufacturers of products covered by multiple patented technologies may challenge in court those patents and, crucially, that the likelihood that a

[1] A weak patent is defined as a patent that may well be invalid, but nobody knows for sure without conclusive litigation (see Llobet (2003) and Farrell and Shapiro (2008)).

[2] "Memorandum Opinion and Order" in the United States Court of Appeals for the Federal Circuit, Ericsson, Inc. v. D-Link Systems, Inc., available at http://www.essentialpatentblog.com/wp-content/uploads/sites/64/2013/08/13.08.06-Dkt-615-Ericsson-v.-D-Link-Order-on-Post-Trial-Motions.pdf. The final decision is available at http://essentialpatentblog.wp.lexblogs.com/wp-content/uploads/sites/234/2014/12/13-1625.Opinion.12-2-2014.1.pdf (downloaded on 8 April 2015).

judge rules in favor of the patent holder is increasing in the number and quality of its patents. This assumption is reasonable. Downstream manufacturers commonly challenge the validity of the patents that cover their products when they litigate in court the licensing terms offered by patent holders. Patent holders with large and high quality patent portfolios will not be constrained by the threat of litigation when setting royalty rates. On the contrary, owners of weak portfolios will have to moderate their royalty claims in order to avoid litigation over patent validity.

More interestingly, our analysis shows that the ability of a patent owner to charge a high royalty without triggering litigation depends on the aggregate royalty charged by all other patent holders: the higher that aggregate rate, the higher the royalty that any patent holder can charge. The intuition is that when the aggregate rate is high the expected gains from invalidating the portfolio of a patent holder are less likely to compensate for the costs incurred by the licensee. This positive relationship is a novel effect that we denote as the *Inverse Cournot effect* and we show that it is very general.[3] This effect provides incentives for unconstrained patent holders (i.e. with strong portfolios) to cut down their royalty rates to force patent holders with weak portfolios to charge, in turn, lower royalties or else face litigation. In so doing, the Inverse Cournot effect becomes a moderating force, offsetting the royalty-stacking problem that arises from the Cournot effect.

This channel becomes less effective, however, among patent holders with weak patent portfolios. To illustrate that, we consider the case in which a licensee decides to litigate patent holders in an endogenous sequence. In that case, it is still true that, by lowering the royalty rate, a patent holder can trigger litigation against other patent holders. This litigation has further consequences, though. Because when the portfolio of a patentee is invalidated the aggregate royalty rate goes down, the incentives for the downstream producer

[3] We use this term to denote a positive externality among owners of complementary inputs (in this case patents) in contrast to the standard Cournot effect which reflects a negative externality.

to litigate the remaining patentees become stronger. As a result of this *litigation cascade*, when a patent holder considers now whether to lower the royalty rate or not it ought to anticipate that, although it might benefit from a smaller royalty stack through an increase in sales, there is a greater probability of itself being litigated. Such a countervailing force implies that the Inverse Cournot effect is more important when patent holdings are more skewed – meaning that patent holders with weak portfolios co-exist with those with strong ones – leading to a lower royalty rate. As a result, we show that the royalty-stacking problem might be mitigated when facing asymmetric but stronger patent holders compared to the case of weak but more similar ones.

Taking into account the enforcement of intellectual property rights implies that whereas in the Cournot model royalty rate decisions are strategic substitutes, once litigation is taken into account these decisions become strategic complements. This paper is, thus, an illustration of the need to proceed with caution when the insights of standard industrial organization models are translated to the context of innovation. The Cournot model may be a good way to describe the pricing of complementary "widgets". When it is the right to use technology that is traded, however, the pricing decision cannot be separated from the decision to challenge the validity of the intellectual property that covers this technology, as this changes the nature of competition. In other words, the Cournot model would not be suitable to understand the pricing in a market where buyers could appropriate the widgets and ask courts not to pay for them.

The results of our model have important implications for the debate regarding standard-setting organizations (SSOs), which determine the specifications of complex products like mobile phones. In those organizations a large number of innovators declare to have Standard Essential Patents (SEPs). Because firms willing to sell a product compatible with the standard need to license all SEPs, many authors have raised concerns about the risk of royalty stacking. Interestingly, this is also a market in which patent holdings are particularly skewed. For example, in the case of the third-generation mobile phones, 80 firms declared to have SEPs, but just ten of them owned about 78% of them.[4] These are, therefore, markets in which the Inverse Cournot effect is most likely to operate and explain why, as Galetovic et al. (2015) argue, technological progress has not slowed down in spite of the large number of patent holders.

The model is also extended to capture some features specific to SSOs. Patent holders typically commit to license their patents according to Fair, Reasonable, and Non-Discriminatory (FRAND) terms. We show that accounting for these commitments and the interpretation that courts could make of them does not alter the main results of the paper. However, some recent court decisions aimed at curtailing the power of some patent holders and, thus, address royalty stacking, might have actually made the royalty-stacking problem worse by weakening the Inverse Cournot effect. We also show that the mechanism to moderate the aggregate royalty rate that this paper uncovers resembles instruments used in other technological areas such as software development (Gambardella and Hall, 2006).

Finally, we discuss how the results of our paper affect the incentives for firms to consolidate their patent holdings either through mergers or patent pools. We argue that patent pools (or mergers) among strong patent holders are likely to have the positive effects emphasized in the literature. However, mergers that involve weak patent holders, motivated in part by the aim to improve their joint power in court, might make the royalty-stacking problem worse. In fact, it could be the case that the total royalty rate increases as a result of the creation of a patent pool.

We start by introducing in section 2 a very stylized model that delivers the main insights

[4] The level of skewness is quite similar in the case of the second-generation (67 firms declared SEPs but ten owned 84% of them) and the fourth-generation (83 firms declared SEPs but ten owned 72% of them) standard.

of the paper. As we discuss in section 3, however, the mechanism driving the main results is very general and similar implications can be drawn in a more general setup, although at the cost of much greater technical complexity. In section 4 we discuss the robustness of the results to changing some of the assumptions and section 5 concludes relating this paper to the debate on patent pools and patent aggregation.

1.1 Literature Review

The literature on SSOs, in works like Lemley and Shapiro (2007), has emphasized that the licensing of complementary and essential patents by many developers could give raise to a royalty-stacking problem. This is not, however, a general result. Spulber (2016), for example, shows that when upstream firms choose quantities but negotiate royalty rates the cooperative outcome will emerge.

Our paper is also related to a long literature on the litigation between a patent holder and firms that might have infringed its probabilistic patents, including papers like Llobet (2003) and Farrell and Shapiro (2008). More recent works have aimed to capture the interaction of these conflicts in contexts like SSOs analyzing the litigation between producers and Non-Practicing Entities (NPEs). This is the case, for example, of Choi and Gerlach (2015*a*) that studies the information externalities that arise when a NPE sequentially litigates against several producers.

The papers closest to ours are Bourreau et al. (2015) and Choi and Gerlach (2015*b*). In the former, the authors study licensing and litigation in SSOs, as well as the decisions of firms to sell their IP to other innovators. The main important difference with our paper, however, is that in their setup litigation occurs after production has taken place. As a result, the total quantity produced does not depend on the outcome of this litigation and the damages paid for infringement are constant. This assumption severs the link between the licensing decision of different patent holders, eliminating the Inverse Cournot effect that plays a crucial role in our model. In their paper the strategic interaction arises from free entry and market competition.

Choi and Gerlach (2015*b*) develop a model where patent holders with weak portfolios facing the threat of litigation moderate their royalty claims so that the aggregate royalty rate falls below the one that would emerge from a patent pool. As in our model the fees charged are strategic complements. However, unlike in our model, that relationship is driven by different economic considerations and only holds off equilibrium. In their paper, the best response to a reduction in the fee charged for a complementary portfolio may be to reduce one's own fee in order to avoid being litigated and induce litigation against others. This effect disappears when one patentee has ironclad patents, i.e. precisely in those circumstances when the Inverse Cournot effect is most significant and relevant in our model.

Finally, our paper is related to the literature on patent pools. Lerner and Tirole (2004) devise a mechanism to weed out welfare-decreasing patent pools that include substitute patents and might induce collusion from welfare-increasing ones that contain only complements. This mechanism consists in allowing upstream firms to license their patents together and separately.[5] This rule leads to a unique equilibrium only when there are two patent holders. Boutin (2015) provides additional conditions on independent licensing to guarantee that there exists a unique equilibrium in which welfare-decreasing pools will not emerge. Rey and Tirole (2013) show that, if we allow for tacit coordination, independent licensing might not be enough to screen patent pools formed by substitutes. As Choi and Gerlach (2015*b*), our work contributes to this literature by showing that even if we restrict ourselves to complementary patents, patent pools may not increase welfare when they essentially include weak or litigation-constrained patent holders which increase their chances in court

[5] Lerner and Tirole (2015) generalizes the previous argument to SSOs.

when they bundle their patents, making royalty stacking more harmful.

2 The Model

Consider a market in which a downstream monopolist, firm D, sells a good to a unique consumer with a unit demand for the product. With probability α the valuation for this unit is 1. With probability $1-\alpha$ the valuation is $v<1$.

The production of the good requires firm D to use the technologies of $N=2$ different pure upstream firms. Upstream firm i holds a portfolio of x_i patents relevant for its own technology, for $i=1,2$, with $x_1 \geq x_2$. Each patent holder charges a per-unit royalty r_i to license the necessary patents to make use of that technology.[6] We denote the total royalty rate as $R \equiv r_1 + r_2$. We assume that there is no further cost of production so that the marginal cost of the final product is also equal to R.

The royalty rate for technology i is set by patent holder i as a take-it-or-leave-it offer. The downstream producer, however, might challenge in court the patents that cover the technology. Litigation between the downstream monopolist and any upstream patent holder involves legal costs L_D and L_U, respectively. As we discuss later, the downstream producer can also choose to litigate more than one patent holder, in an endogenous sequence. When indifferent the downstream producer prefers not to litigate.

The success in court is based on the size of the portfolio of the patent holder. In particular, the probability that a judge rules in favor of patent holder i, denoted as $g(x_i)$ assumed to be increasing in x_i. This assumption can be justified on several grounds. First, one of the most common ways for a downstream producer to dispute in court the licensing terms offered is to challenge the validity of the patents that cover the technology. This strategy is less likely

[6] As pointed out in Llobet and Padilla (2016) royalty-stacking problems are aggravated under per-unit royalties compared to the more frequent ad-valorem royalties, based on firm revenue. However, as discussed in section 4, if royalties are assumed to be ad-valorem the main mechanism is unaffected.

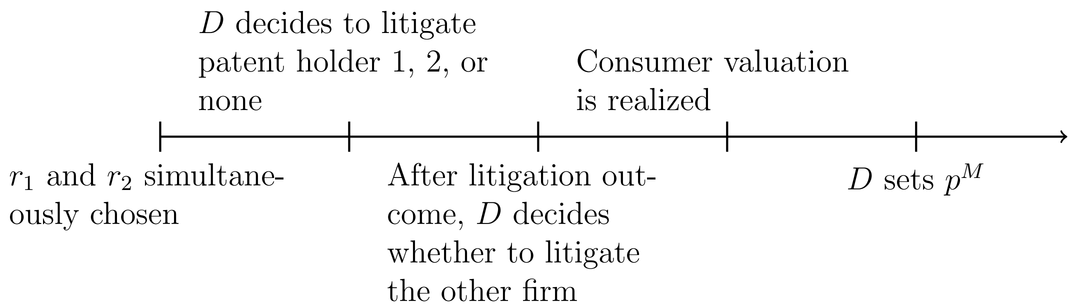

Figure 1: Timing of the model

to succeed if the patent holder is stronger, in the sense of owning a larger portfolio and/or more valuable patents. Second, patent holders do not typically defend their technology with all their patent portfolio but, rather, they choose the patents that are most likely to be upheld in court or that are more relevant for the disputed application. It is more likely to find a suitable patent for litigation if choosing from a larger patent portfolio. Finally, the model is isomorphic to one in which each upstream patent holder i holds a unique patent of quality (or a number of patents of weighted quality) x_i. To the extent that more substantial innovations translate into stronger patents, we can interpret the increasing function $g(x_i)$ as a reflection of this relationship.[7]

It is important to note that, as in Choi and Gerlach (2015*b*), if L_U is sufficiently small our structure is equivalent to a situation in which litigation is initiated by an upstream patent holder as a result of the downstream producer not paying to license its portfolio. In other words, if the legal costs are sufficiently small (compared to the size of the market, of course), the producer by refusing to pay the license anticipates that it will brought to court by the patent holder for patent infringement and the outcome is therefore equivalent to this firm having initiated the lawsuit, as assumed here.

