Pay-for-Delay with Settlement Externalities*

Emil Palikot[†] and Matias Pietola[‡]

May 23, 2018

*We are grateful to Yassine Lefouili, Patrick Rey, Marc Ivaldi, Jorge Padilla, Bruno Jullien, Carl Shapiro, Martin Peitz, Doh-Shin Jeon and Massimo Motta for their valuable comments on various stages of the project. This paper also benefited from comments at seminars in Aalto University, Toulouse School of Economics as well as 15th Annual IIOC, 12th Annual CRESSE and 44th Annual Conference of EARIE.

†Toulouse School of Economics, www.emilpalikot.com, emil.palikot@gmail.com

‡Toulouse School of Economics, www.pietola.com, matiaspietola@gmail.com. Matias Pietola gratefully acknowledges the financial support from the Yrjö Jahnsson Foundation and the Finnish Cultural Foundation.

Abstract

This paper studies settlements between an incumbent patent holder and multiple potential entrants to the market in the shadow of patent litigation. We show that there exists litigation in equilibrium for patents of intermediate strength, whereas sufficiently weak or strong patents are not taken to court. The incumbent uses divide and conquer strategies by paying to delay some entrants, whereas the others either obtain a license for early entry (no litigation) or no settlement deal at all (litigation). We identify a new source of sequentiality in equilibrium entry to markets governed by patents. Settlement externalities between the entrants are the driving force behind our results: when one more entrant is delayed from entering the market, there is less competition and litigation threat from the other entrants is increased. Our results bring new important insights for the hotly debated topic of pay-for-delay agreements witnessed in the pharmaceutical industry and sanctioned by antitrust authorities both in the US and Europe.

1 Introduction

So-called pay-for-delay agreements in the pharmaceutical industry have caught the attention of antitrust authorities around the world. These are settlements of patent litigation between an incumbent patentee and an entrant who seeks to enter the market protected by the patent: In exchange for money from the incumbent, the entrant agrees not to challenge the validity of the patent and stays out of the market, for a certain period of time. Such deals fall on the borderline of antitrust policy, which detects and punishes cartels, and intellectual property rights, which restrict competition by definition. The European Commission considers them strictly anti-competitive and has imposed significant fines on the companies involved, most notably in the Servier case and Lundbeck case.[1] Meanwhile, the US Supreme Court adopted a rule of reason approach in the Actavis case.

Yet paying to delay entry is not the only way to settle a patent dispute. In the European Economic Area, depending on the year, pay-for-delay arrangements have constituted only around 3-12 % of all patent settlements made by pharmaceutical companies.[2] Licensing agreements, where the entrant pays the incumbent and obtains a license for early entry, are more frequent. Although both types of settlement deals often appear simultaneously with different entrants, little attention has been paid to them in this regard. The question why incumbent patentees offer different deals to similar entrants has not been addressed in the current economic literature, to our knowledge. This paper aims to fill the gap.

We develop a setting with one incumbent and multiple identical entrants. The incumbent owns an uncertain patent and enjoys a legal monopoly unless a court of law declares the patent invalid. Each entrant can litigate over patent validity, wait for the market to open or settle with the incumbent. A settlement deal includes a monetary transfer and an entry date for the entrant. We show that, under a credible litigation threat, the incumbent adopts a divide-and-conquer strategy, where she pays to delay some of the entrants, while the others either obtain a license or do not settle. Licensing and litigation are substitutes in reducing the cost of entry delay. We show that patents with intermediate strength are litigated, whereas sufficiently weak or strong patents are licensed. Furthermore, litigation is more likely to occur when litigation costs are low.

[1] These cases are currently on appeal before the EU Court of Justice.

[2] See the Pharma Inquiry by the European Commission. Since 2008, the European Commission has annually monitored patent settlements made by pharmaceutical companies in the EEA area.

In particular, when there is no cost of litigation, there is always litigation.

Our results strike a contrast with the predictions from models with a single entrant. Faced with only one entrant, the incumbent always pays the entrant to stay out, sharing the monopoly profit (Shapiro, 2003). However, with multiple entrants, such a deal imposes an externality to the other entrants. They face less competition in the event of patent invalidation, which increases the litigation threat. In particular, to monopolize the market, the incumbent must compensate each entrant with expected duopoly profits. For sufficiently many entrants, this cost exceeds the gain from monopolization.

We also provide an extension of our model, where the incumbent can offer settlements with entry dates contingent on patent validity. Such conditioning eliminates settlement externalities, reducing the cost of entry delay. If litigation costs are high, the incumbent concludes a pay-for-delay agreement with all entrants, but litigates with one of them if litigation costs are low. As conditional settlement terms lead to more entry delay, they effectively restrict competition in the market.

Related literature So far the literature on patent disputes has separately focused on two types of settlement deals:

- *Licensing:* Farrell and Shapiro (2008); Lemley and Shapiro (2005); Kamien and Tauman (1986); Katz and Shapiro (1987); Amir et al. (2014)

- *Pay-for-delay:* Shapiro (2003); Meunier and Padilla (2015); Elhauge and Krueger (2012); Edlin et al. (2015); Manganelli (2014); Gratz (2012).

However, to our knowledge, licensing and pay-for-delay agreements have not been captured in the same economic model before. Yet assuming that different agreements are made independently from each other obfuscates an important economic mechanism, triggered by settlement externalities between the entrants. We contribute to the economic theory of patent settlements by unifying the two approaches and showing that, due to settlement externalities, the incumbent may benefit from different settlement deals with similar entrants. This relates our work to the literature on contracting with externalities (Segal, 1999, 2003).

The canonical model on pay-for-delay is introduced by Shapiro (2003). He considers a setup with a single entrant who may challenge an incumbent patent holder in a court of law. The parties of the dispute have the opportunity to settle it instead of going to court. They can agree on a reverse payment and an entry date for the entrant. In this setting, the parties will agree on extending the monopoly period and divide the resulting profits. Our approach builds on this by allowing for multiple entrants. This reveals the connection between pay-for-delay agreements, licensing and litigation, and will enable us to capture new a source of sequentiality in market entry.

Previous work on multiple entry to markets protected by a probabilistic patent assumes only one entrant can credibly threaten the patent holder with litigation, while entrants with weaker incentives to challenge the patent holder will free ride on the litigation efforts taken by the more aggressive one (Meunier and Padilla, 2015). Instead of incorporating this economic logic to our model exogenously, we show that in equilibrium at most one entrant follows through with litigation. Others will free ride, even if all entrants are symmetric and have the ability to litigate. Thus we show that the assumption taken in some models is actually an outcome of a strategic interaction between the entrants.

Shapiro (2003) proposes a general rule for evaluating patent settlements in antitrust policy: allowing for settlements should not leave the consumers worse off compared to no settlement and hence litigation. Therefore, welfare analysis simplifies to comparing the duration of exclusion resulting from the settlement (entry date agreed) to the expected entry date when there is no settlement and entry may only occur through patent invalidation. When the reverse payment is higher than the litigation cost incurred by the incumbent, exclusion due to the settlement will exceed the expected delay from litigation (Shapiro, 2003). Elhauge and Krueger (2012) argue that all pay-for-delay settlements with a reverse payment higher than litigation costs should be per se illegal, independently from the probability of the patent being invalid.

Finally, divide-and-conquer strategies have previously been studied in different contexts. Posner et al. (2010) show how a defendant can optimally exploit coordination failures between several plaintiffs. Other related works include Daughety and Reinganum (2002) and Che and Spier (2008). Typically, in these papers, some plaintiffs are offered beneficial treatments and decide to settle with the defendant, which makes the others drop their lawsuits. Perhaps the most natural way to look at the incumbent's problem is that she wants to delay entry in order to benefit from reduced competition, this however becomes increasingly expensive when the number of entrants raises. In order to reduce cost of delaying, the incumbent will accommodate some entry through licensing or pursue litigation.

The paper is organized as follows. The next section is devoted to case studies in the pharmaceutical industry, which sheds light on the environment of patent settlements in practice. Section 3 introduces our baseline model for two entrants, which is then generalized in Section 4 to allow for more than two entrants. Section 5 considers different policy implications and extensions of our model. Section 6 concludes.

2 Pay-for-delay cases in the pharmaceutical industry

Pharmaceutical companies hold patents on branded drugs that they have invented, but face competitive pressure from generic producers of bio-equivalent medicines. Producers of generics can simultaneously contest the brand once the primary patent, protecting the main chemical compound, has expired. Since then, usually several generic producers race to the market neck-to-neck. However, often despite the expiry of the main patent, the legal situation is unclear: the brand has applied for a secondary patent protection and sometimes holds patents on alternative ways of manufacturing the medicine.

Decisions in three cases, Lundbeck, Servier and Actavis, have been referred to as land-marking. We briefly discuss some important elements of these cases.

> The Lundbeck case considers a number of agreements between a Danish pharmaceutical company Lundbeck and several generic drugs producers. In the 1970's and the 1980's Lundbeck developed antidepressant drug citalopram that started being marketed in the 1990's. The medicine was largely successful and constituted a major product for Lundbeck; for example 80-90% of the company's revenues in 2002. At the time of the settlements, years 2002 and 2003, patents related to the chemical compound and the original process had expired. In principle, the market was free for generic producers to enter. However, Lundbeck still had a number of patents related to more efficient or alternative ways of manufacturing the drug. Lundbeck implemented a so-called "generic strategy" that involved different kinds of agreements with several entrants. Overall, the generic strategy was a mixture of reverse payments, takeovers, licensing, accommodation and even introduction of own authorized generic producers. For example, in the UK it allowed one firm to enter the market but offered a reverse payment to another one. In Iceland, Lundbeck allowed a market entry without litigation.
>
> The Perindopril case involves a French pharmaceutical manufacturer Servier and generic producers of perindopril, a medicine for treating high blood pressure developed by Servier in the 1980's. Perindopril became Servier's most

successful product with annual global sales in years 2006 and 2007 exceeding USD 1 billion with average operating margins beyond 90%. Generic entry started to impose a credible threat to Servier once the patent governing the main compound expired in May 2003. Anticipating this, Servier had started to design and implement a generic strategy from the late 1990's. This strategy included acquiring new patents and resulted in five settlement agreements with different generic producers, between the years 2005 and 2007. Four of these agreements were considered pay-for-delay settlements, while the fifth one was a licensing deal. Servier considered litigation and licensing as alternative tools in its strategy.

In the US, the Drug Price Competition and Patent Term Restoration Act of 1984 (the so-called Hatch-Waxman Act) shapes the regulatory approval of generic drugs. The aim of this legislation is to promote entry of generics by guaranteeing the first one of them a duopolistic position. When a generic producer of a bio-equivalent drug files an Abbreviated New Drug Application (ANDA) to the Food and Drug Administration (FDA), the Hatch-Waxman Act requires declaring a relationship to a patent mentioned in the Orange Book, a list of approved drug products together with a catalog of patents related to each of them. If a generic producer states that the relevant patents are no longer valid, or that it is not infringing, certification is granted. The Hatch-Waxman Act provides 180 days of exclusivity for a generic producer who makes such a certification: no other generic producer can obtain approval from the FDA during this time.