The timing of the model is described in Figure 1. First, upstream patent holders simultaneously choose their royalty rates. In the second stage the downstream producer chooses

[7] For simplicity we abstract from situations in which upstream patent holders own the rights for technologies that might be infringed by other upstream patent holders.

which patentees to litigate (if any) and the sequence. In the final stage, once litigation has been resolved, the valuation of the consumer is drawn and the downstream producer chooses the price for the final good.[8]

This timing implies that the downstream producer will always choose a price equal to the realized valuation of the consumer. That is, given R the downstream producer captures all the surplus without generating the losses associated to double marginalization. As a result, expected downstream profits Π_D can be computed as

$$\Pi_D(R) = \begin{cases} \alpha + (1-\alpha)v - R & \text{if } R \leq v, \\ \alpha(1-R) & \text{if } R \in (v,1], \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Notice that these profits are decreasing and convex in R.[9] These are general properties that engender many of the results of the paper as we will see in section 3.

We now characterize the equilibrium of the game depending on the strength of the patent portfolio of each firm. We start with the case in which the parameters imply that litigation never plays a role in the model. This assumption will give raise to the standard *royalty-stacking* result in the literature that we reproduce next.

2.1 Strong Patent Portfolios

Suppose that both patent holders have a portfolio sufficiently strong so that $g(x_1) = g(x_2) = 1$. In this case, litigation by the downstream producer will never be a credible threat.[10] We start by characterizing the royalty rate that maximizes joint profits for the upstream patent holders. This royalty will be used as a benchmark for the case in which patent holders choose

[8] Although, in practice litigation may take several years, the results of the paper are qualitatively unaffected as long as the final-good producer anticipates that after the court decision the price will be adjusted according to the resulting royalty rate.

[9] A dead-weight loss would arise if we assumed that the downstream producer chose the price before the demand is realized. In that case, the threshold value on R in the profit function $\Pi_D(R)$ would change. That is, $p^M(R) = v$ if and only if $R \leq \tilde{R} \equiv \frac{v-\alpha}{1-\alpha} < v$. Since double-marginalization does not interact with the mechanisms explored in this paper, the main results would go through under this alternative assumption although at the cost of more technical complexity.

[10] The same results would arise if, instead, we assumed that L_D is sufficiently high.

their royalty rate independently.

Lemma 1. *The aggregate royalty rate that maximizes total patent holder profits is* $R^M = v$ *if* $v \geq \alpha$ *and* $R^M = 1$ *otherwise.*

The more likely the demand is equal to v (which occurs with probability $1 - \alpha$) or the higher is v, the more likely it is that it is profitable for the patent holders to cater all the demand by choosing a low royalty rate. Notice also that, due to the unit-inelastic demand, the royalty rate $R = v$ also maximizes total social welfare.

We now turn to the situation in which firms choose their royalty rate independently. As in the previous case, it is easy to see that any undominated Nash equilibrium should involve royalties r_1 and r_2 such that $r_1 + r_2$ are either equal to v or to 1.[11]

Proposition 2. *There is a continuum of undominated pure-strategy equilibria. The corresponding royalty rates* (r_1^u, r_2^u) *can be characterized as follows:*

1. *If* $v \geq \frac{2\alpha}{1+\alpha}$, $R^u = r_1^u + r_2^u = v$ *with* $r_i^u \leq \frac{v-\alpha}{1-\alpha}$ *for* $i = 1, 2$.

2. *If* $v \leq \frac{1+\alpha}{2}$, $R^u = r_1^u + r_2^u = 1$ *with* $r_i^u \geq \frac{v-\alpha}{1-\alpha}$ *for* $i = 1, 2$.

Both kinds of equilibria co-exist when $\frac{2\alpha}{1+\alpha} \leq v \leq \frac{1+\alpha}{2}$.

Intuitively, the equilibrium with total royalty of 1 is likely to exist when v is small and α is sufficiently close to 1. A deviation might exist if any patent holder prefers to decrease the royalty rate in order to cater the consumer regardless of her valuation. This deviation is illustrated in Figure 2. Given r_2^u, patent holder 1 can choose to stick with $r_1^u = 1 - r_2^u$ or deviate and choose $\hat{r}_1 = v - r_2^u$ so that the probability of selling increases from α to 1. Such a deviation is unprofitable if r_2^u is sufficiently large and, thus, the low $\hat{r}_1$ does not allow the

[11] As usual, in this family of models there is also a continuum of weakly dominated Nash Equilibria in which $r_1 \geq 1$ and $r_2 \geq 1$. As these are uninteresting, we will ignore them throughout this paper and we will denote the remaining ones as just Nash Equilibria.

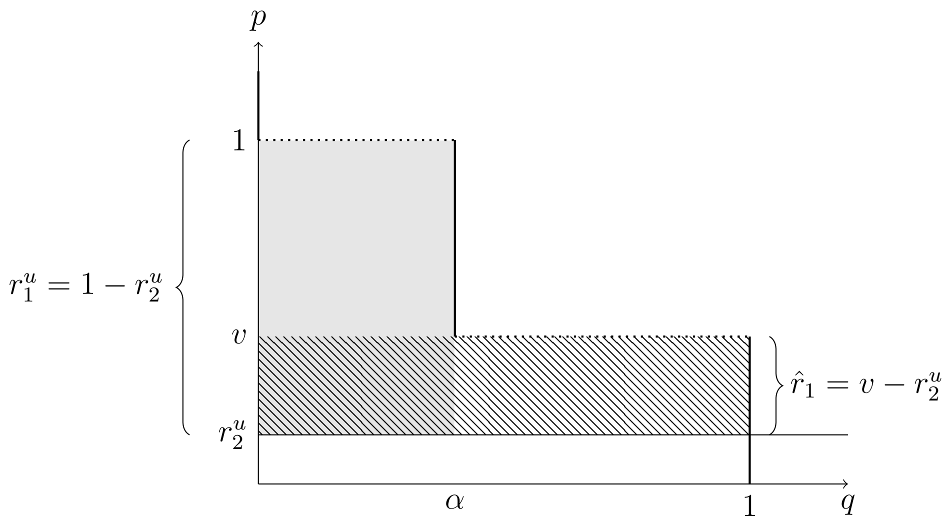

Figure 2: Deviation by patent holder 1 from $r_1^u + r_2^u = 1$.

firm to benefit from the increase in sales. In the limit, when $v = 0$ or $\alpha = 1$ this equilibrium holds for any combination of royalties that sum up to 1.

Similarly, equilibria with a total royalty equal to $R^u = v$ are likely to exist when v is sufficiently high and α is sufficiently small. This time a deviation aims to capture the additional surplus when consumer valuation is 1, even if this surplus is materialized only with probability α. To prevent these deviations each patent holder must charge a modest royalty so that the other firm already obtains sufficiently high profits in equilibrium, thus reducing the appeal of raising the royalty rate and reducing the probability of sale. In the limit, when $v = 1$ or $\alpha = 0$ any combination of royalties that sum up to v would constitute an equilibrium.

The next result shows that the equilibrium total royalty – and the corresponding final-good price – might be higher than in the case in which royalties were chosen by a monopolist. In other words, there are values of v for which patent holders would separately induce a total royalty $R^u = 1$ and the final price would become $p^M(R^u) = 1$ and yet they would benefit from coordinating and choosing a royalty rate $R^M = v$, leading to a final price $p^M(R^M) = v$.

Corollary 3 (Royalty Stacking). *When $\alpha < v \leq \frac{1+\alpha}{2}$ inefficient equilibria with $r_1^u + r_2^u =$*

$R^u = 1$ exist even though joint profits are maximized when the total royalty is v. When $\alpha \le v < \frac{2\alpha}{1+\alpha}$, all equilibria lead to $R^u = 1$. However, there are no parameter values for which $R^M = 1$ but $R^u = v$.

This is a version of the Cournot-complements result in which firms choosing quantities of complementary products induce final prices even higher than the monopoly one. The intuition has already been discussed in the context of patent licensing and it has been referred to, in papers like Lemley and Shapiro (2007), as the royalty-stacking problem. The decision of a patent holder to increase the royalty rate trades off the higher margin with the lower quantity sold but without internalizing the fact that this decrease in quantity has a negative effect on the royalty revenues of the other patent holder.

This result holds for a generic number of firms and under general assumptions regarding the demand functions and it emerges whenever litigation is irrelevant, as most of the previous literature has implicitly assumed.[12] As a result, whereas the profit-maximizing rate is independent of the number of firms, in the equilibrium we have that the royalty-stacking problem becomes more severe when the total number of patents is fragmented in the hands of more firms. Also importantly, if litigation is irrelevant, meaning that patents are always enforced, the size of a patent portfolio also becomes irrelevant and each patent holder should charge the same royalty rate. This prediction, however, seems quite implausible in practice.

We now discuss the effects of the litigation threat. We analyze two prototypical situations. First, we consider the case in which only one patentee is constrained by this threat. Later we study the situation in which both patentees are equally constrained.

[12] As we discuss in section 3, a sufficient condition is the log-concavity of the demand function. This condition guarantees that the patent holder's problem is quasiconcave and also that royalty rates become strategic substitutes, which is enough for this result to arise.

2.2 One Constrained Patent Holder

Suppose that $g(x_1) = 1$ and $g(x_2) < 1$ so that the downstream producer may only be interested in litigating patent holder 2. We restrict our discussion to the case where $v \in \left(\alpha, \frac{1+\alpha}{2}\right]$ so that, according to Corollary 3, in the previous benchmark a combination of royalties for which $r_1^* + r_2^* = 1$ constituted an equilibrium with royalty stacking.

When litigation is feasible, the first additional condition that r_2^* must satisfy is that

$$(1 - g(x_2)) \left[\Pi_D \left(1 - r_2^*\right) - \Pi_D(1)\right] \leq L_D. \tag{2}$$

That is, it is not profitable for the downstream producer to go to court against patentee 2. In this expression the downstream firm trades off the legal costs, L_D, with the increase in profits when the portfolio of patent holder 2 is invalidated and the royalty r_2^* goes to zero. Portfolio invalidation occurs with probability $1 - g(x_2)$ and it increases profits from $\Pi_D(1)$ to $\Pi_D(1 - r_2^*)$. Using (1) we have that $\Pi_D(1) = 0$, $\Pi_D\left(1 - r_2^*\right) = \alpha r_2^*$ if $1 - r_2^* \geq v$, and $\Pi_D\left(1 - r_2^*\right) = \alpha + (1-\alpha)v - (1 - r_2^*)$ if $1 - r_2^* < v$, meaning that the previous condition will hold if r_2^* is sufficiently small. In particular, litigation against patent holder 2 is unprofitable given $r_1^* + r_2^* = 1$ if

$$r_2^* \leq \bar{r}_2 = \begin{cases} \frac{L_D}{\alpha(1-g(x_2))} & \text{if } \frac{L_D}{1-g(x_2)} < \alpha(1-v), \\ (1-\alpha)(1-v) + \frac{L_D}{1-g(x_2)} & \text{otherwise.} \end{cases} \tag{3}$$

Failure of the previous condition constitutes a sufficient but not necessary condition to rule out some equilibria with royalty stacking. In particular, suppose that both patent holders choose a royalty rate higher than $\frac{v-\alpha}{1-\alpha}$, so that the conditions that guarantee an equilibrium with $R = 1$ in Proposition 2 are satisfied. Furthermore, suppose that equation (2) holds, so that the downstream producer is not interested in litigating patent holder 2. As we discuss next, there might still be strategic considerations that compel patentee 1 to deviate and choose an alternative royalty rate $\hat{r}_1$ that induces litigation against the other patentee, leading to a reduction in the total royalty rate. In particular, given r_2 patentee 2 will be litigated if $\hat{r}_1 \leq \bar{r}_1(r_2)$, implicitly determined by

$$(1 - g(x_2)) \left[\Pi_D (\bar{r}_1) - \Pi_D(\bar{r}_1 + r_2)\right] = L_D. \tag{4}$$

Replacing the profit function of the downstream producer we have that

$$\bar{r}_1(r_2) = v + \frac{\alpha}{1-\alpha} r_2 - \frac{L_D}{(1-\alpha)(1-g(x_2))} \text{ if } \underline{r}_2 < r_2 \leq \bar{r}_2, \tag{5}$$

where $\bar{r}_1(r_2) \leq v$ and $\bar{r}_2$ is defined in (3). It is important to point out that if the royalty rate of patentee 2 is sufficiently low, defined as

$$r_2 < \underline{r}_2 = \begin{cases} \frac{L_D}{1-g(x_2)} & \text{if } \frac{L_D}{1-g(x_2)} \leq v, \\ \frac{L_D}{\alpha(1-g(x_2))} - \frac{1-\alpha}{\alpha} v & \text{otherwise,} \end{cases} \tag{6}$$

even a royalty $\hat{r}_1 = 0$ is not enough to induce litigation by the downstream producer since the gain from facing a 0 royalty rate is smaller than the legal costs involved. For higher values of r_2 litigation against patentee 2 arises for $\hat{r}_1$ sufficiently small. The threshold value $\bar{r}_1(r_2)$ is weakly increasing in r_2 and weakly decreasing in L_D. The positive relationship between r_2 and $\bar{r}_1$ makes royalty rates strategic complements.

Suppose now that patent holder 1 chooses a royalty rate $\hat{r}_1 < \bar{r}_1(r_2)$ so that the downstream producer has incentives to litigate patent holder 2. The threat has a moderating effect on this patentee, which can avoid litigation by lowering r_2. We call this mechanism the *Inverse Cournot effect*. This is one of the main insights of this paper and it constitutes the reason why an equilibrium with royalty stacking might fail to exist in the presence of a litigation threat. As we discuss later, this relationship is very general and it applies to demand functions of all classes and to a generic number of firms.

Following the previous argument, patent holder 1 might benefit from a royalty rate $\hat{r}_1$ below $\bar{r}_1$ only if, by causing litigation against patentee 2, it induces an expansion in the quantity sold from α to 1 with probability $1 - g(x_2)$. Hence, $\hat{r}_1$ must be lower than v. Since

$\bar{r}_1(\bar{r}_2) \leq v$ it follows that the optimal deviation for patent holder 1 when patentee 2 sets $r_2^* \leq \bar{r}_2$ is the highest royalty rate which guarantees that patentee 2 is litigated, $\hat{r}_1 = \bar{r}_1(r_2^*)$.[13] Patent holder 1's profits in that case would become

$$\hat{\Pi}_1 = [\alpha + (1-\alpha)(1-g(x_2))]\,\hat{r}_1. \tag{7}$$

That is, a deviation will lead to profits equal to $\hat{r}_1$ either because the valuation of the consumer is 1 or because the valuation is v but patent holder 2 is successfully litigated by the downstream producer. This deviation will not take place if profits, $\hat{\Pi}_1$, are lower than those in the candidate equilibrium, $\Pi_1^* = \alpha r_1^*$. Notice that the lower are r_1^* or $g(x_2)$ the more binding this condition becomes. The next proposition characterizes the circumstances under which it is not possible that $\Pi_1^* \geq \hat{\Pi}_1$ holds while, as Proposition 2 requires, $r_2^* \geq \frac{v-\alpha}{1-\alpha}$. In those situations, an equilibrium with royalty stacking will fail to exist.

Proposition 4. *Suppose that $v > \alpha$. If $\frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$ there is no pure strategy equilibrium with royalty stacking. However, if $g(x_2)$ is sufficiently small the efficient equilibrium exists and it involves $r_2^* \leq \frac{L_D}{1-g(x_2)} < v$ and $r_1^* = v - r_2^*$.*

The previous result indicates that when L_D and/or $g(x_2)$ are sufficiently low, royalty stacking will not arise in equilibrium. That is, in instances in which a monopolist patent holder prefers to choose a royalty $R^M = v$ – when $v > \alpha$ – there would be no equilibrium with $R^* = 1$.

In order to interpret this result it is useful to start by considering the case under which such an equilibrium may exist. From (6) we know that if $r_2^* \leq \frac{L_D}{1-g(x_2)}$ the Inverse Cournot effect has no bite since there is no positive value of $\hat{r}_1$ that triggers litigation. When $\frac{L_D}{1-g(x_2)} \geq \frac{v-\alpha}{1-\alpha}$ it is also possible to find $r_2^* \geq \frac{v-\alpha}{1-\alpha}$, satisfying the conditions of Proposition 2. Hence, it

[13] More precisely, given our assumptions, $\hat{r}_1$ should be slightly lower than $\bar{r}_1(r_2^*)$

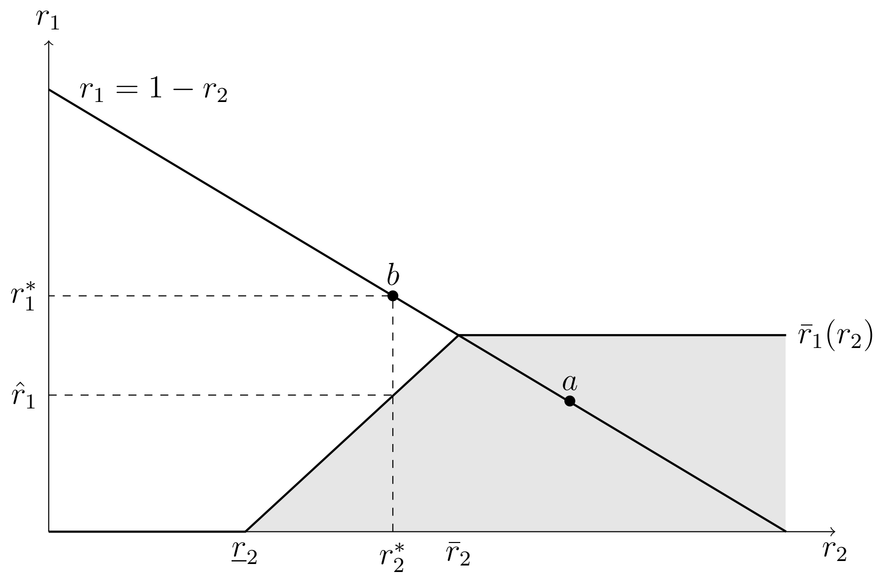

Figure 3: Deviation from an equilibrium with Royalty Stacking.

is optimal for patent holder 1 to choose $r_1^* = 1 - r_2^*$ and an equilibrium with royalty stacking will arise in that case.

This proposition shows that the condition $\frac{L_D}{1-g(x_2)} \geq \frac{v-\alpha}{1-\alpha}$ is not only sufficient but also necessary for a royalty-stacking equilibrium to exist. In other words, consider a combination of royalties (r_1^*, r_2^*) with $r_1^* + r_2^* = 1$ such that $r_i^* \geq \frac{v-\alpha}{1-\alpha}$ for $i = 1, 2$. If $r_2^* \geq \bar{r}_2$, we know that patent holder 2 is litigated since $r_1^* = 1 - r_2^* < \bar{r}_1(r_2^*)$. This combination of royalties is illustrated as point a in Figure 3, where the shaded area shows all the royalty pairs (r_1, r_2) that induce litigation against patentee 2. If, instead, $r_2^* \leq \bar{r}_2$ so that $r_1^* \geq \bar{r}_1(r_2^*)$ – point b in the figure – absent any strategic considerations, patentee 2 would not be litigated. However, the previous proposition indicates that patent holder 1 could always increase profits by lowering the royalty rate – and choose $\hat{r}_1$ as indicated in the figure – and, due to the Inverse Cournot effect, foster litigation against patent holder 2 by the downstream producer. The reason for this result is, precisely, that when $v > \alpha$ total profits increase when there is no royalty stacking and patentee 1 expects to appropriate this increase in total surplus.

The second part of the proposition also indicates that when the probability of success in court of patentee 2 is small two results concur. First, the royalty rate is commensurate to the strength of the patent portfolio and the cost of challenging those rights by the downstream producer, $r_2^* \leq \frac{L_D}{1-g(x_2)}$. This result arises from the fact that when $g(x_2)$ is small patentee 2 must choose a low royalty rate to discourage the downstream producer from engaging in litigation that will, most likely, result in a zero royalty. Second, and more interestingly, the profit maximizing equilibrium, consisting of $R^M = v$, may exist. The reason is that the low value of r_2 makes patent holder 1 the residual claimant of the surplus generated. This can be seen using Figure 2, where we showed that when r_2 is low patent holder 1 internalizes the losses that a deviation towards a larger royalty rate entail.

2.3 Two Constrained Patent Holders

Suppose now that both firms have identical patent holdings which do not confer full protection against litigation, $g(x_1) = g(x_2) = g(x) < 1$. As in the previous case we focus on the situation in which royalty stacking was an equilibrium when no litigation was feasible, $v \in \left(\alpha, \frac{1+\alpha}{2}\right]$. As opposed to what happened in the previous case, litigation here might involve one or both upstream patent holders. We assume that litigation occurs in sequence and this sequence is chosen by the downstream producer. Importantly, the decision of whether to litigate a second patent holder or not might be contingent on the first court decision.

As in the previous case, we study whether litigation affects the existence of an equilibrium with royalty stacking, so that $r_1^* + r_2^* = 1$. For the purpose of presenting the results in this section it is enough to focus on the symmetric case in which $r_1^* = r_2^* = 1/2$ as if this equilibrium did not exist no asymmetric equilibrium would exist either.[14] Suppose first that only patent holder 2 is litigated. Using (2), the expected gain of the downstream producer from going

[14] As discussed in previous sections, an equilibrium may fail to exist because one of the royalties is too low and, as a result, either the patent holder decides to deviate and raise it even at the cost of being litigated or the other patent holder may benefit from lowering its own royalty and serve the whole market. By focusing on the symmetric royalty rate we are minimizing the profitability of these deviations.

to court, to be compared with the cost L_D, is equal to $(1 - g(x))\,[\Pi_D\,(1/2) - \Pi_D(1)]$.

Suppose that after patent holder 2 has been litigated the downstream producer is considering whether to also litigate patentee 1 or not. Since litigation is sequential, the decision ought to be contingent on the success in court against patentee 2. If that patentee prevails, the expected gains from another trial are identical to the ones described above. This implies that if litigating patentee 2 only were profitable, litigating patentee 1 after the downstream producer had been defeated in court in the first stage would be equally profitable.