The FTC vs. Actavis case was brought to the US Supreme Court by the Federal Trade Commission (FTC) in 2013. The case considers a deal made between a Belgian pharmaceutical company Solvay Pharmaceuticals and a generic producer Actavis, Inc. Solvay was granted a new patent for AndroGel in 2003. Later on, Actavis filed an ANDA to the FDA and stated that Solvay's new patent was invalid and the generic version produced by Actavis did not infringe upon the AndroGel patent. Solvay settled the case with Actavis: the settlement agreement included a reverse payment from Solvay to Actavis in return for an exclusion period during which Actavis agreed to stay out of the market. The agreed entry date was 65 months before the AndroGel patent expired. The FTC considered the arrangement between Solvay and Actavis as an antitrust violation and brought a lawsuit against them. The District Court and the appellate court, the Eleventh

Circuit, dismissed the case. However ,the US Supreme Court overturned their decisions and held that it is not sufficient to base the legal analysis on patent law policy and that the antitrust question must be addressed. The Supreme Court argued that: (1) FTC's complaint could not have been dismissed without analyzing the potential justifications for such decision; (2) the patentee is likely to have enough power to implement antitrust harm in practice; (3) the antitrust action is likely to prove more feasible administratively than the Eleventh Circuit believed: a large, unexplained payment from the patentee to the generic producer can provide a workable surrogate for a patent's weakness; (4) the parties could have made another type of settlement agreement, by allowing the generic producer to enter the market before the patent expires without a need for a reverse payment.

Key lessons: Even a weak patent can be useful for the incumbent brand, due to free-riding between the generic companies in their litigation efforts. Success in trial opens the market for everybody, not merely for the generic producer who took litigation effort and incurred, often a significant, litigation cost. The generic entrants have expressed their concern to "win the battle, but lose the war" due to follow-up entry to the market.[3] The incentives to settle litigation are pronounced when other generic producers lie in wait in the shadow of the litigation. The terms of a settlement have to reflect both the competitive situation in the market and the strength of the patent. In order to reach mutually beneficial settlement parties need to have a similar assessment of the strength of the patent. To ensure that these assessments reflect the actual probability of patent invalidation parties make laboratory tests and seek third-party counsels.[4]

We can make two other observations from the case studies. Firstly, the entry game starts after the expiry of a certain patent, and this date is common knowledge. Secondly, the entrants arrive at the market simultaneously, and their subsequent sequential entry is an outcome of an interplay between the patent holder and the entrants. Some entrants may receive a license or go to court, while the others are delayed. Such divide and conquer strategies are achieved through a myriad of decisions: pay-for-delay agreements, licensing deals, litigation, and take-overs.

[3]See paragraph 493 of the Perindoprildecision by the European Commission.
[4]See paragraph 709 of the Perindoprildecision and 522 of the Lundbeckdecision

3 Basic model

We consider a market with two symmetric entrants $\{1, 2\}$ and one incumbent. The incumbent owns a patent and enjoys a legal monopoly until the patent expires unless one of the entrants litigates and a court of law declares the patent invalid. An entrant who litigates incurs litigation cost $c \geq 0$, while the incumbent has to pay $C \geq 0$. It is common knowledge that, if at least one of the entrants litigates, the court declares the patent invalid with probability $1-\theta$ but upholds the patent otherwise, where θ captures patent strength. We assume the outcomes from litigation are perfectly correlated.[5]

The patent goes from date zero to date one when it expires. At date zero, the firms play a two-stage negotiation game. The incumbent moves first and offers each entrant $i \in \{1, 2\}$ a settlement deal, which includes a payment $p_i \in \mathbb{R}$ from the entrant to the incumbent and an entry date t_i, which for the simplicity of the basic model, can be either zero or one. Hence there are two types of settlement deals:

- If $t_i = 0$, the incumbent offers a license in exchange for a licensing fee p_i.
- If $t_i = 1$, the incumbent makes a reverse payment $-p_i$ in exchange for entry delay.

After having observed the offers, each entrant either accepts or rejects the settlement offer. An entrant who rejects, either litigates over the validity of the patent or waits for the market to open.

At date one, free entry to the market drives all profits to zero. Absent entry at date zero, the incumbent makes a monopoly profit $\Pi(1)$ while the entrants obtain zero profit each. If one entrant enters the market at date zero, the incumbent has a duopoly profit $\Pi(2)$, the active entrant gets a duopoly entry profit $\pi(2)$ and the one who stays out makes zero profit. If both entrants enter the market at date zero, the incumbent makes a triopoly profit $\Pi(3)$ and the entrants obtain a triopoly entry profit $\pi(3)$ each.[6] We assume that industry profit and individual profits are decreasing in the number of firms. All firms are risk-neutral, and their payoffs are linear in payments and profits. Finally, there are no entry costs.

Our equilibrium concept is sub-game perfect Nash equilibrium in pure strategies and we use backward induction to solve the game.

[5]We expect the outcomes of two separate court cases on the validity of a patent to be highly correlated. Indeed, courts often bundle similar cases.

[6]Note that the incumbent may have different profit than the entrants, capturing product differentiation between a branded drug and a generic drug in the pharmaceutical industry.

Entrants' decisions

Keeping the settlement offers fixed, consider the second stage of the negotiation game. First, by going to court, entrant $i \in \{1,2\}$ incurs litigation cost c and enters the market only in the event of patent invalidation. The litigation payoff is thus given by

$$-c + (1-\theta)\begin{cases} \pi(2) & \text{if } j \text{ is delayed} \\ \pi(3) & \text{otherwise} \end{cases} \tag{1}$$

The litigation payoff is higher when the other entrant j is delayed from entering the market. Indeed, by accepting a pay-for-delay deal, entrant j imposes a positive settlement externality to entrant i. This settlement externality is greater when the competitive effect of subsequent entry to the market is higher.[7]

Instead of going to court, entrant i may wait. Waiting allows to save on the litigation cost and yields a waiting payoff

$$(1-\theta)\begin{cases} \pi(3) & \text{if } j \text{ litigates} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The waiting payoff is higher than the litigation payoff if the other entrant j litigates. It thus follows immediately that, due to perfectly correlated litigation outcomes, litigate is never the best response to litigate.

Finally, entrant i may accept the settlement offer. The settlement payoff is $-p_i$ if $t_i = 1$ and

$$-p_i + \begin{cases} \theta\pi(2) + (1-\theta)\pi(3) & \text{if } j \text{ litigates} \\ \pi(2) & \text{if } j \text{ is delayed or waits} \\ \pi(3) & \text{if } j \text{ gets a license} \end{cases} \tag{3}$$

if $t_i = 0$. By getting a license to enter the market, the entrant is able to earn profit in the market regardless of patent strength. However, when the rival entrant litigates,

[7] We assume entrant j stays out of the market even though the court declares the patent invalid. Entrant j is still tied to the contract made by incumbent even though the patent is declared invalid. This is in line with the legal principle of *pacta sunt servanda*. Indeed, it is the ex-ante view of patent strength that matters for reaching a settlement before the court, not the validity of the patent resolved ex-post after litigation. Still, we provide an extension of our model in which entry delay is contingent on patent validity.

her profits depend on the outcome of that court case, and hence indirectly on patent strength. On the other hand, when the rival entrant settles, the patent is never challenged and there is a settlement externality between the entrants. This externality is positive when the rival entrant is delayed, whereas licensing has a negative settlement externality, reflecting non-exclusivity of the licensing deal.

Credibility of the litigation threat

For a sufficiently strong patent, the litigation payoff is negative and there is no credible litigation threat. The relevant threshold is given by

$$\hat{\theta} := 1 - \frac{c}{\pi(3)}.$$

If $\theta > \hat{\theta}$, by making unacceptable settlement offers to both entrants, the incumbent is able to monopolize the market with no cost, because waiting is a dominant strategy for both entrants. We say that the litigation threat is credible when $\theta \leq \hat{\theta}$.

Entrant 1 \ Entrant 2	litigate	wait	settle
litigate		×	×
wait	×		
settle	×		×

Table 1: Potential subgame equilibria under a credible litigation threat

When the litigation threat is credible, there are essentially three potential sub-game equilibria: (litigate,wait), (litigate,settle) and (settle,settle). These are summarized in Table 1. First, due to free-riding in litigation efforts, litigating is never a best response to litigating, because waiting allows to save on the litigation cost but otherwise yields the same payoff as going to court. The best response of entrant i to litigate is given by

$$BR_i(\text{litigate}) = \begin{cases} \text{settle} & \text{if } p_i \leq \theta(1-t_i)\pi(2) - (1-\theta)t_i\pi(3), \\ \text{wait} & \text{otherwise,} \end{cases} \tag{4}$$

where the upper bound on the payment depends on the type t_i of the settlement deal. The highest possible licensing fee equals the expected entry profit only attainable through a licensing agreement. This amount is given by a share θ of the duopoly entry profit made by the entrant when the court upholds the patent. On the other hand, the lowest possible reverse payment captures the payoff from waiting and free-riding on the litigation effort, which is given by a share $1-\theta$ of the triopoly entry profit.

Second, under a credible litigation threat, waiting is never a best response to waiting. This is because the payoff from waiting is zero, whereas the litigation payoff is always positive. The best response to wait is given by

$$BR_i\left(\text{wait}\right)=\begin{cases}\text{settle} & \text{if } p_i \leq c+\left(1-t_i\right)\pi\left(2\right)-\left(1-\theta\right)\pi\left(3\right), \\ \text{litigate} & \text{otherwise.}\end{cases} \tag{5}$$

Now the highest possible licensing fee captures the avoided litigation cost as well as the duopoly profit the licensee obtains while the rival entrant waits until patent expiration. The lowest possible reverse payment equals the litigation payoff when the rival entrant free rides on the litigation effort.

Third, as the litigation payoff is always positive, waiting is never a best response to settling either. The best response to settling is given by

$$BR_i\left(\text{settle}\right)=\begin{cases}\text{settle} & \text{if } p_i \leq c+\left(\theta-t_i\right)\left[t_j\pi\left(2\right)+\left(1-t_j\right)\pi\left(3\right)\right], \\ \text{litigate} & \text{otherwise,}\end{cases} \tag{6}$$

where the upper bound on the payment now depends on both entry dates. When both entrants settle, the highest possible licensing fee captures the litigation cost and a share θ of the entry profit made by the licensee. This is the duopoly profit when the other entrant stays out and the triopoly profit when both entrants obtain a license. The lowest possible reverse payment accounts for the litigation payoff. This is proportional to the duopoly entry profit if the rival entrant stays out and the triopoly profit if the rival entrant gets a license.