Suppose now that patentee 2 lost in court, which occurs with probability $1 - g(x)$. The expected profits of the downstream producer of litigating against patentee 1 are now evaluated when $r_2 = 0$ and they would be equal to $(1 - g(x))\,[\Pi_D(0) - \Pi_D\,(1/2)]$. Since $\Pi_D(R)$ is convex, $\Pi_D\,(1/2) - \Pi_D(1) \leq \Pi_D(0) - \Pi_D\,(1/2)$. That is, litigating patentee 1 after victory against patentee 2 would always be more profitable than if the downstream producer had lost.

An important implication of this result is that in the symmetric case it will never be optimal to litigate one of the patent holders only. That is, it would also be at least as profitable to litigate the other one. For this reason, when $r_1^* = r_2^* = 1/2$, the downstream producer will prefer to avoid going to court against patent holder 2 if and only if

$$(1 - g(x))\,[\Pi_D\,(1/2) - \Pi_D(1)] + (1 - g(x))\{(1 - g(x))[\Pi_D(0) - \Pi_D(1)] - L_D\} \leq L_D. \quad (8)$$

The first term in the previous expression is identical to the one that governs the decision of the downstream producer to litigate in the case of one constrained patent holder, as described in equation (2). The second term captures the *option value* that litigation now may bring. That is, if the downstream producer wins the first trial the profitability of going to court against the other patent holder increases. We call this effect a *litigation cascade*.[15]

In order to interpret this constraint it is useful to consider the situation in which the condition is satisfied with equality and the downstream producer is indifferent between engaging in litigation or not. In this scenario, equation (8) implies that litigating patentee 2 only must result in an increase in expected market revenues lower than the cost L_D or, else, litigating both patent holders would lead to strictly positive profits. Since litigation against patentee 2 is unprofitable and the problem the downstream producer faces against patentee 1 is the same when it has not succeeded in court before, it will only litigate a second time upon an initial success. Indifference between going to court or not implies, thus, that the profits from this second trial, which occurs with probability $1 - g(x)$, must compensate the losses from the first one.[16] That is, when indifferent between bringing a patent holder to court or not, the downstream producer is motivated to litigate only due to the prospect of invalidating the portfolio of both patent holders.

From the previous arguments it is immediate that equation (8) is less likely to be satisfied than the one that drives the decision to litigate patent holder 2 when only this firm is constrained, as illustrated in equation (2). The downstream producer benefits from having the option to litigate against a second patent holder contingent upon the success of the first trial. This comparison would suggest that before we introduce strategic considerations in the patent holders' royalty choices – that is, before we account for the optimal response of the patent holders to the increased litigation risk associated with that option –, royalty stacking is less likely when they both have a weak portfolio. This conclusion is inaccurate and, as we

[15] Notice that using the possibility to delay litigation against one patent holder until the outcome of the previous trial has been revealed is always optimal for the downstream producer. In practice, litigation might take years and the firm might decide to engage in a second trial when the first one has not concluded but the information uncovered during the process indicates that the revised probability of success is sufficiently high. The implications of such a strategy are very similar to the sequential setup assumed here.

[16] Notice that here we are abstracting from the informational spillovers that a court outcome may have on future court rulings.

will see next, once we introduce these strategic considerations the opposite may hold.

In particular, suppose that the litigation constraint in (8) is not satisfied and, thus, it is unprofitable for the downstream producer to go to court if patent holders charge a royalty rate $r_1^* = r_2^* = 1/2$. We now consider the incentives for patentee 1 to deviate. In order to simplify the exposition we will consider only the case in which if litigation occurs in equilibrium the downstream producer prefers to start by challenging the portfolio of the patentee with the highest royalty rate. As we prove later in the paper in a more general setup – see Lemma 8 – this is the only relevant situation, since it is the order that maximizes profits for the downstream producer.

As in the previous case, a necessary condition for a deviation by patentee 1 to be profitable is that it spurs litigation against patentee 2. Because the downstream producer litigates first the patent holder with the highest royalty rate, such a deviation must entail a decrease in r_1. It turns out that, in spite of the lower royalty, patentee 1 might now be litigated afterwards. The reason is that although it is not profitable to litigate patent holder 1 initially, it might be worthwhile to do it if and when the downstream producer prevails against patentee 2, which occurs with probability $1 - g(x)$.

The next lemma characterizes the threshold values of $\hat{r}_1$ for which patentee 1 expects to be litigated in case patentee 2 loses in court.

Lemma 5. *Suppose that under $r_1^* = r_2^* = \frac{1}{2}$ it is not profitable for the downstream producer to engage in litigation. If by deviating to $\hat{r}_1 < r_1^*$ patent holder 2 is litigated, patent holder 1 will also be litigated if and only if patent holder 2 lost in court and $\hat{r}_1 > \frac{L_D}{1-g(x)}$.*

The deviations that this lemma characterizes determine two regions depending on whether $\hat{r}_1$ is higher or lower than $\frac{L_D}{1-g(x)}$. Both deviations are less profitable than in the case in which $g(x_1) = 1$, albeit for different reasons. In one of the regions, by choosing a low $\hat{r}_1$, patentee 1 eludes litigation but at the cost of reducing the royalty revenues that the firm might obtain.

In the second region, when $\hat{r}_1$ is higher, the lower profitability of the deviation arises from the probability that the patent holder might not accrue any licensing revenues from the portfolio if the court declares it invalid, together with the corresponding litigation costs. In particular, in this last region, the profits from a deviation are

$$\hat{\Pi}_1 = g(x)\alpha\hat{r}_1 + (1-g(x))\left[g(x)\hat{r}_1 - L_U\right].$$

When the portfolio of the other patent holder is upheld in court the expected quantity is α. When the portfolio of patentee 2 is invalidated and the downstream producer also decides to litigate patent holder 1, the quantity sold is 1 but the royalty $\hat{r}_1$ is only paid if the portfolio is upheld in the second trial.

It is easy to see that the risk of a litigation cascade might foster the existence of an equilibrium with royalty stacking. As an illustration, take the case in which L_U is significant, which makes the threat of litigation particularly relevant for the upstream patent holders, and consider the situation in which $v \leq \frac{1}{2}$.

Given $r_1^* = r_2^* = 1/2$, two conditions must be satisfied for such an equilibrium to exist. First, equation (8) should not hold, so that the downstream producer is not interested in litigating, which in this case implies

$$\frac{L_D}{1-g(x_2)} \geq \frac{1}{2-g(x)}\left[g(x)\frac{\alpha}{2} + (1-g(x)(\alpha+(1-\alpha)v)\right]. \tag{9}$$

Second, the cost of a litigation cascade implies that the optimal deviation of patent holder i, for $i = 1, 2$, involves $\hat{r}_i = \min\left\{v, \frac{L_D}{1-g(x)}\right\}$ and such a deviation is unprofitable if and only if $\hat{\Pi}_1 \leq \Pi^*$ or

$$\left[\alpha + (1-\alpha)(1-g(x))\right]\hat{r}_i \leq \frac{\alpha}{2}. \tag{10}$$

Notice that, as in the case of one constrained patent holder $\hat{r}_i \leq v$ so that demand expands if the portfolio of the other patent holder is invalidated.

These two conditions provide a lower and upper bound, respectively, on $\frac{L_D}{1-g(x)}$ for an equilibrium with royalty stacking to exist. That is, the legal costs of the downstream producer must be sufficiently large to discourage this firm from litigating but they must also be sufficiently small so that the decrease in the royalty rate necessary for a deviating firm to fend off litigation is large.

Although the previous conditions are highly non-linear in the main parameters of the model it is easy to see that it is possible to find combinations that satisfy them. More interestingly, we can also find situations in which this equilibrium with a total royalty equal $R^* = 1$ is sustainable when both patent holders have a very strong or a very weak portfolio but not in the case in which firms are asymmetric.

Example 1. *Consider the parameter values* $\alpha = 0.1$*,* $v = 0.3$*,* $g(x_2) = 0.7$*,* $L_D = 0.035$*, and* L_U *sufficiently large. If litigation were not possible, the parameter values would satisfy the conditions of Proposition 2 and an equilibrium with royalty stacking,* $R^u = 1$*, would exist.*

Next, consider the case in which $g(x_1) = 1$ *so that only the second patent holder is potentially constrained. By construction,* $\frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$*, and according to Proposition 4, the royalty-stacking equilibrium does not exist in this case.*

Finally, consider the case in which $g(x_1) = g(x_2) = 0.7$*. It can be verified that equations* (9) *and* (10) *are satisfied and, thus, the royalty-stacking equilibrium exists when both patent holders are similarly constrained.*

We can, thus, conclude that once we introduce litigation in the model the royalty rate is not necessarily monotonic in the strength of the patent portfolios. In fact, we have just shown that when portfolios are weaker but patents are more evenly distributed the royalty stacking problem might become more relevant.

3 Generality of the Results

We now show that the main forces at work in the previous model hold more generally. In particular, we assume a continuously differentiable demand function $D(p)$.[17]

Consider the case of N patent holders. Each of them sets a royalty rate r_i for $i = 1, .., N$, so that $R = \sum_{i=1}^{N} r_i$. The expression for profits of the downstream producer arises from

$$\Pi_D(R) = \max_p \, (p - R)D(p).$$

Standard calculations show that the optimal price $p^M(R)$ is increasing in R and, therefore, the profit function is decreasing and convex in R: $\Pi'_D(R) = -D(p^M) < 0$ and $\Pi''_D(R) = -D'(p^M)\frac{dp^M}{dR}(R) > 0$.

In order to guarantee that the profit function of the patent holders is well-behaved with respect to the royalty rate we introduce the following standard regularity condition.

Assumption 1. *$D(p^M(R))$ is log-concave in R.*

The profits of patent holder i can be defined as

$$\Pi_i(R_{-i}) = \max_{r_i} \; r_i D\left(p^M\left(\sum_{j=1}^{N} r_j\right)\right)$$

where $R_{-i} = \sum_{j \neq i} r_j$. We denote the royalty rate that corresponds to the Nash Equilibrium of the game when firms are unconstrained by litigation as $r_i^u = r^u$ for all i. It can be obtained from

$$D(p^M(Nr^u)) + r^u D'(p^M(Nr^u))\frac{dp^M}{dR}(Nr^u) = 0. \tag{11}$$

For completeness, we reproduce the standard royalty-stacking result. It is important to notice that Assumption 1 not only guarantees concavity of the patent holder's problem but it also implies that royalty rates are strategic substitutes, delivering the result.

[17] A minor difference here is the assumption that the downstream producer chooses a unique price and, therefore, an inefficiency due to double marginalization will arise for any price $p > 0$. This difference will only have implications in the welfare interpretations of the model.

Proposition 6 (Royalty Stacking)**.** *If litigation is sufficiently costly for the downstream producer, in the unique equilibrium of the game all patent holders choose* $r_i^u = r^u$*, defined by* (11)*, independently of the size of their portfolio. In this equilibrium* $r^u(N)$ *is decreasing in* N *but* $R^u(N) = Nr^u(N)$ *and* $p^M(R^u(N))$ *are increasing in* N*.*

We now discuss how the two main forces that drive the results in the previous section generalize in this context. We start by talking about the Inverse Cournot effect and we later analyze how litigation cascades manifest in more general demand setups. For simplicity we return to the $N = 2$ case.

3.1 The Inverse Cournot Effect

We first generalize the results of the previous section to the case in which only patentee 2 is constrained by litigation. That is, we assume that $g(x_2) < g(x_1) = 1$. As in the benchmark model, the downstream producer prefers not to litigate patentee 2 if and only if

$$(1 - g(x_2)) \left[\Pi_D(r_1) - \Pi_D(r_1 + r_2)\right] \leq L_D. \tag{12}$$

Litigation will be unprofitable if the expected gains from avoiding to license the patent portfolio of patentee 2 are lower than the legal costs involved. The highest royalty that induces litigation against patentee 2, $\bar{r}_1$, is still determined by (4). The next lemma characterizes how this threshold on the royalty rate depends on the parameters of the model.