At this point, it is convenient to divide the equilibrium candidates into two groups: with and without litigation. Thus, we will, firstly, analyse equilibrium candidates in which both of the entrants settle, and, secondly, when one of them litigates, and the other one settles or waits. Finally, we will discuss which of the candidates will actually constitute an equilibrium.

Equilibrium candidates without litigation

We start by focusing on the cases when the incumbent settles with both entrants. The following lemma characterizes settlement offers, which are both accepted in the equilibrium.

Lemma 1. *For any settlement offers under a credible litigation threat, there exists a subgame equilibrium in which both entrants settle if and only if*

$$p_1 \leq c + (\theta - t_1)\left[t_2 \pi(2) + (1 - t_2)\pi(3)\right],$$
$$p_2 \leq c + (\theta - t_2)\left[t_1 \pi(2) + (1 - t_1)\pi(3)\right].$$

Furthermore, the equilibrium is unique if the inequalities are strict.

Proof. Action profile (settle,settle) constitutes an equilibrium if and only if for both entrants the best response to settle is settle, being equivalent to the inequalities. To show uniqueness, suppose the inequalities are strict. This immediately rules out (settle,litigate) and (litigate,settle) as equilibria. Furthermore,

$$\begin{aligned}(\theta - t_i)\left[t_j \pi(2) + (1 - t_j)\pi(3)\right] &= (1 - t_i)\left[t_j \pi(2) + (1 - t_j)\pi(3)\right] \\ &\quad - (1 - \theta)\left[t_j \pi(2) + (1 - t_j)\pi(3)\right] \\ &\leq (1 - t_i)\pi(2) - (1 - \theta)\pi(3)\end{aligned}$$

Thus

$$p_i < c + (1 - t_i)\pi(2) - (1 - \theta)\pi(3)$$

for both i. This litigate is not a best response to wait, ruling out (litigate,wait) and (wait,litigate) as equilibria. But then the equilibrium (settle,settle) is unique. □

It is never profitable for the incumbent to make an entrant strictly better off from a settlement. Thus, the payments in equilibria with settlements are pinned down by indifference: an entrant who settles in equilibrium must be indifferent between accepting or rejecting the settlement deal, keeping the strategy adopted by the rival entrant fixed.

In any equilibrium with two settlements, for both i the payments thus satisfy[8]

$$p_i = c + (\theta - t_i)\left[t_j \pi(2) + (1 - t_j)\pi(3)\right] \tag{7}$$

Therefore, a reverse payment captures the litigation payoff that an entrant i would obtain by rejecting the deal and going to court, whereas a licensing fee stands for the difference between the profits made by the licensee and the payoff, she would obtain from litigation. The incumbent charges a higher licensing fee for exclusivity as compared to the situation where both entrants obtain a license, and must pay a higher reverse payment when both entrants are delayed. The incumbent's payoff is given by the payments and the profit she earns from the market:

$$t_1 t_2 \Pi(1) + (t_1 + t_2 - t_1 t_2)\Pi(2) + (1 - t_1 t_2)\Pi(3) + p_1 + p_2 \tag{8}$$

where

$$p_1 + p_2 = 2c + \left[2\theta - (t_1 + t_2)\right]\pi(3) + \underbrace{\left[\theta(t_1 + t_2) - 2t_1 t_2\right]\left[\pi(2) - \pi(3)\right]}_{\text{settlement externalities}} \tag{9}$$

gives the total payment. Given the competitive effects of entry and patent strength, by comparing the incumbent's payoffs from different settlement strategies, we have the following result.

Proposition 1. *Suppose the litigation threat is credible and there is no litigation in equilibrium. Then, there exist thresholds of patent strength*

$$\overline{\theta} := 1 - \frac{\min\left\{\Pi(1) - \Pi(2) - \pi(2), \frac{1}{2}\left[\Pi(1) - \Pi(3) - 2\pi(3)\right]\right\}}{\left[\pi(2) - \pi(3)\right]}$$

$$\underline{\theta} := 1 - \frac{\max\left\{\frac{1}{2}\left[\Pi(1) - \Pi(3) - 2\pi(3)\right], \Pi(2) + \pi(2) - \Pi(3) - 2\pi(3)\right\}}{\left[\pi(2) - \pi(3)\right]}$$

such that the unique equilibrium of the game is (delay,delay) if $\theta > \overline{\theta}$, (license,delay) if $\underline{\theta} < \theta < \overline{\theta}$ and (license,license) if $\theta < \underline{\theta}$. Furthermore, $\underline{\theta} = \overline{\theta}$ if

$$\Pi(1) - \Pi(2) - \pi(2) \geq \Pi(2) + \pi(2) - \Pi(3) - 2\pi(3)$$

[8] Note that if the incumbent wants to ensure the uniqueness of the equilibrium she can just decrease the payment be an ϵ that is close to zero.

(the industry profit is convex in the number of entrants) and $\underline{\theta} < \overline{\theta}$ otherwise (the industry profit is concave in the number of firms). In particular, $\overline{\theta} \leq 0$ if

$$\Pi(1) \geq \overline{\Pi} := 2\pi(2) + \max\{\Pi(3), \Pi(2) - \pi(3)\}$$

Proof. Follows immediately by comparing the incumbent's payoffs from different settlement strategies and choosing the offers with the highest payoff. □

Assuming that there is no litigation in equilibrium, Proposition 1 shows that the number of delayed entrants is weakly increasing in patent strength. Since a reverse payment accounts for a share $1 - \theta$ of the profits a delayed entrant loses, the cost of entry delay decreases in patent strength. Thus, intuitively, entry delay is more likely to occur when the patent is strong. However, the licensing revenue also increases in patent strength, because the incumbent is able to charge a share θ of the profits a licensee makes in the market. As the industry profit is decreasing in the number of entrants, the former effect dominates the latter, so that entry delay is more likely to occur when the patent is strong.

Furthermore, when subsequent entry reduces industry profit more than the first entrant (the industry profit is concave in the number of entrants), an incumbent with an intermediate patent may adopt a divide-and-conquer strategy in which one entrant is delayed while the other entrant gets a license. On the other hand, when the first entrant reduces the industry profit more then the second one (the industry profit is convex in the number of entrants), such settlement strategy is never optimal for the incumbent.

Both thresholds of patent strength are decreasing in the gain from monopolization and increasing in the difference between duopoly and triopoly entry profit, which captures the intensity of the settlement externality between the two entrants. If the gain from monopolization is high and entry profit is not very sensitive to subsequent entry, even a weak patent can be enough to monopolize the market. In particular, if $\Pi(1) \geq \overline{\Pi}$, the incumbent always delays both entrants regardless of patent strength. Note that $\overline{\Pi}$ is more than twice the duopoly entry profit.[9]

[9] Later in the general model we shall see that this threshold increases in the number of potential entrants to the market.

Equilibrium candidates with litigation

Suppose now that there is litigation in equilibrium, there are essentially two possible equilibria: (litigate,settle) and (litigate,wait). First, by offering unacceptable settlement deals to both entrants, the incumbent ensures that one entrant litigates while the other one waits. The incumbent's payoff is then

$$\theta\Pi(1) + (1-\theta)\Pi(3) - C \tag{10}$$

With probability θ the incumbent wins the court case and keeps her monopoly, but with probability $1-\theta$ the court declares the patent invalid and both entrants enter the market. No matter what is the outcome, the incumbent incurs the litigation cost C.

Second, by making an unacceptable settlement offer to one of the entrants while keeping the other one indifferent between accepting the settlement deal and waiting, the incumbent can ensure that the entry game has a unique equilibrium in which one entrant litigates while the other one settles.

Lemma 2. *For any settlement offers under a credible litigation threat, there exists a subgame equilibrium in which entrant $i \in \{1,2\}$ settles while the rival entrant j litigates if and only if*

$$\begin{aligned} p_i &\leq \theta(1-t_i)\pi(2) - (1-\theta)t_i\pi(3), \\ p_j &\geq c + (\theta - t_j)[t_i\pi(2) + (1-t_i)\pi(3)]. \end{aligned}$$

Furthermore, the equilibrium is unique if the inequalities are strict and $p_j > c$, $t_j = 1$.

Proof. The strategies (settle,litigate) constitute an equilibrium if and only if entrant i's best response to litigate is to settle and entrant j's best response to settle is to litigate. This is equivalent to the first two inequalities and rules out (wait,litigate) and (settle,settle) as equilibria if the inequalities are strict. Furthermore,

$$\begin{aligned} \theta(1-t_i)\pi(2) - (1-\theta)t_i\pi(3) &= (1-t_i)\pi(2) - (1-\theta)[t_i\pi(3) + (1-t_i)\pi(2)] \\ &\leq (1-t_i)\pi(2) - (1-\theta)\pi(3) \end{aligned}$$

rules out (litigate,wait) as an equilibrium, leaving us with (litigate,settle) as the only candidate for another equilibrium. This is ruled out when $p_j > c$ and $t_j = 1$, because then entrant j's best response to litigate is to wait. □

Again, it is never profitable for the incumbent to make an entrant strictly better off from a settlement. When entrant j litigates, entrant i is indifferent between accepting the settlement deal or rejecting and waiting when

$$p_i = \theta (1 - t_i) \pi (2) - (1 - \theta) t_i \pi (3) \tag{11}$$

By pursuing litigation against entrant j, the incumbent can pay a small reverse payment but gets a high licensing fee. The reverse payment captures the expected triopoly profit the entrant would obtain by waiting, whereas the licensing fee stands for the expected duopoly profit the entrant can make only through the settlement when the patent stays valid. This is the difference between the profit made by the licensee and the waiting payoff. The incumbent's payoff is now given by

$$\theta [t_i \Pi (1) + (1 - t_i) \Pi (2)] + (1 - \theta) [t_i \Pi (2) + (1 - t_i) \Pi (3)] + p_i - C \tag{12}$$

We observe that licensing and litigation are substitutes for the incumbent, in the sense of having the same impact on the outside option of the another entrant. Comparing the payoffs between (litigate,license) and (litigate,wait) we immediately see that (litigate,license) is never an equilibrium. Since monopoly maximizes industry profit, the incumbent is always better off by not offering a licensing agreement. Then both entrants will enter the market only if the incumbent loses in court. If the incumbent wins, she monopolizes the market. This means that in equilibrium with litigation, a settlement is always a pay-for-delay agreement.

Proposition 2. *Suppose the litigation threat is credible and there is litigation in equilibrium. Then, the equilibrium of the game is (litigate,delay) if*

$$\Pi (2) > \Pi (3) + \pi (3)$$

and (litigate,wait) if the reverse holds.