Lemma 7. *The downstream producer will litigate patent holder* 2 *if* $r_1 < \bar{r}_1(L_D, x_2, r_2)$*, as defined by* (4)*. This threshold royalty* $\bar{r}_1$ *is strictly increasing in* r_2 *and strictly decreasing in* L_D *and* x_2*.*

This result implies that the decision to litigate a patent holder also depends on the royalty rate set by the other patent holder and it generalizes the expression in equation (5) for the benchmark model. Denoted before as the *Inverse Cournot effect*, it implies that if r_1 is high, profits for the downstream producer are low, independently of whether the patent portfolio of firm 2 is upheld in court or not. Thus, it is less likely that the gains from litigation offset the legal costs involved. In the benchmark model we showed that this effect is an important counterbalancing force to the conventional Cournot effect and it is, indeed, a reason why a royalty-stacking equilibrium would fail to exist when $\frac{L_D}{1-g(x_2)}$ took an intermediate value, but it would arise if strategic decisions were not taken into account. Of course, this effect immediately generalizes to the case of N patent holders with a portfolio sufficiently strong so that they will never be litigated. In that case, the Inverse Cournot effect would indicate that the highest royalty that patentee 2 can charge is increasing in the sum of the royalty of all the other patent holders, denoted as R_{-2}.

A direct consequence of this mechanism is that higher legal costs or a stronger portfolio of patentee 2 makes this constraint less relevant. Patentee 1 needs to set an even lower royalty to make litigation against patentee 2 profitable for the downstream producer. As a result, a deviation is less likely to be profitable and royalty stacking is more likely to arise in equilibrium.

In any equilibrium with royalties r_1^* and r_2^* patentee 2 will avoid being litigated if (12) holds. However, this condition also implies that there will never be a Nash Equilibrium in which the downstream producer is indifferent between litigating patentee 2 or not. The reason is that patentee 1 always has incentives to lower slightly the royalty rate, so that (12) does not hold and induce litigation on patentee 2. At essentially no cost, it becomes, with probability $1 - g(x_2)$, the only firm licensing the technology. This deviation is profitable as it generates a discrete reduction in the royalty stack. If, instead, equation (12) holds with inequality, patent holder 2 will raise its royalty unless it is equal to r^u. A consequence of this insight is that unless L_D is so high that litigation is irrelevant when $r_1^* = r_2^* = r^u$, there will be no pure-strategy equilibrium.

This is in contrast with what occurs in our benchmark model. In that case, an equilibrium in pure strategies other than the one that generated royalty stacking, $R^u = r_1^u + r_2^u$, could arise when $\frac{L_D}{1-g(x_2)}$ was small because demand was constant when prices were sufficiently low. For this reason, when $R \leq v$ no patent holder had incentives to lower its royalty rate since it would not generate an increase in demand.

In the general case, when demand is strictly decreasing in the price and L_D is sufficiently small, only a Nash equilibrium in mixed strategies will exist. Patent holders randomize in a support $[r_i^L, r_i^H]$ and according to a distribution $F_i(r_i)$ (with density $f_i(r_i)$) for $i = 1, 2$. Patentee 2 when choosing a higher r_2 trades off a lower probability of being litigated with a higher payoff when litigation occurs but the firm succeeds in court. This trade-off means that patentee 2 will choose a lower expected royalty rate than when litigation was not a threat. In the case of patentee 1 two effects go in opposite directions. On the one hand, due to Inverse Cournot effect the patent holder has incentives to lower the royalty rate r_1 in order to enjoy monopoly profits with a higher probability. On the other hand, there is a probability that the portfolio of the other patent holder is invalidated and, since in that case royalty rates are strategic substitutes, it is optimal to raise r_1. Our benchmark model suggests that the first effect is likely to dominate and the royalty rate is likely to be lower when litigation is a relevant threat.

3.2 Litigation Cascades and its Strategic Effects

We now turn to the case in which both patent holders have a weak portfolio, with $g(x_1) = g(x_2) = g(x) < 1$. In the next lemma we describe the order in which the downstream producer might litigate both patent holders and it validates the sequence postulated in Section 2.3.

Lemma 8. *If the two patent holders have a portfolio of the same strength, the downstream* *producer always prefers to litigate first the one that has set the highest royalty.*

The intuition of this result is as follows. The higher the royalty rate of a patent holder the more likely it is that litigation pays off irrespective of the outcome of the litigation with the other patent holder. So, while the downstream monopolist may want to litigate against the high royalty patent holder, it may not want to litigate against the patent holder charging a low rate unless the high royalty rate patent portfolio is invalidated. In that case the optimal strategy of the downstream monopolist is to litigate against the high rate patent holder first since that creates the option to litigate against the low royalty patent holder; an option that will be exercise if the first trial is successful.

The previous result is useful in order to anticipate the changes in the probability that patent holders are litigated as a result of variations in the royalty rate. In particular, we now explore conditions under which a symmetric equilibrium $r_1^* = r_2^* = r^*$ exists. Because Π_D is a convex function of the total royalty rate, we have that

$$\Pi_D(r^*) - \Pi_D(2r^*) \leq \Pi_D(0) - \Pi_D(r^*).$$

This implies that if it is profitable to litigate one of the patent holders it will also be profitable to litigate the other one upon winning in court. This also means that in a symmetric equilibrium, r^*, litigation against both firms will take place if

$$(1 - g(x)) \left[\Pi_D(r^*) - \Pi_D(2r^*)\right] + (1 - g(x)) \left\{(1 - g(x)) \left[\Pi_D(0) - \Pi_D(2r^*)\right] - L_D\right\} > L_D.$$

This expression has the same interpretation as (8). Notice that because the downstream producer is indifferent between litigating any patent holder first, if this condition does not hold the probability that each one eventually faces a trial is $\frac{1}{2} + \frac{1}{2}(1 - g(x))$.

We now characterize the incentives to litigate when patent holder 1 deviates from the symmetric candidate equilibrium. Given r_1 and r_2 and the endogenous ordering that they imply we can define the gains arising from litigation contingent on success in the first trial as

$$\Phi(r_1, r_2) \equiv \begin{cases} \Pi_D(r_2) - \Pi_D(r_1 + r_2) + (1 - g(x)) \left[\Pi_D(0) - \Pi_D(r_2)\right] & \text{if } r_1 > r_2, \\ \Pi_D(r) - \Pi_D(2r) + (1 - g(x)) \left[\Pi_D(0) - \Pi_D(r)\right] & \text{if } r_1 = r_2 = r, \\ \Pi_D(r_1) - \Pi_D(r_1 + r_2) + (1 - g(x)) \left[\Pi_D(0) - \Pi_D(r_1)\right] & \text{otherwise.} \end{cases}$$

The first two terms correspond to the increase in profits accruing after the initial trial, whereas the last term is the additional increase in profits due to further litigation. Thus, litigation is profitable if $(1 - g(x)) \left[\Phi(r_1, r_2) - L_D\right] > L_D$. From Lemma 8, we know that if $r_1 > r_2$ the downstream producer litigates against patentee 1 first. The gains compared to the initial situation $\Pi_D(r_1 + r_2)$ accrue with probability $1 - g(x)$. Further litigation occurs in that case. Success against patentee 2, with probability $1 - g(x)$, results in profits $\Pi_D(0)$. If the downstream producer is defeated in court profits become $\Pi_D(r_2)$. The expression for profits is reversed when $r_2 > r_1$.

Consider how these profits change with r_1. In that case,

$$\frac{\partial \Phi}{\partial r_1} = \begin{cases} -\Pi'_D(r_1 + r_2) & \text{if } r_1 \geq r_2, \\ \Pi'_D(r_1) - \Pi'_D(r_1 + r_2) - (1 - g(x))\Pi'_D(r_1) & \text{otherwise.} \end{cases}$$

This implies that increases and decreases of r_1 around r_2 have a different effect on the willingness of the downstream producer to litigate. Consider an initial situation in which $r_1 = r_2$. As expected, an increase in r_1 raises the profitability of challenging the portfolio of patentee 1 as the downstream profits without litigation are smaller. Decreases in r_1 below r_2, however, lead to two opposing effects as shown in the previous expression. On the one hand, the first two terms correspond to the Inverse Cournot effect which implies that patent holder 2 is more likely to be litigated. As a result, a litigation cascade might ensue. On the other hand, contingent on the portfolio of patent holder 2 being invalidated, a lower r_1 reduces the expected gains from trying to invalidate the portfolio of patent holder 1 by $(1 - g(x))\Pi'_D(r_1)$. Hence, the total effect of a decrease in r_1 in the chances that patentee 1 is litigated is in general ambiguous.

Example 2. *Under a linear demand function, $D(p) = 1-p$, and symmetric royalty rates, the Inverse Cournot effect dominates the litigation cascade and, hence, a decrease in the royalty rate lowers the return from litigation of the downstream producer if and only if $r > \frac{1-g(x)}{2-g(x)}$. Notice that the unconstrained equilibrium royalty rate is $r_1^u = r_2^u = \frac{1}{3}$.*

As opposed to the case of one firm being threatened by litigation, the fact that a patent holder that chooses a lower royalty might be more likely to face litigation implies that sometimes a symmetric Nash Equilibrium in pure strategies may exist. The next proposition characterizes one such case.

Proposition 9. *With identical patent holders and a linear demand function, in a symmetric equilibrium in pure strategies, $r_1^* = r_2^* = r^*$, either $r^* = r^u$ or $r^* < r^u$ and it is defined as*

$$g(x)\Pi_D(r^*) + (1-g(x))\Pi_D(0) - \Pi_D(2r^*) = \frac{L_D}{1-g(x)} + L_D.$$

This last equilibrium arises when $g(x)$ and L_D are sufficiently small and $L_U \geq 0$. The equilibrium royalty is increasing in $g(x)$ and L_D.

In order to discuss the previous result it is useful to consider the possible deviations of any patent holder. First, only large increases in the royalty rate might compensate the sure litigation cost L_U and the probability that the patent portfolio is invalidated. When $g(x)$ is small the costs are likely to outweigh the benefits. Second, lowering the royalty rate below r^* implies that the other patentee is litigated first. However, given that $g(x)$ is small, a litigation cascade might affect the deviating patent holder, making the move less profitable. Finally, as discussed in the benchmark case, a significant decrease in the royalty rate is necessary in order to prevent litigation if the downstream producer is successful against patent holder 2. The lower is L_D the lower this royalty rate must be and, again, the less profitable this deviation becomes.

4 Robustness and Extensions

The results in this paper are based on a very stylized model of patent licensing. In this section we study the effect of changing some of the maintained assumptions throughout the paper.

4.1 Ad-Valorem Royalties

Although most of the literature in innovation has assumed that royalties are paid per-unit sold in the downstream market, in many technological industries patents are typically licensed using ad-valorem royalties, understood as a percentage of the revenue of the licensee.[18] As Llobet and Padilla (2016) show, absent litigation, royalty stacking also exists in the case of ad-valorem royalties although the impact is much reduced.

The moderating force introduced by the Inverse-Cournot effect also exists under ad-valorem royalties. In particular, consider the generic case in which the downstream producer faces a demand $D(p)$ and ad-valorem royalties s_1 and s_2 and it incurs in a marginal cost of production $c > 0$. If the total royalty is $S \equiv s_1 + s_2$, the problem of this producer can be written as

$$\Pi_D(S) = \max_p \left[(1 - S)p - c\right] D(p).$$

The monopoly price, p^M, is increasing in S under standard regularity conditions, such as the log-concavity of the demand function. This requirement is also enough to show that $\Pi_D(S)$ is decreasing and convex in S and it allows us to show that a counterpart of Lemma 7 holds in this case. That is, if $g(x_2) < g(x_1) = 1$, the downstream producer will litigate patent holder 2 if s_1 is lower than a threshold level $\bar{s}_1$. As a result, patent holder 1 will have an incentive to lower s_1 in order to constrain s_2 down – i.e. the Inverse-Cournot effect will

[18] See, for example, Bousquet et al. (1998). Interestingly, lump-sum payments are not common. Of course, if firms relied only on them the royalty-stacking problem would not be a relevant concern, and they would only have implications for the distribution of surplus.

diminish the royalty stack.

4.2 Downstream Competition

Introducing downstream competition is likely to moderate the Inverse-Cournot Effect and, in that case, the royalty stacking concern might reemerge. This result is due to two reasons. First, a free-riding problem arises. If courts invalidate the portfolio of one of the patent holders the royalty rate that all downstream producers pay is also reduced to 0. This means that the firm that challenges the portfolio does not benefit from a cost advantage with respect to its competitors but it must incur in the corresponding legal expenditures. As a result, the relevance of the litigation constraint underlying the Inverse-Cournot effect is reduced. Second, increased competition decreases downstream profits of any producer, both when the portfolio of a patent holder is invalidated and when it is not, making equation (12) more likely to hold.

In the extreme, if the downstream market were perfectly competitive, the threat of litigation would become irrelevant. However, in most technological markets product differentiation is typically important, mitigating the second force. Furthermore, the problem associated to free-riding is also likely to have a limited impact if firms are sufficiently heterogeneous in size. In that case, the profits from going to court will exceed its costs only for a subset of firms (usually the large ones), meaning that those firms will internalize most of the benefits from invalidating the portfolio of an upstream developer.

4.3 Royalty Renegotiation

The timing of the model assumes that once patent holder i chooses the royalty rate r_i, the downstream producer will end up paying either that amount or 0, if the patent portfolio is litigated and invalidated. In other words, if the validity of the portfolio is upheld in court the patent holder has no chance to increase the royalty rate. Nevertheless, allowing for this possibility, does not affect qualitatively the results of the paper. In the benchmark model, equation (4) sets the maximum royalty rate that patent holder 1 could charge and induce litigation on patent holder 2, denoted as $\bar{r}_1$. Allowing patent holder 2 to revise the royalty rate after the portfolio has been considered valid implies that this royalty would become $\tilde{r}_2 = 1 - \bar{r}_1$. Equation (4) would then be replaced by

$$(1 - g(x_2)) [\Pi_D(\bar{r}_1) - \Pi_D(\bar{r}_1 + r_2)] + g(x_2) [\Pi_D(1) - \Pi_D(\bar{r}_1 + r_2)] = L_D.$$

It would still be true that $\bar{r}_1$ increases in r_2 but only when the size of the portfolio of patent holder 2, x_2, is sufficiently small. So, the Inverse-Cournot effect will continue to operate when the distribution of patent ownership is sufficiently skewed.

4.4 Sequential Royalty Setting

In the benchmark model firms choose their royalty rates simultaneously. We consider now the case in which one patent holder has a large portfolio and the other has a small one and, thus, it is constrained by litigation. It can be verified that the equilibrium royalty rate is unchanged if we assume, instead, that patent holder 1 behaves like a Stackelberg leader and chooses first. The reason is that the royalty rate of patent holder 2 in both cases is set as a result of the action of the downstream producer, who moves after patent holder 1.

This equivalence is useful to explain the behavior of large innovators that participate in SSOs. These firms devote substantial resources in developing technologies that depend on the success in the final-good market of the products that embed them. The announcement of a low royalty rate early in the standardization process can, thus, be understood as a commitment that the royalty rate of complementary technologies developed by firms with a smaller patent portfolio would also be low, reducing the risk of royalty stacking. This interpretation is consistent with the adoption of some standards in recent years. For example, in the case of the fourth-generation mobile telecommunications technology (also denoted as

LTE) its main sponsors announced the licensing condition for their (essential) patents very early in the process.

The mechanism used to spur the adoption of a product that this paper uncovers achieves results similar to what we observe through contractual arrangements in other technological contexts. Gambardella and Hall (2006) study the public-good problem faced in the software development when placed in the public domain. Developers create improvements to the software but use them to launch commercial applications instead of making them available to the rest of the users. In this context they analyze the option that the leader of the project has to attach a General Public License (GPL) to the software, which forces all improvements to be contributed back to the project. They show that a GPL has two opposite effects. On the one hand, it discourages applications to be created since their developers lose the competitive advantage that their improvements generate as they become available to everybody. On the other hand, the quality of the software improves, since some developers that would otherwise create commercial applications now decide to contribute to the project and improve it. More recently, in the case of encryption technologies the risk that non-practicing entities might try to enforce their patents has encouraged agents more invested in the development of software to make it open source and, therefore, royalty free.[19]

In our model, a large patent holder can use the litigation threat of the downstream producer, by means of the Inverse-Cournot Effect, to limit the incentives of other patent holders to charge a high royalty rate, fostering the adoption of the technology. In these other examples, the choice of a GPL or a royalty-free arrangement allows the leaders of a technology to internalize part of the distortions generated. Imposing restrictions on the behavior of other developers reduces the free-riding problem, promotes the contribution to a technology and helps in its take-up. Leaders might even find optimal to forgo royalty

[19]See "A rush to patent the blockchain is a sign of the technologys promise" (2017, 14 January), The Economist (downloaded on 8 February 2017).

revenues altogether if they obtain profits from other sources related to the take-up of the technology, for example, through complementary services.

4.5 FRAND Licensing

Most SSOs request participating firms to license the patents that are considered essential to the standard according to Fair, Reasonable, and Non-discriminatory (FRAND) terms. The ambiguity of this term and the different interpretation of patent holders and licensees has made FRAND a legally contentious issue. Courts have sometimes been asked to decide whether a royalty rate is FRAND or not and in some instances to determine the FRAND rate.

The goal of this section is not to assert whether a royalty is FRAND or not but, rather, to study what is the effect of courts determining it on the previous results and, in particular, on the Inverse Cournot effect. In order to do so, we now extend the basic model and assume that the downstream producer can litigate a patent holder arguing, as before, that the portfolio is invalid and, in case it is not, to ask the court to rule that the patents are essential to the standard and the royalty requested is not FRAND. We assume that the larger is a patent portfolio the more likely it is that the technology it covers is considered essential to the standard. This probability is defined as $h(x_i)$, increasing in x_i. We also generalize the previous setup by considering the case of N firms, where R_{-i} corresponds to the sum of the royalty rate of all patentees other than i.

If the portfolio is declared to include patents that are essential to the standard the court will determine the appropriate royalty. We assume that this royalty, $\rho(x_i, r_i, R_{-i})$, is an increasing function of the quality of the patent portfolio, x_i. As we discuss later, we also allow for the possibility that the court's decision depends on the royalty announced by the patent holder or the total royalty established by the other patent holders.

Following the analysis in the benchmark model, the downstream monopolist will be interested in litigating patentee i only if

$$\begin{aligned}&(1-g(x_i))\left[\Pi_D(R_{-i})-\Pi_D(R_{-i}+r_i)\right]\\&+g(x_i)h(x_i)\left[\Pi_D(R_{-i}+\rho(x_i,r_i,R_{-i}))-\Pi_D(R_{-i}+r_i)\right]>L_D.\end{aligned}$$

The previous expression has a straightforward interpretation. The producer might benefit from litigation either because the patents are invalidated, which occurs with probability $1-g(x_i)$, or because they are considered valid and essential to the standard, with probability $g(x_i)h(x_i)$. In this latter case, the royalty rate is decreased from r_i to $\rho(x_i, r_i, R_{-i})$.

Lemma 10. *Suppose that $\rho(x_i, r_i, R_{-i})$ is independent of r_i and R_{-i}. Then, there exists a unique critical value $\bar{r}_i(x_i, R_{-i}, L_D)$ such that the producer prefers to litigate patentee i if and only if $r_i > \bar{r}_i$. Furthermore, this threshold is increasing in R_{-i} and L_D.*

This result indicates that the Inverse Cournot effect is qualitatively unaffected as long as the court determines the FRAND royalty as only a function of the quality of the portfolio. The main difference, however, is that the result does not guarantee that patent holders with a stronger portfolio can indeed charge a higher royalty without enticing the producer to litigate. The reason is that although a higher x_i reduces the probability that the court invalidates the patent portfolio, it also increases the probability that it considers the patents essential and, thus, that the royalty would be decreased from r_i to $\rho(x_i, r_i, R_{-i})$. This second effect dominates when increases in x_i have a large impact on $h(x_i)$ but a small one on $\rho(x_i, r_i, R_{-i})$.[20]

It is plausible, however, that $\rho(x_i, r_i, R_{-i})$ is increasing in r_i. Our results establish sufficient conditions and they might still hold even if ρ increases in r_i. An interesting case that

[20] It stands to reason that if the latter effect dominated, large patent holders would anticipate it and decide to license some of their patents at a rate of 0 in order to prevent their portfolio being deemed as essential.

it is worth to mention is the following: Suppose that a court would determine the FRAND royalty as a function of x_i but it will never choose $\rho(x_i, r_i, R_{-i})$ higher than r_i. It can be shown that the results are preserved in this case.

Finally, there have been instances in which courts have used existing royalties in order to pin down the FRAND royalty rate for a patent portfolio. Interestingly, they have been used in two directions. In some cases, courts have adopted the so-called *comparables approach* and set the royalty rate according to the rate negotiated for comparable patent portfolios, even in the same standard.[21] In those cases increases in R_{-i} would have a positive effect on $\rho(x_i, r_i, R_{-i})$ and strengthen the Inverse Cournot effect.

In other cases, and more concretely in the Microsoft v. Motorola case,[22] it has been argued that the FRAND royalty rate of a patent holder should be lowered due to the already large royalty stack. This reasoning would make $\rho(x_i, r_i, R_{-i})$ non-increasing in R_{-i}. Interestingly, this result would undermine the Inverse Cournot effect and, it might even have the effect of reversing its sign, with self-defeating consequences. Large patent holders would anticipate that by choosing a larger royalty, weaker competitors facing litigation would be forced by the court to set a lower rate, making worse the royalty-stacking problem that courts were trying to mitigate in the first place.

5 Concluding Remarks and Policy Implications

The existence of royalty stacking has been argued by translating the insights that arise from the idea of Cournot complements to the context of technology licensing. This paper shows, however, that these insights do not carry through when we explicitly consider patent litigation and, most specifically, the incentives that firms have to make strategic use of it.

[21]See Leonard and Lopez (2014) for a discussion of this and other approaches used to determine FRAND royalty rates.

[22]Microsoft Corp v. Morotola Inc, 854 F. Supp 2d 933 - Dist Court WD Washington 2012.

The implications of reconsidering the idea of royalty stacking through the lens of a model of patent litigation are far-reaching. One of the main contexts in which these changes apply is in the case of SSOs. Royalty stacking has been used to assess the desirability of patent consolidation or disaggregation. The concern about privateers, spin-offs of existing firms aimed at enforcing their intellectual property, and patent assertion entities has been seen as a way to increase the royalty stack. In contrast, consolidation efforts through patent acquisitions or the creation of patent pools have been encouraged as they contribute to lower the aggregate royalty rate.

In this model, as a result of the assumption that enforcement depends on the strength of the patent portfolio, if firms pool their patents they are likely to make enforcement more effective. As we discuss next, this last effect might imply that, contrary to common wisdom, the formation of a patent pool or the merger of two patent holders might make the royalty-stacking problem worse. By the same token, to the extent that disaggregation creates more asymmetric patent holdings, it might be socially beneficial.

To illustrate this point take the following simple example. Consider the case in which originally $N = 3$ patent holders decide independently on their royalty rate, with $g(x_1) = 1$ and $g(x_2) = g(x_3) \leq 1$. For simplicity, assume also that $g(x_2 + x_3) = 1$ so that the sum of their portfolios is big enough to guarantee their sure success in court if their patents are consolidated.