Proof. By the assumption that the industry profit is decreasing in the number of firms, the incumbent's payoff from (litigate,wait) always exceeds that from (litigate,license) so that litigation and licensing never coexist. The rest follows by comparing the incumbent's payoff from (litigate,wait) and (litigate,delay). □

Proposition 2 shows that entry delay is independent of patent strength when litigation costs are sufficiently low for litigation to occur in equilibrium. Given that there is litigation, the incumbent monopolizes the market with probability θ. Otherwise there is entry to the market, and the incumbent can either compete in a triopoly or duopoly, but needs to compensate the delayed entrant with a reverse payment that is proportional to $1-\theta$. Therefore, the incumbent chooses to delay one of the entrants independently of patent strength. Only the intensity of competition matters. If the triopoly profit of an entrant is less than the change in incumbent's profit when moving from duopoly to triopoly, then the incumbent delays the entrant.

Litigation for patents with intermediate strength

When both entrants obtain a license, there is a negative settlement externality between them, because they make less profits in the market. As the licensing fees capture a share θ of these profits, the incumbent gets less licensing revenue. Furthermore, the incumbent's own profits are decreased. By the same logic, a delayed entrant must be compensated more when the other entrant is delayed as well: there is a positive settlement externality that increases the individual reverse payments.

However, a licensee is still willing to pay a high licensing fee when the other entrant litigates, because the litigator does not enter the market when the patent is upheld in court. Still, the incumbent gets away with a small reverse payment, as the waiting payoff equals the expected triopoly profit. On the downside the incumbent is not able to capture the avoided litigation costs under litigation and on top of that incurs her own litigation cost. The optimal settlement design compares the benefits of litigation to the cost of litigation, because the incumbent effectively loses the total litigation costs in court. We have the following result.

Proposition 3. *For any patent strength $\theta \leq \hat{\theta}$, the equilibrium of the game is (delay,delay) if the gain from monopolization is sufficiently high,*

$$\Pi(1) \geq \overline{\Pi} = 2\pi(2) + \max\{\Pi(3), \Pi(2) - \pi(3)\} \qquad (13)$$

Otherwise, there exists an interval $\Theta \subseteq \left[0, \hat{\theta}\right]$ of patent strength such that there is litigation in equilibrium if $\theta \in \Theta$ and no litigation otherwise. The interval shrinks in total litigation costs $C + 2c$ and is empty for costs large enough. In particular, for zero litigation costs, $\Theta = [0, 1]$.

Proof. See the Appendix. □

If the gain from monopolizing the market is sufficiently high, $\Pi(1) \geq \overline{\Pi}$, the incumbent will always delay both entrants, regardless of patent strength. Otherwise, patent strength plays a role in determining the equilibrium strategies: there exists an interval of patent strength such that benefits from litigation exceed the total cost of going to court; in particular, for zero litigation costs, this is true for any patent strength. The intuition why litigation occurs for intermediate patent strength is that, by going to court, the incumbent has the chance of monopolizing the market, but spends money in litigation. For a sufficiently weak patent, the chance of winning in court and monopolizing the market is too low compared to the litigation cost. For a sufficiently strong patent, the cost of litigation is too high compared to the small reverse payments needed to delay entry with settlements. Figure 1 depicts equilibria of the game as a function of patent strength and costs of litigation.

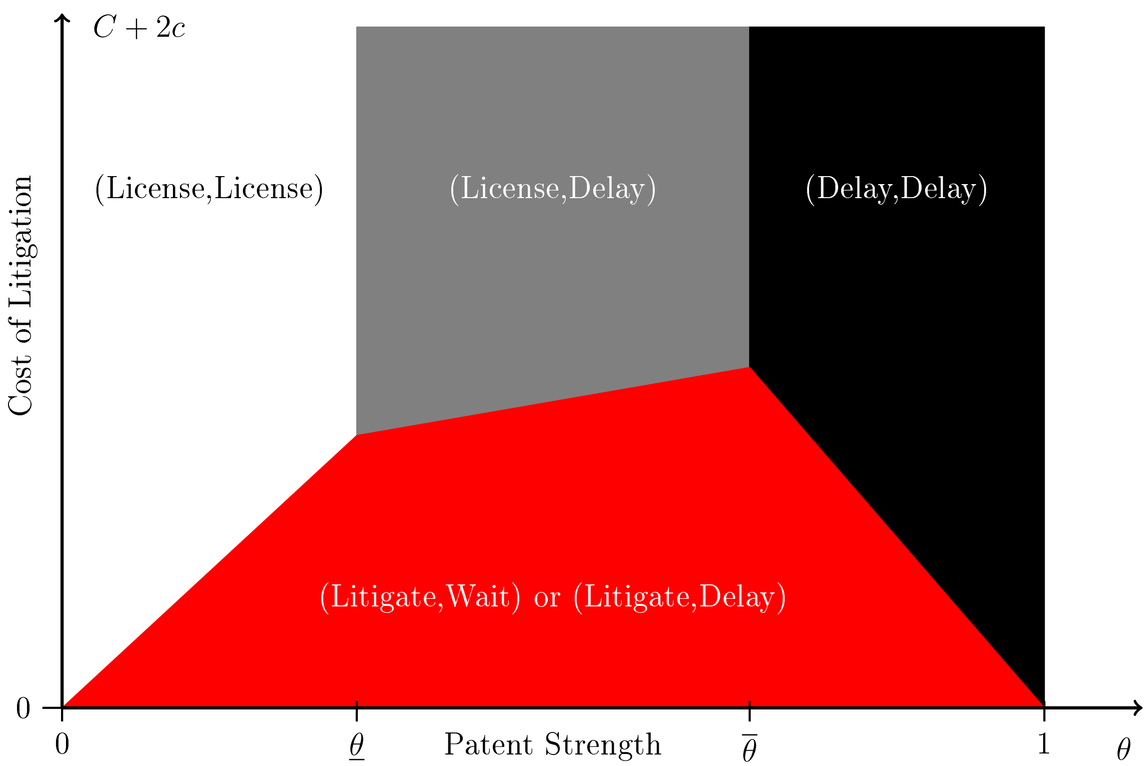

Figure 1: The Equilibrium of the Game

4 General model

We shall now consider a generalized version of the basic model, allowing for $N \geq 2$ symmetric entrants and settlement offers with intermediate entry dates. We will, firstly, show how design of optimal settlement depends on the number of players. Secondly, we will argue that a basic logic described in the two-entrants model extends to a general case. The timing of the game is analogous to the baseline model. The incumbent moves first and offers each entrant $i \in \{1, 2, \ldots, N\}$ a settlement deal, which consists of an entry date $t_i \in [0, 1]$ and a payment $p_i \in \mathbb{R}$ from the entrant to the incumbent. After the entrants have observed these offers, they decide whether to accept or reject the settlement offer. If an entrant rejects, she chooses to wait or litigate.

Payoffs are modeled, in a similar manner as in the basic model, using instantaneous profit functions $\pi(\cdot)$ for the entrants and $\Pi(\cdot)$ for the incumbent. As before, we assume that the industry profit and individual profits are decreasing in the number of entrants. Furthermore, we introduce a litigation period that ends at $l \in [0, 1)$ when the court decides whether the patent is valid or not.

Finally, considering entrants other than i, it will be convenient to denote $s_i(t)$ as the number of rival entrants who settle with an entry date t or later. Then, at any date t we can categorize entrants into three groups:

$$s(t) = \begin{cases} s_i(t) + 1 & \text{if entrant } i \text{ settles with } t_i \geq t \\ s_i(t) & \text{otherwise} \end{cases}$$

delayed entrants, $s(0) - s(t)$ number of entrants who settle and enter the market before t, and $N - s(0)$ are those who have rejected settlement offer.

Entrants' decisions

We start by taking settlement offers as given and compare payoffs from different actions taken by an entrant. We are particularly interested in the impact of increasing the number of entrants on the strength of externalities due to settlement and litigation decisions. Accordingly, the payoff from starting litigation depends, not only on patent strength and associated costs, but also on the number of entrants that have not settled (and thus could enter the market if litigation is successful), and on how long it takes for a court to decide in the case. The litigation payoff of entrant i is given by

$$(1-\theta)\int_{l}^{1}\pi\left(1+N-s_{i}\left(t\right)\right)\mathrm{d}t-c \tag{14}$$

where $s_i(t)$ denotes the number of entrants other than i who settle with an entry date t or later, while the others either litigate or wait and free-ride on the litigation effort. Each delayed entrant imposes a positive settlement externality to entrant i whose litigation payoff is increased: the more entrants are delayed at any given point of time, the higher is the expected payoff from litigation. In particular, the litigation payoff is the lowest when no entrant is delayed beyond litigation period, i.e. $s_i(l) = 0$.

As before, waiting allows to save on the litigation cost and free-ride on the litigation effort taken by some other entrant. If a rival entrant litigates, the waiting payoff is

$$\int_{l}^{1}\pi\left(1+N-s_{i}\left(t\right)\right)\mathrm{d}t \tag{15}$$

and zero otherwise. Finally, by accepting the settlement deal, the entrant gets a settlement payoff

$$\begin{aligned}-p_i+&\int_{t_i}^{\max\{t_i,l\}}\pi\left(2+s_i\left(0\right)-s_i\left(t\right)\right)\mathrm{d}t+\theta\int_{\max\{t_i,l\}}^{1}\pi\left(2+s_i\left(0\right)-s_i\left(t\right)\right)\mathrm{d}t \\ &+\left(1-\theta\right)\int_{\max\{t_i,l\}}^{1}\pi\left(2+\tilde{N}_i-s_i\left(t\right)\right)\mathrm{d}t\end{aligned} \tag{16}$$

where $\tilde{N}_i = N-1$ if there is an entrant who litigates and $\tilde{N}_i = s_i(0)$ otherwise. During the litigation period entry may occur only through a settlement. This means that an entrant with an early entry date $t < l$ is first active in the market only with other entrants who settled. However, after the litigation period, with probability $1-\theta$ there is also entry through litigation, provided that at least one of the entrants went to court.

Credibility of the litigation threat

Similarly as before, for a sufficiently strong patent, there is no credible litigation threat. In a game with N entrants and time-consuming litigation, the relevant threshold is given by

$$\hat{\theta}\left(N\right):=1-\frac{c}{\left(1-l\right)\pi\left(1+N\right)}, \tag{17}$$

which is weakly decreasing in N. Due to free-riding, the litigation threat is less likely to be credible when the number of rival entrants is high. By making unacceptable settlement offers, the incumbent ensures that a litigator faces competitive pressure from N entrants after winning in court.

From now on we assume the litigation threat is credible: $\theta \leq \hat{\theta}(N)$. Unless all entrants accept settlement offers, $N - s(0)$ entrants play a game of chicken in going to court, so that in equilibrium only one of those entrants litigates while the others wait and free ride on this effort.[10]

Lemma 3. *Under a credible litigation threat, a litigation game has exactly $N - s(0)$ pure strategy equilibria, in which one entrant litigates while others wait.*

Proof. Suppose there are at least two entrants who have rejected settlement offers, and that one of them litigates. All others have the best response to wait because waiting saves on their own litigation effort:

$$(1-\theta)\int_l^1 \pi(1+N-s_i(t))\,\mathrm{d}t \geq (1-\theta)\int_l^1 \pi(1+N-s_i(t))\,\mathrm{d}t - c$$

However, in this case, the best response of the litigator is to litigate, as otherwise, nobody would challenge the patent:

$$(1-\theta)\int_l^1 \pi(1+N-s_i(t))\,\mathrm{d}t - c \geq (1-\theta)(1-l)\,\pi(1+N) - c \geq 0$$

where the last inequality follows from the assumption that the litigation threat is credible. □

An increase in the number of entrants makes the free-riding problem more severe. However, the logic from the basic model is still valid. The game has multiple equilibria in pure strategies, since any entrant that does no settle may end up litigating while the others wait. Accordingly, as before, under credible litigation threat we can rule out some sets of strategies from equilibria: litigate is never a best response to litigate, and wait is never a best response to wait or settle when no other entrant litigates. Hence there is no litigation only when all entrants settle with the incumbent.

The best responses can be conveniently characterized for two types of strategy profiles taken by the rival entrants. First, when none of the other $N-1$ entrants litigates,

[10] We focus here on pure strategy equilibria, but there is also a mixed strategy equilibrium in which each mixes between litigate and wait, and obtains the litigation payoff.

the best response of entrant i is to settle if

$$p_i \le c + \int_{t_i}^{1} \pi \left(2 + s_i\left(0\right) - s_i\left(t\right)\right) \mathrm{d}t - \left(1-\theta\right) \int_{l}^{1} \pi \left(1 + N - s_i\left(t\right)\right) \mathrm{d}t \tag{18}$$

and to litigate otherwise. The upper bound for the payment captures the litigation cost and then two terms, where the first term represents the profits made by the entrant after her entry date and the second term is the litigation payoff the entrant would obtain by rejecting the settlement deal and going to court. Second, when at least one of the other $N-1$ entrants litigates, entrant i has a best response to settle if

$$\begin{aligned} p_i \le & \int_{t_i}^{\max\{t_i,l\}} \pi \left(2 + s_i\left(0\right) - s_i\left(t\right)\right) \mathrm{d}t + \theta \int_{\max\{t_i,l\}}^{1} \pi \left(2 + s_i\left(0\right) - s_i\left(t\right)\right) \mathrm{d}t \\ & - \left(1-\theta\right) \int_{l}^{\max\{t_i,l\}} \pi \left(1 + N - s_i\left(t\right)\right) \mathrm{d}t \end{aligned} \tag{19}$$

and to wait otherwise. Now the upper bound on the payment takes into account the possibility that, even if entrant i settles, $N-1-s_i\left(0\right)$ entrants enter the market after the litigation period with probability $1-\theta$.

In our basic model with two entrants and no litigation period we restricted attention to two types of settlement deals, licensing with $t_i = 0$ and pay-for-delay with $t_i = 1$. The objective is now to derive formally optimally designed settlement offers and to show how the logic from the basic model extends to a more general case. As in the basic model, we proceed in two steps, looking at equilibrium candidates with and without litigation. We begin by considering the game without litigation.

Equilibrium candidates without litigation

Assuming that the litigation threat is credible, the incumbent must settle with all N entrants to avoid a court case on patent validity. The following lemma characterizes subgame equilibria with N settlements.

Lemma 4. *For any settlement offers under a credible litigation threat, there exists a* *subgame equilibrium in which all N entrants settle if and only if*

$$\begin{aligned} p_i \leq & c + \int_{t_i}^{\max\{t_i,l\}} \pi\left(1+N-s\left(t\right)\right) \mathrm{d}t + \theta \int_{\max\{t_i,l\}}^{1} \pi\left(1+N-s\left(t\right)\right) \mathrm{d}t \\ & - \left(1-\theta\right) \int_{l}^{\max\{t_i,l\}} \pi\left(2+N-s\left(t\right)\right) \mathrm{d}t \end{aligned}$$

for all $i \in \{1, 2, \ldots, N\}$. Furthermore, this equilibrium is unique if the inequalities are strict.

Proof. See the Appendix. □

Lemma 4 generalizes Lemma 1 and shows that there exists a unique subgame equilibrium in which all N entrants settle, provided that the payments charged by the incumbent are not too high. The relevant upper bound on p_i depends not only on the deal offered to entrant i but also on the deals offered to the rival entrants.

As before, it is never profitable for the incumbent to make an entrant strictly better off from a settlement. As such, in equilibrium with N settlements, the payment offered to entrant i is pinned down as a function of her entry date (and the settlement strategies of rival entrants):

$$\begin{aligned} p_i\left(t_i\right) = & c + \int_{t_i}^{\max\{t_i,l\}} \pi\left(1+N-s\left(t\right)\right) \mathrm{d}t + \theta \int_{\max\{t_i,l\}}^{1} \pi\left(1+N-s\left(t\right)\right) \mathrm{d}t \\ & - \left(1-\theta\right) \int_{l}^{\max\{t_i,l\}} \pi\left(2+N-s\left(t\right)\right) \mathrm{d}t. \end{aligned} \tag{20}$$

Naturally, the payment decreases with the number of entrants present at any time in the market. By refusing to accept a licensing agreement, an entrant stays out of the market for the entire duration of the litigation period; therefore, when she accepts, a licensing fee will amount to entirety of the profits $\pi\left(1+N-s\left(t\right)\right)$ in this period at any time $t \in [0, l]$. The payment is increasing in the number of delayed entrants, reflecting the positive settlement externality of entry delay to the licensees, that is captured by the incumbent in licensing fees.

However, after the litigation period lapses, the entrant could successfully challenge patent validity in the court of law, which has two implications. Firstly, the licensing fee must be discounted by patent strength, so the incumbent captures only a share θ of the positive settlement externalities due to delaying. Secondly, the entrant expects a positive profit if the patent is invalidated. This ”foregone” profit is increasing in the number of delayed entrants. Importantly, this effect has a negative impact on the profit of the incumbent, since each payment associated with delaying entry increases with each additional delayed entrant.

Given the settlement offers, the incumbent's payoff when all N entrants settle can be written as

$$\int_0^1 \Pi\left(1+N-s\left(t\right)\right)\mathrm{d}t+\sum_{i=1}^{N} p_i\left(t_i\right). \tag{21}$$

The incumbent could potentially offer settlement agreements, that involve both licensing and entry delay, so that exclusion period is less than the scope of the patent. However, it turns out that the incumbent will always prefer to offer a mix between licensing and fully delaying contracts rather than intermediate ones.

Proposition 4. *Suppose the litigation threat is credible and there is no litigation in equilibrium. Then, in equilibrium, $N-s\left(\theta\right)$ entrants make a licensing deal with an early entry date equal to the court ruling date l, whereas*

$$s\left(\theta\right)\in\arg\max_{s\leq N}\left\{\Pi\left(1+N-s\right)+\underbrace{\left(N-s\right)\theta\pi\left(1+N-s\right)}_{\textit{licensing fees}}-\underbrace{s\left(1-\theta\right)\pi\left(2+N-s\right)}_{\textit{reverse payments}}\right\}$$

entrants conclude a pay-for-delay agreement with a late entry date equal to the patent expiration date 1. *Furthermore, the number of delayed entrants $s\left(\theta\right)$ is weakly increasing in patent strength θ.*

Proof. See the Appendix. □

Proposition 4 generalizes Proposition 1 to the case of N entrants. The settlement externalities are more pronounced when the number of potential entrants to the market is large. This means that the extreme strategies of licensing to all or delaying everyone are less likely to constitute an equilibrium. On the one hand, the incumbent needs to pay a high total reverse payment for delaying many entrants, because more entrants need to be compensated and individual reverse payments go up when there is less competition in the market. On the other hand, licensing becomes more profitable when the number of delayed entrants is higher. Therefore, the optimal settlement strategy of the incumbent is typically a divide-and-conquer strategy, where some entrants receive a license while the others are delayed.

By not allowing entry to the market during the litigation period, the incumbent secures monopoly profits, which are weakly higher than the sum of competitive profits and licensing fees. Thus there is no entry to the market during the litigation period. After the litigation period, the incumbent has to account for the possibility of entry in the event of patent invalidation, hence the need to lock some entrants into delaying agreements. Reverse payments needed to delay entry are increasing in number of delayed entrants, positive externality imposed by an entrant who accepts a delaying settlement on profits of all those entrants that decide to enter the market, either through litigation or licensing. In the extreme case of delaying all N entrants until patent expiry each of them has to receive $(1-l)(1-\theta)\pi(2)-c$, for sufficiently strong competitive effect (or large N) incumbent will accommodate some entry either through licensing or litigation, as these decrease the payment associated with delaying all other entrants to expected (triopoly) profit.

Equilibrium candidates with litigation

Consider now that there is litigation in equilibrium. By making unacceptable settlement offers to a subset of the entrants, the incumbent ensures that these entrants will not settle and by Lemma 3 one of them pursues litigation. The following lemma characterizes subgame equilibria under litigation.

Lemma 5. *For any settlement offers under a credible litigation threat, there exists a subgame equilibrium in which $s(0)$ entrants settle, one entrant litigates while the other entrants wait only if*

$$p_i \leq \int_{t_i}^{\max\{t_i,l\}} \pi\left(1+s(0)-s(t)\right)\mathrm{d}t + \theta\int_{\max\{t_i,l\}}^{1} \pi\left(1+s(0)-s(t)\right)\mathrm{d}t \\ -(1-\theta)\int_{l}^{\max\{t_i,l\}} \pi\left(2+N-s(t)\right)\mathrm{d}t$$

holds for each entrant i who settles. Furthermore, the equilibrium exists if $p_j > c$ and $t_j = 1$ for all other j.[11]

Proof. See the Appendix. □

Again, it is never profitable for the incumbent to make an entrant strictly better off from a settlement. When entrant j litigates, entrant i is indifferent between accepting

[11] There are $N - s(0)$ equilibrium candidates, because any one of the entrants who refuses to settle may litigate in equilibrium.

the settlement deal or rejecting and waiting when the payment satisfies

$$\begin{aligned} p_i\left(t_i\right) = &\int_{t_i}^{\max\{t_i,l\}} \pi\left(1+s\left(0\right)-s\left(t\right)\right)\mathrm{d}t + \theta\int_{\max\{t_i,l\}}^{1} \pi\left(1+s\left(0\right)-s\left(t\right)\right)\mathrm{d}t \\ &-\left(1-\theta\right)\int_{l}^{\max\{t_i,l\}} \pi\left(2+N-s\left(t\right)\right)\mathrm{d}t. \end{aligned}$$

The incumbent's payoff is now written as

$$\begin{aligned} \int_0^l \Pi\left(1+s\left(0\right)-s\left(t\right)\right)\mathrm{d}t + \theta\int_l^1 \Pi\left(1+s\left(0\right)-s\left(t\right)\right)\mathrm{d}t& \\ +\left(1-\theta\right)\int_l^1 \Pi\left(1+N-s\left(t\right)\right)\mathrm{d}t + \sum_{i=1}^{N} p_i\left(t_i\right)& \end{aligned} \tag{22}$$

As in the basic model, licensing and litigation are substitutes for the incumbent, in the sense of having the same impact on the outside option of the another entrant. The following result summarizes equilibrium when litigation costs are low.