Extending the results in the benchmark model, patentees 2 and 3 are more likely to be restricted when L_D is small, leading to a lower royalty rate r_2 and r_3. We also know that when these rates are sufficiently low, the royalty stacking problem is likely to disappear, as patent holder 1 internalizes all the aggregate gains from a moderate r_1.

Consider now the decision of two patent holders to consolidate their portfolios in a patent pool. If this decision involves patentee 1, royalty stacking is less likely to arise. This observation is due to two reasons. First, by assumption, the strength of the resulting portfolio does not increase and, therefore, the bargaining power of the downstream producer against the pool is not affected. Second, suppose that the consolidation eliminates patentee 2 as a player. Because the Cournot effect implies that the merged firm will choose a royalty rate lower than $r_1 + r_2$, we have that patentee 3 will be more constrained by the threat of litigation and, due to the Inverse Cournot effect, it will need to decrease r_3. As a result of both effects, the large patent holder is likely to internalize a larger proportion of the surplus and, thus, moderate the royalty demands to prevent royalty stacking from emerging. It is important to notice that this consolidation is likely to be profitable for the parties involved precisely because the lower total royalty rate increases total surplus.

The previous positive effects are in opposition to what we find if patent holder 2 and 3 consolidate their portfolios and form a patent pool. Due to our assumptions, this new situation is akin to having two large patent holders and, as we discussed in the main part of the paper, in this situation royalty stacking is more likely to occur. In particular, if L_D is small the total royalty was low before consolidation but, as a result of it, the decrease in the number of patent holders leads to royalty stacking.

The application of these arguments to the opposite phenomenon, the creation of patent spin-offs, suggests that the welfare implications depend on their size. If they control large patent portfolios and decrease the skewness of the patent distribution they could increase the risk of a litigation cascade. As a result, they would discourage large patent holders from choosing a lower royalty rate, with a detrimental effect on welfare.

This paper also has implications for the incentives of downstream producers to merge with upstream patent holders. As usual, the vertical integration of a large patent holder with a downstream producer would mitigate the double-marginalization problem by eliminating the royalty payment. More interestingly, this change, by reducing the aggregate royalty rate, would strengthen the Inverse Cournot effect and, as a result, it would also reduce the royalty demands of other small patent holders.

Finally, our setup has intriguing implications for the incentives to innovate. Accounting for the Inverse Cournot effect implies that the returns from further innovation will be lower for weak patent holders and increase fast with the strength of the portfolio. As a consequence, innovation is likely to be more intense for already larger patent holders, leading over time to more concentrated patent ownership, which may result to a higher or lower royalty stack depending on the resulting number of patent holders.

References

Bourreau, Marc, Yann Meniere and Tim Pohlman, "The Market for Standard Essential Patents," 2015, working Paper.

Bousquet, Alain, Helmuth Cremer, Marc Ivaldi and Michel Wolkowicz, "Risk sharing in licensing," *International Journal of Industrial Organization*, September 1998, *16*(5), pp. 535–554.

Boutin, Aleksandra, "Screening for Good Patent Pools through Price Caps on Individual Licenses," *American Economic Journal: Microeconomics*, 2015, *8*(3), pp. 64–94.

Choi, Jay Pil and Heiko Gerlach, "A Model of Patent Trolls," 2015*a*, working Paper.

—, "Patent Pools, Litigation and Innovation," *RAND Journal of Economics*, 2015*b*, *46*(3).

—, "A Theory of Patent Portfolios," *American Economic Journal: Microeconomics*, February 2017, *9*(1), pp. 315–351.

Cournot, Antoine Augustin, *Researches into the Mathematical Principles of the Theory of Wealth*, Macmillan, 1838, Nathaniel T. Bacon (trans.), 1987.

Farrell, Joseph and Carl Shapiro, "How Strong Are Weak Patents," *American Economic Review*, 2008, *98*(4), pp. 1347–1369.

Galetovic, Alexander, Stephen Haber and Ross Levine, "An Empirical Evaluation of Patent Holdup," *Journal of Competition Law and Economics*, 2015, *11*, pp. 549–578.

Gambardella, Alfonso and Bronwyn H. Hall, "Proprietary versus Public Domain Licensing of Software and Research Products," *Research Policy*, 2006, *35*(6), pp. 875–892.

Lemley, Mark A. and A. D. Melamed, "Missing the Forest for the Trolls," *Columbia Law Review*, 2013, *113*, pp. 2117–2189.

Lemley, Mark A. and Carl Shapiro, "Patent Holdup and Royalty Stacking," *Texas Law Review*, 2007, *85*, pp. 1991–2049.

Leonard, Gregory K. and Mario A. Lopez, "Determining RAND Royalty Rates for Standard-Essential Patents," *Antitrust*, 2014, *29*(1), pp. 86–94.

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

—, "Standard-Essential Patents," *Journal of Political Economy*, 2015, *123*(3), pp. 547–586.

Llobet, Gerard, "Patent litigation when innovation is cumulative," *International Journal of Industrial Organization*, October 2003, *21*(8), pp. 1135–1157.

Llobet, Gerard and Jorge Padilla, "The Optimal Scope of the Royalty Base in Patent Licensing," *Journal of Law and Economics*, 2016, *59*(1), pp. 45–73.

Rey, Patrick and Jean Tirole, "Cooperation vs. Collusion: How Essentiality Shapes Co-opetition," 2013, TSE Working Paper 439.

Spulber, Daniel F., "Complementary Monopolies and Bargaining," 2016, Northwestern Law & Econ Research Paper No. 16-10.

A Proofs

The main results of the paper are proved here.

Proof of Lemma 1: Immediate from the fact that when $R = v$ total profits are v whereas under $R = 1$ profits are α. $\square$

Proof of Proposition 2: Regarding the first case, contingent on selling with probability 1 the sum of royalties must be equal to v or otherwise any patent holder would deviate and increase the royalty rate. Hence, take r_1^u and $r_2^u = v - r_1^u$ and suppose without loss of generality that $r_1^u \geq \frac{v}{2} \geq r_2^u$. The optimal deviation for patentee i is $\hat{r}_i = 1 - r_j^u$ for $j \neq i$ and it would be unprofitable if $v - r_j^u \geq \alpha(1 - r_j^u)$ or $r_j^u \leq \frac{v-\alpha}{1-\alpha}$. Such a combination of royalties is only possible as long as $\frac{v}{2} \leq r_1^u \leq \frac{v-\alpha}{1-\alpha}$ or $v \geq \frac{2\alpha}{1+\alpha}$.

For the second case, take r_1^u and $r_2^u = 1 - r_1^u$ and suppose without loss of generality that $r_1^u \geq \frac{1}{2} \geq r_2^u$. The optimal deviation for patentee i is $\hat{r}_i = v - r_j^u$ for $j \neq i$ if it leads to a positive royalty and it would be unprofitable if $\alpha(1 - r_j^u) \geq v - r_j^u$ or $r_j^u \geq \frac{v-\alpha}{1-\alpha}$. Such a combination of royalties will be possible as long as $\frac{v-\alpha}{1-\alpha} \leq r_2^u \leq \frac{1}{2}$ or $v \leq \frac{1+\alpha}{2}$.

Finally, notice that $\frac{2\alpha}{1+\alpha} < \frac{1+\alpha}{2}$ for all $\alpha \in [0, 1]$ so both equilibria can co-exist. $\square$

Proof of Proposition 4: Assume towards a contradiction that $\frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$ and there exists an equilibrium with $r_1^* + r_2^* = 1$. Notice that $\frac{v-\alpha}{1-\alpha} \leq v$ implies $\frac{L_D}{1-g(x_2)} < v$, and, from Proposition 2, an equilibrium with royalty stacking requires $r_2^* \geq \frac{v-\alpha}{1-\alpha}$.

A necessary condition for an equilibrium with royalty stacking to exist is $r_2^* \leq \bar{r}_2$. Also notice that since $\underline{r}_2 = \frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$, $r_2^* > \underline{r}_2$.

Thus, assume that $\underline{r}_2 < r_2^* \leq \bar{r}_2$ so that patentee 2 is litigated if patentee 1 deviates and chooses $\hat{r}_1 = \bar{r}_1(r_2^*)$. Patentee 1's profits become

$$\hat{\Pi}_1(x_2, r_2^*) = [\alpha + (1-\alpha)(1-g(x_2))] \left[v + \frac{\alpha}{1-\alpha} r_2^* - \frac{L_D}{(1-\alpha)(1-g(x_2))} \right],$$

strictly decreasing in x_2 and L_D and strictly increasing in r_2^*. This deviation will be unprofitable if $\hat{\Pi}_1(x_2, r_2^*) \leq \alpha r_1^* = \alpha(1 - r_2^*)$ or, rearranging terms, if

$$r_2^* \leq \tilde{r}_2(x_2, L_D) = \frac{(1-\alpha)(\alpha - Gv) + G\frac{L_D}{1-g(x_2)}}{\alpha(G + (1-\alpha))}$$

where $G \equiv \alpha + (1-\alpha)(1-g(x_2)) \in [\alpha, 1]$. This expression can be rewritten as

$$r_2^* \leq \left[1 - \frac{G}{\alpha(G+(1-\alpha))}\right] \frac{vG-\alpha}{G-\alpha} + \frac{G}{\alpha(G+(1-\alpha))} \frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$$

since $\frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$ and $\frac{vG-\alpha}{G-\alpha}$ is increasing in G for $v \leq 1$. Thus, we reach a contradiction.

Hence, we only need to show that the equilibrium with $R = v$ exists. Consider the case $r_2^* = \frac{L_D}{1-g(x_2)} < \frac{v-\alpha}{1-\alpha}$ and $r_1^* = v - r_2^*$. From (4) r_2^* avoids litigation and by Proposition 2 patentee 1 has no incentive to deviate. Thus, the only deviation we need to consider from patentee 2 is such that $R > v$. However, notice that

$$r_1^* = v - \frac{L_D}{1-g(x_2)} = v + \frac{\alpha}{1-\alpha} r_2^* - \frac{L_D}{(1-\alpha)(1-g(x_2))} = \bar{r}_1(r_2^*),$$

and so any higher r_2 will induce litigation. Hence, an equilibrium in pure strategies exists if and only if such a deviation is not profitable

$$\frac{L_D}{1-g(x_2)} \geq \alpha g(x_2) \left(1 - v + \frac{L_D}{1-g(x_2)}\right) - L_U.$$

This condition is guaranteed if $g(x_2)$ is sufficiently small or L_U sufficiently large. □

Proof of Lemma 5: First notice that if patent holder 2 loses in court patent holder 1 will be litigated if and only if

$$\Pi_D(0) - \Pi_D(\hat{r}_1) > \frac{L_D}{1-g(x)}$$

or $\hat{r}_1 > \frac{L_D}{1-g(x)}$. Also notice that, from the arguments in the text, if originally it was not optimal to engage in litigation it has to be that

$$\Pi_D(1/2) - \Pi_D(1) \leq \frac{L_D}{1-g(x)}.$$

Patent holder 1 would be litigated after downstream producer loses against patent holder 2 if

$$\Pi_D(1/2) - \Pi_D(1/2 + \hat{r}_1) > \frac{L_D}{1-g(x)}$$

which is incompatible with the previous condition. □

Proof of Proposition 6: Define $F(R) \equiv D(p^M(R))$ so that $D(p^M(R))$ is quasiconcave if $F'(R)^2 \geq F''(R)F(R)$. The optimal royalty of patentee i is the result of

$$\max_{r_i} \; r_i F(R),$$

with first-order condition

$$F(R) + r_i^* F'(R) = 0. \Longrightarrow r_i^* = -\frac{F(R)}{F'(R)}.$$

Replacing $r_i^* = r^* = \frac{R^*}{N}$ we can use the Implicit Function Theorem to compute

$$\frac{dR^*}{dN} = \frac{\frac{R^*}{N} F'(R^*)}{F'(R^*) + \frac{R^*}{N} F'(R^*) + \frac{1}{N} F'(R^*)} \geq 0.$$