Proposition 5. *Suppose the litigation threat is credible and there is litigation in equilibrium. Then, in equilibrium, one entrant litigates, $N-s_0-1$ entrants wait and*

$$s_0 \in \arg\max_{s<N} \left\{ \Pi\left(1+N-s\right) - \underbrace{s\pi\left(2+N-s\right)}_{\text{reverse payments}} \right\}$$

entrants conclude a pay-for-delay agreement with a late entry date equal to the patent expiration date 1. *Furthermore, the number of delayed entrants is independent of patent strength.*

Proof. See the Appendix. □

Licensing and litigation do not coexist in equilibrium, because, firstly, incumbent always prefers to preserve chance of monopolization and add another non-settling entrant, when there is an ongoing litigation; and secondly, incumbent cannot extract litigation costs anymore. It is interesting to note that, allocation between delaying and litigating does not depend on patent strength, but only on the profit functions. This is because both reverse payments and chances of patent invalidation are scaled with the probability of invalidation.

Litigation for intermediate patents

Now we compare the incumbent's payoffs and provide conditions under which litigation takes place. There are three factors that drive these result: extracting litigation costs in equilibrium without litigation, competitive effects due to entry and delay, and costs of it determined by θ. We have the following result, which generalizes Proposition 3.

Proposition 6. *For any patent strength $\theta \leq \hat{\theta}(N)$, there are N delayed entrants in equilibrium if the gain from monopolization is sufficiently high,*

$$\Pi(1) \geq \overline{\Pi}(N) := N\pi(2) + \max_{s<N} \left\{ \Pi(1+N-s) - \underbrace{s\pi(2+N-s)}_{\text{cost of entry delay}} \right\}$$

Otherwise, there exists an interval $\Theta \subseteq \left[0, \hat{\theta}\right]$ of patent strength such that there is litigation in equilibrium when $\theta \in \Theta$ and no litigation otherwise. The interval shrinks in total litigation costs and is empty for costs large enough. In particular, for zero litigation costs, $\Theta = [0, 1]$.

Proof. See the Appendix □

We offer an intuition for the proof here. The proof of the result has the following steps. We first show that the equilibrium payoff of an incumbent who settles with everybody is increasing and convex in patent strength, whereas the incumbent's payoff in equilibrium with litigation is linear and increasing in patent strength. Therefore, the difference between the litigation and no litigation equilibrium payoffs is concave in patent strength. Furthermore, when the incumbent settles with everyone, she saves on the litigation costs, which means that for costs high enough, litigation will not occur in equilibrium. At the other extreme, when total litigation costs go to zero, there will always be litigation in equilibrium.

From the Proposition 6 we infer that patents can be both too strong or too weak to be litigated. Intermediate strength patents are challenged and the likelihood of litigation decreases in total costs of litigation. In another words, when a patent is too strong (weak), chances for the patent to be invalidated (upheld) are too low compared to the litigation costs.

5 Extensions and policy implications

Conditional settlements

So far our analysis has rested on the assumption that settlement agreements remain in force even in the event of patent invalidation. A justification comes from the legal principle of *pacta sunt servanda*, which states that a contract should stay in power despite an *expected* change of environment. Any settlement agreement is reached based on perceived patent strength, before the judgment in the court of law, so the parties of such a contract had been aware of embedded risks. Therefore, it is the ex ante belief about the patent strength, rather than the ex post validity that matters for reaching a settlement. Our understanding is thus that our assumption well describes general practice.

However, settlement agreements that are conditional on patent validity are also possible (such a contract could be explicitly formulated), and sometimes used (for example, in the Servier case). Furthermore, parties of such contracts have argued that from the consumer welfare perspective conditional settlements are beneficial, as they allow more entry as soon as the patent is invalidated. In this section we argue to the contrary: we will show that allowing settlements to be conditional on patent's validity, leads to more exclusion.

Permitting all delayed entrants to enter the market if a patent has been invalidated, has immediate implication to the payoffs. Firstly, the litigation payoff of entrant i is now given by

$$(1-\theta)(1-l)\,\pi(1+N)-c \tag{23}$$

no matter what strategies the rival entrants take. Secondly, the waiting payoff is

$$(1-\theta)(1-l)\,\pi(1+N)$$

if one of the rival entrants litigates and zero otherwise. Thirdly, the settlement payoff is

$$-p_i+\int_{t_i}^{1}\pi\left(2+s_i(0)-s_i(t)\right)\mathrm{d}t \tag{24}$$

if none of the rival entrants litigates but

$$
\begin{aligned}
-p_i + & \int_{t_i}^{\max\{t_i,l\}} \pi\left(2+s_i\left(0\right)-s_i\left(t\right)\right)\mathrm{d}t + \theta\int_{\max\{t_i,l\}}^{1} \pi\left(2+s_i\left(0\right)-s_i\left(t\right)\right)\mathrm{d}t \qquad (25)\\
& + \left(1-\theta\right)\left(1-l\right)\pi\left(1+N\right)
\end{aligned}
$$

otherwise. Note that the litigation payoff is positive under a credible litigation threat and waiting still allows to avoid the litigation cost if a rival entrant litigates. As such, to litigate is never a best response to litigate, and to wait is never a best response to wait nor to settle. In particular, when none of the other entrants litigates, an entrant i has a best response to settle if

$$
p_i \leq c + \int_{t_i}^{1} \pi\left(2+s_i\left(0\right)-s_i\left(t\right)\right)\mathrm{d}t - \left(1-\theta\right)\left(1-l\right)\pi\left(1+N\right) \qquad (26)
$$

and litigate otherwise. We have the following lemma on the existence of subgame equilibria with N settlements.

Lemma 6. *For any settlement offers under a credible litigation threat and conditional settlement terms, there exists a subgame equilibrium in which all N entrants settle if and only if*

$$
p_i \leq c + \int_{t_i}^{1} \pi\left(1+N-s\left(t\right)\right)\mathrm{d}t - \left(1-\theta\right)\left(1-l\right)\pi\left(1+N\right)
$$

for all $i \in \{1, 2, \ldots, N\}$. This equilibrium is unique if the inequalities are strict. Furthermore, there exists a subgame equilibrium in which $s\left(0\right)$ entrants settle, one entrant litigates while the other entrants wait only if

$$
p_i \leq \int_{t_i}^{\max\{t_i,l\}} \pi\left(1+s\left(0\right)-s\left(t\right)\right)\mathrm{d}t + \theta\int_{\max\{t_i,l\}}^{1} \pi\left(1+s\left(0\right)-s\left(t\right)\right)\mathrm{d}t
$$

holds for each entrant i who settles. This equilibrium exists if $p_j > c$ and $t_j = 1$ for all other j.

Proof. See the Appendix. □

The key difference to Lemmas 4 and 5 for unconditional settlement terms is that the upper bound on the payments is increased. Firstly, an entrant who settles no longer foregoes profits in the event that a rival entrant litigates and wins in court. Secondly, the entrant's own litigation payoff is reduced, because the entrants who settle now enter the market after patent invalidation.

As earlier, the payments in any equilibrium are pinned down by indifference, because the incumbent never leaves an entrant strictly better off from a settlement. The following result now characterizes equilibria under conditional settlement terms.

Proposition 7. *Suppose the litigation threat is credible. Then, under conditional settlement terms, in equilibrium the incumbent pays to delay all entrants until patent expiration if*

$$(1-\theta)\left[\Pi(2) + N\pi(1+N) - \Pi(1)\right] \leq \frac{C+Nc}{1-l}$$

and otherwise pays to delay $N-1$ entrants until patent expiration while the remaining entrant litigates.

Proof. See the Appendix. □

Proposition 7 shows that, when total litigation cost is high, the incumbent pays to delay all entrants, but otherwise pursues litigation against one of the entrants. The benefit of this is that the reverse payments under litigation are zero (as the entrant is allowed to enter the market when the rival entrant wins), whereas the incumbent needs to compensate each entrant with the litigation payoff (positive by the assumption that the litigation threat is credible) to delay all entry to the market. Furthermore, as opposed to the previous result that litigation occurs for patents with intermediate strength, litigation is now more likely when the patent is strong.

The implication of proposition 7 is that conditional settlement terms tend to reduce entry to the market. Furthermore, the incumbent secures monopoly position with lower reverse payments.

Antitrust limits to patent settlements

Welfare analysis of patent settlements has to overcome two main challenges. Firstly, incorporating incentives to innovate of both the incumbent and of entrants is necessary, as these are influenced by expected profits. This is outside of the scope of this paper, and we leave it for future research. Secondly, one has to acknowledge that the patent system is imperfect, so a flawless one granting only ironclad patents is not a relevant benchmark. In reality patents are probabilistic property rights and can be invalidated in a court of law. We analyze the welfare implications of patent settlements

Evidence and theory presented in this study tackle the consequences of an imperfect patent system, these being on one hand dubious settlements and wasted litigation costs, but on the other hand, welfare improving licensing agreements. Therefore, these effects are subject of our discussion.

To analyze welfare implications of patent settlements, economic literature (Shapiro, 2003),(Elhauge and Krueger, 2012) as well as antitrust policy (Lundbeck and Servier cases) have focused on environments where there is only one entrant or a single settlement. Yet, different types of settlements occur in practice. Therefore, the overall effect of patent settlements can be ambiguous.

To study the effects of patent settlements on consumer welfare, we shall model consumer surplus in reduced form, using a weakly increasing instantaneous function $\sigma(\cdot)$ that maps the number of entrants in the market to the surplus of the consumers. In the benchmark case of no patent settlements, there is entry to the market only when the litigation threat is credible, $\theta \leq \hat{\theta}$, in which case all entrants enter the market after the expected entry date of litigation. The consumer surplus writes

$$CS = l\sigma(0) + (1-l)\left[\theta\sigma(0) + (1-\theta)\sigma(n)\right] \tag{27}$$

Proposition 8. *Suppose $\theta \in \Theta$, then, settlements decrease consumer surplus. Otherwise, settlements increase consumer surplus before the expected entry date from litigation, $\theta + l(1-\theta)$, and reduce consumer surplus afterwards; taking into account both effects, the consumers strictly benefit from settlements if and only if*

$$\sigma(n - s_1) \geq \theta\sigma(0) + (1-\theta)\sigma(n) \tag{28}$$

Proof. See Appendix. □

Proposition 8 shows that consumers are strictly better off from patent settlements when $\theta \notin \Theta$ (litigation costs are high) and the inequality 28 is satisfied.

As an illustration consider, baseline model with two entrants. In the equilibrium without litigation, consumers are better of when both entrants receive licensing settlements ($\theta \leq \underline{\theta}$), and worse off when each of the entrants receives a delaying settlement ($\theta \geq \bar{\theta}$). For intermediate patent strength consumers are better off for all $\theta > \tilde{\theta} \equiv \frac{\sigma(3)-\sigma(2)}{\sigma(3)-\sigma(1)}$

Proposition 8 has an immediate policy implication: strict prohibition of pay-for-delay settlements may not be the best policy for a competition authority to pursue.

Impossibility to enter into delaying settlements undermines incentives to license patents, and might thus decrease consumer welfare.

Numerical examples

Now we provide two numerical examples in order to illustrate some of our results, and also to show how they depend on particular market set-up.

Example 1 Suppose that the instantaneous profit functions are determined by textbook Cournot quantity-setting game with a simple inverse demand $1 - Q$, where Q denotes industry profit, firm-level outputs are $q(1) = \frac{1}{2}$, $q(2) = \frac{1}{3}$ and $q(3) = \frac{1}{4}$, resulting in equilibrium profits $\pi(1) = \frac{1}{4}$, $\pi(2) = \frac{1}{9}$ and $\pi(3) = \frac{1}{16}$ Thresholds for patent strength are $\underline{\theta} = \frac{2}{7}$ and $\bar{\theta} = \frac{3}{7}$. Furthermore, $\pi(2) - 2\pi(3) = \frac{1}{9} - \frac{1}{8} < 0$ so the equilibrium with litigation is (litigate,wait).

The instantaneous consumer surpluses can be calculated by integrating the demand function from zero to the equilibrium industry output: $\sigma(1) = \frac{3}{8}$, $\sigma(2) = \frac{4}{9}$ and $\sigma(3) = \frac{15}{32}$. So consumer are strictly better off from settlements for all $\theta > \tilde{\theta} = \frac{7}{27}$. Consumers are strictly better off from settlements if and only if there is no litigation and the patent is sufficiently weak: $\theta \leq \overline{\theta} = \frac{3}{7}$. Only if the patent is strong enough to delay both entrants from entering the market, consumers are worse off from patent settlements. The positive effect of one licensing agreement outweighs the negative effect of one pay-for-delay agreement.

Example 2 Let us now consider an oligopoly model of (Dixit, 1979). The firms have constant marginal costs normalized to zero and we assume fixed costs are sunk. The consumer utility is defined as

$$U = \sum_{i=1,2,3} [\alpha - p_i] \, q_i - \frac{\beta \sum_{i=1,2,3} q_i^2 + \gamma \, [q_1 q_2 + q_2 q_3 + q_1 q_3]}{2}$$

where q_i is the quantity purchased from firm i, who charges a price p_i. We assume $\alpha, \beta > 0$, $\beta^2 > \gamma^2$ and $\alpha \, [\beta - \gamma] > 0$. Maximization with respect to q_i allows us to compute a linear demand structure, where:

$$p_i = \alpha - \beta q_i - \sum_j q_j \gamma$$

where $j \neq i$. Notice that the goods are substitutes (complements) when γ is positive (negative). The ratio γ/β^2 gives the level of product differentiation and the market size is captured by α.

Firstly, an important observation is that the threshold of patent strength $\tilde{\theta}$ is increasing in elasticity of demand β and increasing (decreasing) in substitutability γ if goods are complements, $\gamma < 0$ (substitutes, $\gamma > 0$).

Secondly, one can note that the question whether consumer benefits from allowing settlements depends on the level of product differentiation and the elasticity of demand. In the markets where consumers have high elasticity patent settlements are more likely to benefit consumers. Moreover, the magnitude of the effect increases in the market size α, and decreases in the length of litigation.[12]

European Commission in Lundbeck and Servier cases had considered pay-for-delay settlements as restrictive to competition by object, thus illegal; US Supreme Court in the case *FTC v. Actavis, Inc.* ruled that the reverse payments are subject to a complete rule of reason analysis. The above discussion of consequences of banning settlements that an incumbent can propose, shows that welfare implications are unclear, and in fact consumers can benefit from pay-for-delay agreements as long as these are accompanied by sufficient licensing activity. Therefore, we believe that a rule of reason approach is the proper one.

Tax on reverse payment In the light of the discussion above banning patent settlements could have a negative outcome of shutting down licensing, and thus reducing consumer welfare. Here we want to propose a policy response that could limit delaying entry and promote accelerating it, namely a tax on reverse payments. As noted before the primary reason for the incumbent to accommodate entry is gaining ability to delay. Taxing payments from incumbent to entrants would make this mechanism weaker, and therefore the incumbent would have to accommodate more entrants, which increases consumer surplus.

Consider imposing tax τ, hence incumbent's problem in the equilibrium when she

[12] See the Appendix for the details

settles with all entrants becomes:

$$s(\theta,\tau) \in \arg\max_{s \le N} \left\{ \Pi(1+N-s) + \underbrace{(N-s)\,\theta\pi\,(1+N-s)}_{\text{licensing fees}} - (1+\tau)\underbrace{s\,(1-\theta)\,\pi\,(2+N-s)}_{\text{reverse payments}} \right\}$$

and in the case of equilibrium with litigation:

$$s_0 \in \arg\max_{s<N} \left\{ \Pi(1+N-s) - (1+\tau)\underbrace{s\pi\,(2+N-s)}_{\text{reverse payments}} \right\}$$

Proposition 9. *Imposing a tax τ on reverse payments weakly increases number of accommodated entrants in both equilibria.*

Proof. See the Appendix. □

Proposition 9 shows that imposing tax weakly reduces number of delayed entrants in each equilibrium. Key issue is however to investigate whether imposing such taxes will not move the equilibrium from the one with licensing to the one with litigation, which would be detrimental to welfare. This is however an empirical question, because as discussed above imposing taxes might change number of delayed entrants to varying degree and hence, we cannot draw general conclusions.

6 Concluding remarks

Since consumer welfare depends largely on the date of entry to the market governed by a patent, settlements aimed at delaying or accelerating entry have a serious economic impact. We propose a model of patent settlements in an environment with multiple entrants. This allows us to study an important phenomenon, which is new to the literature on patent disputes, namely settlement externalities between entrants to the market. So far mainly entry of a single firm has been a topic of economic scrutiny with respect to pay-for-delay agreements. It has been shown that an incumbent patent holder can agree with a potential entrant to delay her entry with a reverse payment that compensates for the missed profits (Shapiro, 2003). Allowing for more than one entrant highlights an important externality of such an agreement, as it makes entry for the other entrants more attractive. In fact, for all entrants to accept a pay-for-delay agreement, the incumbent must offer expected duopoly profits, as these are attainable by rejecting the settlement offer and pursuing litigation instead.

There are two types of potential equilibria: with and without litigation. Both types of equilibria are characterized by a divide-and-conquer strategy implemented by the incumbent, who divides entrants into two groups: those who make a pay-for-delay agreement and the others, who are either obtain a licensing deal or do not settle. This is so even if all entrants are completely identical in our model. We show that litigation occurs for patents with intermediate strength. At the extremes, when the patent is either sufficiently weak or strong, the cost of litigation is too high compared to the small gain in letting the court decide on patent validity.

Our results shed light on the discussion about economic consequences of pay-for-delay agreements. We develop extensions that seek to clarify the most important policy dilemmas. Firstly, we prove that settlements which are conditional on patent validity should not be allowed, because the elimination of settlement externalities leads to more entry delay. Secondly, we address the question whether there should be limits to patent settlements. We find that when there is no licensing in equilibrium, allowing for settlements is disadvantageous to consumer welfare. On the other hand, when there is licensing, the overall effect is ambiguous. This brings us to a conclusion that pay-for-delay settlements should be subject to a rule of reason analysis. Thirdly, we study consequences of imposing a tax on reverse payments. We show that such a tax would increase number of accommodated entrants conditioned on a given equilibrium type (with/without litigation). However, the overall equilibrium effect depends on a particular form of profit function.

A particular feature of the pharmaceutical market in the US is the Hatch-Waxman Act. This law aims at promoting entry of generics by guaranteeing the first entrant a duopoly position. In light of our results, such a policy response should not be effective. Once an exclusivity for the first entrant is granted, the incumbent has to just delay her entry. There are no settlement externalities, because the other entrants are excluded from the market by law.

References

Amir, R., Encaoua, D., and Lefouili, Y. (2014). Optimal licensing of uncertain patents in the shadow of litigation. *Games and Economic Behavior*, 88:320–338.

Che, Y.-K. and Spier, K. E. (2008). Exploiting plaintiffs through settlement: Divide and conquer. *Journal of Institutional and Theoretical Economics JITE*, 164(1):4–23.

Daughety, A. F. and Reinganum, J. F. (2002). Informational externalities in settlement bargaining: confidentiality and correlated culpability. *RAND Journal of Economics*, pages 587–604.

Dixit, A. (1979). Quality and quantity competition. *The Review of Economic Studies*, 46(4):587–599.

Edlin, A., Hemphill, S., Hovenkamp, H., and Shapiro, C. (2015). The actavis inference: Theory and practice. *Rutgers University Law Review*, 67:585–635.

Elhauge, E. and Krueger, A. (2012). Solving the patent settlement puzzle. *Texas Law Review*, 91:284–330.

Farrell, J. and Shapiro, C. (2008). How strong are weak patents? *American Economic Review*, 98(4):1347–69.

Gratz, L. (2012). Economic analysis of pay-for-delay settlements and their legal ruling. Discussion papers in economics, University of Munich, Department of Economics.

Kamien, M. I. and Tauman, Y. (1986). Fees versus royalties and the private value of a patent. *The Quarterly Journal of Economics*, 101(3):471–491.

Katz, M. and Shapiro, C. (1987). R&d rivalry with licensing or imitation. *American Economic Review*, 77(3):402–20.

Lemley, M. A. and Shapiro, C. (2005). Probabilistic patents. *Journal of Economic Perspectives*, 19(2):75–98.

Manganelli, A.-G. (2014). Delay competition to increase competition: Should reverse payments be banned per se?

Meunier, V. and Padilla, J. (2015). Should reverse payments be prohibited per se?

Posner, E. A., Spier, K. E., and Vermeule, A. (2010). Divide and conquer. *Journal of Legal Analysis*, 2(2):417–471.

Segal, I. (1999). Contracting with externalities. *The Quarterly Journal of Economics*, 114(2):337–388.

Segal, I. (2003). Coordination and discrimination in contracting with externalities: Divide and conquer? *Journal of Economic Theory*, 113(2):147–181.