The last inequality arises from a negative numerator due to $F'(R) \leq 0$ and a negative denominator that it is also negative due to the quasiconcavity of $F(R)$. □

Proof of Lemma 7: From equation (4) we can see, using the fact that $\Pi_D'(R) < 0$ and $\Pi_D''(R) > 0$, that

$$\frac{d\bar{r}_1}{dL_D} = \frac{1}{\Pi_D'(\bar{r}_1) - \Pi_D'(\bar{r}_1 + r_2)} < 0,$$
$$\frac{d\bar{r}_1}{dx_2} = \frac{g'(x_2)\left[\Pi_D(\bar{r}_1) - \Pi_D(\bar{r}_1 + r_2)\right]}{\left[\Pi_D'(\bar{r}_1) - \Pi_D'(\bar{r}_1 + r_2)\right]} < 0,$$
$$\frac{d\bar{r}_1}{dr_2} = \frac{\Pi_D'(\bar{r}_1 + r_2)}{\Pi_D'(\bar{r}_1) - \Pi_D'(\bar{r}_1 + r_2)} > 0.$$

□

Proof of Lemma 8: Suppose without loss of generality that $r_1 > r_2$. The optimal policy of the downstream producer can be described as arising from the following two stages. In the first stage, it decides whether to litigate patentee 1 or 2 or none at all. Upon observing the outcome of the first trial the patent holder must decide whether to litigate the other patent holder or not.

Suppose that in the first stage patentee i was litigated. Then, if it is optimal for the downstream producer to litigate patentee j upon the defeat it is also optimal to litigate upon victory since, by convexity of $\Pi_D(R)$,

$$\Pi_D(r_i) - \Pi_D(r_i + r_j) \leq \Pi_D(0) - \Pi_D(r_j),$$

for $i = 1, 2$ and $j \neq i$. Furthermore, notice that

$$\Pi_D(r_1) - \Pi_D(r_1 + r_2) \leq \Pi_D(r_2) - \Pi_D(r_1 + r_2),$$
$$\Pi_D(0) - \Pi_D(r_2) \leq \Pi_D(0) - \Pi_D(r_1).$$

Hence, two possible orderings can arise depending on whether $\Pi_D(r_2) - \Pi_D(r_1 + r_2)$ is higher or lower than $\Pi_D(0) - \Pi_D(r_2)$. In order to determine the profits of the downstream producer in each case, we need to see how these profits compare with $\Lambda \equiv \frac{L_D}{1-g(x)}$.

i Suppose that when 1 is litigated first it is always optimal to litigate 2 afterwards. Obviously, if litigating 1 after the litigation of 2 is also optimal, both options are equivalent and profits are identical.

ii Suppose that when 1 is litigated first it is only optimal to litigate 2 after victory. This implies that $\Pi_D(r_1) - \Pi_D(r_1 + r_2) < \Lambda \leq \Pi_D(0) - \Pi_D(r_2)$. Profits are

$$g(x)\left[\Pi_D(r_1 + r_2) - L_D\right] + (1 - g(x))\left[g(x)\Pi_D(r_2) + (1 - g(x))\Pi_D(0)\right] - L_D.$$

These profits are, by definition, higher than those that arise in the first case. If after litigation of patent holder 2 it is then optimal to litigate firm 1 always, this option would be, therefore, dominated by (i).

Alternatively, it could be that when 2 is litigated first it is only optimal to litigate 1 upon victory. Profits would be in that case,

$$g(x)\left[\Pi_D(r_1 + r_2) - L_D\right] + (1 - g(x))\left[g(x)\Pi_D(r_1) + (1 - g(x))\Pi_D(0)\right] - L_D,$$

which are lower than when 1 is litigated first.

iii Suppose that when 1 is litigated first it is never optimal to litigate 2 afterwards. This implies profits are

$$g(x)\Pi_D(r_1 + r_2) + (1 - g(x))\Pi_D(r_2) - L_D.$$

If when 2 is litigated first, it is optimal to litigate 1 always, these profits are lower because, as in the previous case, they coincide with profits in the first option. If instead it was optimal to litigate only upon success, again, these profits are dominated by the second option as seen before. Finally, if it is never optimal to litigate firm 1, profits are

$$g(x)\Pi_D(r_1 + r_2) + (1 - g(x))\Pi_D(r_1) - L_D,$$

which are again lower.

iv Using the same argument, if Λ is sufficiently high so that it is never optimal to litigate 1 only, litigating 2 only must also be dominated.

$\square$

Proof of Proposition 9: Consider a symmetric equilibrium in which 1 are 2 constrained. This implies that $\Phi(r^*, r^*) = \frac{L_D}{1-g(x)} + L_D$. Profits are $r^* D(p^M(2r^*))$. It is immediate that r^* is increasing in L_D and $g(x)$.

Three possible deviations of any patent holder, say patentee 1, can come about:

i Patentee 1 might increase its royalty to $r_1 > r^*$. In that case, Patentee 1 will be litigated first. Profits become $\max_{r_1}\ g(x) r_1 D(p^M(r_1 + r^*)) - L_U$.

ii Patentee 1 might deviate by lowering the royalty slightly. In this case, the sign of $\frac{\partial \Phi}{\partial r_1}$ becomes relevant. In particular,

$$\frac{\partial \Phi}{\partial r_1}(r_1, r_2) \geq 0 \Longleftrightarrow g(x)\Pi'_D(r_1) - \Pi'_D(r_1 + r_2) = D(p^M(r_1 + r_2)) - g(x) D(p^M(r_1)) \geq 0,$$

If $\frac{\partial \Phi}{\partial r_1} \geq 0$, decreases in r_1 reduce the incentives for the downstream firm to litigate. Since royalties are strategic substitutes and r^* is below the unconstrained royalty this strategy can never be optimal.

Alternatively, if $\frac{\partial \Phi}{\partial r_1} < 0$, a deviation consisting in a slight decrease in r_1 induces litigation, first against patentee 2 and, upon success, against patentee 1. This implies that the profits of patentee 1 become

$$g(x) r^* D(p^M(2r^*)) + (1 - g(x)) \left[g(x) r^* D(p^M(r^*)) - L_U \right],$$

This deviation is unprofitable if

$$r^* D(p^M(2r^*)) - g(x) r^* D(p^M(r^*)) < -L_U,$$

which holds if L_U is sufficiently large, since the left-hand side is negative when $\frac{\partial \Phi}{\partial r_1}(r^*, r^*) < 0$ which occurs when r^* is large.

iii Finally, patent holder 1 could lower r_1 enough so that $(1 - g(x)) [\Pi_D(0) - \Pi_D(r_1)] \leq L_D$. In that case, patent holder 1 would not be litigated. Again, two possibilities can arise here depending on whether the downstream producer is interested in litigating patentee

2 or not. Notice that only if patentee 2 is litigated this deviation might be profitable. Hence, the optimal deviation $\tilde{r}_1 = \min\{r_1^A, r_1^B\}$, where

$$(1 - g(x)) \left[\Pi_D(0) - \Pi_D(r_1^A)\right] = L_D, \tag{13}$$

and

$$(1 - g(x)) \left[\Pi_D(r_1^B) - \Pi_D(r^* + r_1^B)\right] = L_D. \tag{14}$$

When r^* is sufficiently high the first constraint will be binding. Profits in either case will be $g(x)r_1D(p^M(r^* + \tilde{r}_1)) + (1 - g(x))r_1D(p^M(\tilde{r}_1))$.

When $g(x)$ is sufficiently small it is clear that the first deviation is always dominated since it would imply profits of $-L_U$. The second deviation is also unprofitable since when $g(x) = 0$, $\frac{\partial \Phi}{\partial r_1} \geq 0$.

Regarding the last deviation, we know that $\tilde{r}_1 \leq r_1^B$. Under a linear demand when $g(x) = 0$, we have that $\Pi_D(0) - \Pi_D(2r^*) = 2\left[\Pi_D(r_1^B) - \Pi_D(r_1^B + r^*)\right]$ implies $r_1^B = \frac{r^*}{2}$. Thus, for the deviation not to be profitable we only require

$$r^*D(p^M(2r^*)) \geq \frac{r^*}{2}D\left(p^M\left(\frac{r^*}{2}\right)\right).$$

When L_D is 0, $r^* = 0$ and the result holds trivially. The derivative of the profit functions evaluated at $r^* = 0$ are $D(p^M(0))$ and $\frac{1}{2}D(p^M(0))$ for the left-hand side and the right-hand side expression, respectively. Thus, the deviation is not profitable when L_D is sufficiently small.

We now show that there is no other symmetric pure strategy equilibrium when the litigation constraint is relevant. First, notice that if $r_1 = r_2$ are lower than r^*, each firm has incentives to increase its royalty since their problem is the same as they would face if they were unconstrained and royalties are strategic substitutes. If, instead, $r_1 = r_2 = \tilde{r}$ are higher than r^* each firm obtains profits

$$\frac{1}{2}\left[g(x)\tilde{r}D(p^M(2\tilde{r})) - L_U\right] + \frac{1}{2}\left[g(x)\tilde{r}D(p^M(2\tilde{r})) + (1 - g(x))\left[g(x)\tilde{r}D(p^M(2\tilde{r})) - L_U\right]\right]$$

where each firm is litigated first with probability $\frac{1}{2}$ and the second firm is litigated only if the downstream producer succeeds against the first. Notice that in this case it is always optimal for one firm, say patentee 1, to undercut the other patentee. As a result profits increase to

$$g(x)\tilde{r}D(p^M(2\tilde{r})) + (1 - g(x))\left[g(x)\tilde{r}D(p^M(2\tilde{r})) - L_U,\right]$$

leading to higher profits. $\square$

Proof of Lemma 10: Define

$$\Phi(r_i, x_i, L_D, R_{-i}) \equiv (1 - g(x_i)) \left[\Pi_D(R_{-i}) - \Pi_D(R_{-i} + r_i)\right] +$$
$$g(x_i)h(x_i) \left[\Pi_D(R_{-i} + \rho(x_i, r_i, R_{-i})) - \Pi_D(R_{-i} + r_i)\right] - L_D$$

Obviously, $\frac{\partial \Phi}{\partial L_D} = -1$. We can also compute

$$\frac{\partial \Phi}{\partial r_i} = -(1 - g(x_i))\Pi'_D(R_{-i} + r_i) + g(x_i)h(x_i) \left[\Pi'_D(R_{-i} + \rho(x_i, r_i, R_{-i}))\frac{\partial \rho}{\partial r_i} - \Pi'_D(R_{-i} + r_i)\right]$$
$$\frac{\partial \Phi}{\partial R_{-i}} =(1 - g(x_i)) \left[\Pi'_D(R_{-i}) - \Pi'_D(R_{-i} + r_i)\right]$$
$$+ g(x_i)h(x_i) \left[\Pi'_D(R_{-i} + \rho(x_i, r_i, R_{-i})) \left(1 + \frac{\partial \rho}{\partial R_i}\right) - \Pi'_D(R_{-i} + r_i)\right]$$

Given that Π_D is convex, $\rho(x_i, r_i, R_{-i}) \leq r_i$ and the assumption that $\rho(x_i, r_i, R_{-i})$ is independent of r_i and R_{-i}) we can show that $\frac{\partial \Phi}{\partial r_i} \geq 0$ and $\frac{\partial \Phi}{\partial R_{-i}} \leq 0$. $\square$