Shapiro, C. (2003). Antitrust limits to patent settlements. *RAND Journal of Economics*, 34(2):391–411.

Appendix

Proof of Lemma 4 When none of the other entrants litigates, entrant i has a best response to settle if

$$p_i \leq c + \int_{t_i}^{1} \pi\left(2 + s_i(0) - s_i(t)\right) \mathrm{d}t - (1-\theta) \int_{l}^{1} \pi\left(1 + N - s_i(t)\right) \mathrm{d}t$$

and litigate otherwise. Note that, because $s_i(0) - s_i(t)$ is weakly increasing and $N - s_i(t)$ is weakly decreasing in the number $s_i(0)$ of other entrants who settle, the upper bound on p_i is weakly decreasing in $s_i(0)$. This means that, when entrant i has a best response to settle if $s_i(0)$ entrants settle while $N - 1 - s_i(0)$ wait, the entrant also has a best response to settle if $s_i(0) - x$ settle while $N - 1 - s_i(0) + x$ wait, where $x \leq s_i(0)$. Consequently, if each entrant has a best response to settle when all other $N-1$ entrants settle, there exists a unique subgame equilibrium with N settlements. This is because in any equilibrium without N settlements there is exactly one entrant who litigates, in which case the other $N-1$ entrants either settle or wait. But then, by the above arguments, the litigator wants to settle instead of going to court. Using the above inequality and changing notation to $s_i(t) = s(t) - 1$ if $t \leq t_i$ and $s_i(t) = s(t)$ otherwise, and letting $s(0) = N$, concludes the proof.

Proof of Proposition 4 Differentiating $p_i(t_i)$ with respect to t_i we obtain

$$\frac{\mathrm{d}p_i(t_i)}{\mathrm{d}t_i} = \begin{cases} -\pi\left(1 + N - s(t)\right) & \text{if } t_i < l \\ -\theta\pi\left(1 + N - s(t)\right) - (1-\theta)\,\pi\left(2 + N - s(t)\right) & \text{if } t_i > l \end{cases}$$

and

$$p_i(1) = c - (1-\theta) \int_{l}^{1} \pi\left(2 + N - s(t)\right) \mathrm{d}t$$

As such, the total payment can be expressed as the number of settlements times the payment associated with full entry delay, subtracted with changes in payments due to early entry dates:

$$
\begin{aligned}
\sum_{i=1}^{N} p_i(t_i) =& p_i(1)\, N - \int_0^1 (N - s(t))\, \mathrm{d}p_i(t) \\
=& \int_0^l (N - s(t))\, \pi(1 + N - s(t))\, \mathrm{d}t \\
&+ \theta \int_l^1 (N - s(t))\, \pi(1 + N - s(t))\, \mathrm{d}t \\
&- (1 - \theta) \int_l^1 s(t)\, \pi(2 + N - s(t))\, \mathrm{d}t + Nc
\end{aligned}
$$

We plug this into the incumbent's payoff to see impact of the incumbent's choice of entry dates on her profits. During the litigation period incumbent has a possibility to obtain entire industry profit:

$$
\int_0^l \left[\Pi(1 + N - s(t)) + (N - s(t))\, \pi(1 + N - s(t))\right] \mathrm{d}t
$$

Industry profit is maximized under monopoly, thus the incumbent has no incentive to allow for entry, so $s(t) = N$ for all $t \leq l$. After the litigation period, incumbent's profit depends on the outcome of litigation:

$$
\begin{aligned}
& \theta \int_l^1 \left[\Pi(1 + N - s(t)) + (N - s(t))\, \pi(1 + N - s(t))\right] \mathrm{d}t \\
+ (1 - \theta) & \int_l^1 \left[\Pi(1 + N - s(t)) - s(t)\, \pi(2 + N - s(t))\right] \mathrm{d}t \\
& + Nc
\end{aligned}
$$

Notice that the expected instantaneous profit depends on time only through entry dates, so the problem is linear. If incumbent finds profitable to delay entry until any $t > l$ she will also want to delay to $t + \epsilon > t$ and so on until patent expiry date. Consequently, the incumbent's problem reduces to selecting an entry schedule

$$
s(t) = \begin{cases} N & \text{for all } t \in [0, l] \\ s & \text{for all } t \in (l, 1] \end{cases}
$$

that maximizes its expected payoff

$$Nc + l\Pi(1) + (1-l)\left[\Pi(1+N-s) + (N-s)\theta\pi(1+N-s) - s(1-\theta)\pi(2+N-s)\right]$$

subject to $s \leq N$.

Proof of Lemma 5 First, when at least one of the entrants litigates, entrant i has a best response to settle if and only if

$$p_i \leq \int_{t_i}^{\max\{t_i,l\}} \pi(2+s_i(0)-s_i(t))\,\mathrm{d}t + \theta\int_{\max\{t_i,l\}}^{1} \pi(2+s_i(0)-s_i(t))\,\mathrm{d}t - (1-\theta)\int_{l}^{\max\{t_i,l\}} \pi(1+N-s_i(t))\,\mathrm{d}t$$

This can be rewritten using

$$s(t) = \begin{cases} s_i(t)+1 & \text{if } i \text{ settles with } t_i \geq t \\ s_i(t) & \text{otherwise} \end{cases}$$

for all $t \in [0,1]$. Furthermore, when $p_j > c$ and $t_j = 1$ hold for $N - s(0)$ entrants, they never settle in equilibrium, so that one of them litigates while the other entrants wait.

Proof of Proposition 5 Differentiating $p_i(t_i)$ with respect to t_i we obtain

$$\frac{\mathrm{d}p_i(t_i)}{\mathrm{d}t_i} = \begin{cases} -\pi(1+s(0)-s(t)) & \text{if } t_i < l \\ -\theta\pi(1+s(0)-s(t)) - (1-\theta)\pi(2+N-s(t)) & \text{if } t_i > l \end{cases}$$

and

$$p_i(1) = -(1-\theta)\int_l^1 \pi(2+N-s(t))\,\mathrm{d}t$$

As such, the total payment can be expressed as the number of settlements times the payment associated with full entry delay, subtracted with changes in payments due to early entry dates:

$$\begin{aligned}\sum_{i=1}^{N} p_i(t_i) =& p_i(1)\, s(0) - \int_0^1 (s(0) - s(t))\, \mathrm{d}p_i(t) \\ =& \int_0^l (s(0) - s(t))\, \pi(1 + s(0) - s(t))\, \mathrm{d}t \\ &+ \theta \int_l^1 (s(0) - s(t))\, \pi(1 + s(0) - s(t))\, \mathrm{d}t \\ &- (1-\theta) \int_l^1 s(t)\, \pi(2 + N - s(t))\, \mathrm{d}tc\end{aligned}$$

We plug this into the incumbent's payoff to see impact of the incumbent's choice of entry dates on her profits. During the litigation period incumbent has a possibility to obtain entire industry profit:

$$\int_0^l \left[\Pi(1 + s(0) - s(t)) + (s(0) - s(t))\, \pi(1 + s(0) - s(t))\right] \mathrm{d}t$$

Industry profit is maximized under monopoly, thus the incumbent has no incentive to allow for entry, so $s(0) = s(t)$ for all $t \leq l$. After the litigation period, incumbent's profit depends on the outcome of litigation:

$$\begin{aligned}&\theta \int_l^1 \left[\Pi(1 + s(0) - s(t)) + (s(0) - s(t))\, \pi(1 + s(0) - s(t))\right] \mathrm{d}t \\ &+ (1-\theta) \int_l^1 \left[\Pi(1 + N - s(t)) - s(t)\, \pi(2 + N - s(t))\right] \mathrm{d}t \\ &-C\end{aligned}$$

Notice that the expected instantaneous profit depends on time only through entry dates, so the problem is linear. If incumbent finds profitable to delay entry until any $t > l$ she will also want to delay to $t + \epsilon > t$ and so on until patent expiry date. Consequently, the incumbent's problem reduces to selecting an entry schedule

$$s(t) = \begin{cases} s(0) & \text{for all } t \in [0, l] \\ s & \text{for all } t \in (l, 1] \end{cases}$$

that maximizes its expected payoff

$$-C + l\Pi(1) + \theta(1-l)\Pi(1+s(0)-s) + \theta(s(0)-s)\pi(1+s(0)-s) \\ + (1-\theta)(1-l)(\Pi(1+N-s) - s\pi(2+N-s))$$

subject to $s \leq s(0) < N$. Since industry profit is decreasing in the number of firms in the market, $s(0) = s$ is optimal, where s is chosen optimally to maximize $\Pi(1+N-s) - s\pi(2+N-s)$ subject to $s < N$.

Proof of Proposition 6 (and 3) First, when all entrants settle, the incumbent's payoff

$$V_S = Nc + l\Pi(1) \\ + (1-l)[\Pi(1+N-s) + (N-s)\theta\pi(1+N-s) - s(1-\theta)\pi(2+N-s)]$$

is maximized at

$$s(\theta) \in \arg\max_{s \leq N} \{\Pi(1+N-s) + (N-s)\theta\pi(1+N-s) - s(1-\theta)\pi(2+N-s)\}$$

with the associated value function $V_S(\theta)$. Furthermore, V_S defines a family of affine functions of θ parametrized by s. The epigraph of an affine function is a half-space and any intersection of half-spaces is convex set. The value function $V_S(\theta)$ is convex if and only if its epigraph is convex. The epigraph of the value function is convex, therefore the value function is convex too. Notice that

$$\frac{\partial V_S}{\partial \theta} = (1-l)[(N-s)\pi(1+N-s) + s\pi(2+N-s)] > 0$$

Therefore $V_S(\theta)$ is strictly increasing, convex function of θ. Second, when there is litigation in equilibrium, the incumbent's payoff

$$V_L = -C + l\Pi(1) + \theta(1-l)\Pi(1) \\ + (1-\theta)(1-l)(\Pi(1+N-s) - d\pi(2+N-s))$$

is maximized at

$$s_0 \in \arg\max_{s<N} \{\Pi (1+N-s) - s\pi (2+N-s)\}$$

with the associated value function $V_L(\theta)$. Clearly s_0 does not depend on θ and therefore the value function $V_L(\theta)$ is linear and nondecreasing in θ. Since $V_S(\theta)$ is convex and $V_L(\theta)$ is linear, the difference

$$\triangle V(\theta) = V_L(\theta) - V_S(\theta)$$

is concave in θ. Notice that

$$N \in \arg\max_{s} \{\Pi (1+N-s) - s\pi (2+N-s)\}$$

implies $s(\theta) = N$ and $V_S(N) \geq V_L(s_0)$. Otherwise, $\triangle V(1) = V_L(s_0) - V_S(s(1)) = -C - Nc \leq 0$ and $\triangle V(0) = V_L(s_0) - V_S(s(0)) = -C - Nc \leq 0$. Therefore, if $C + Nc > 0$, $\{0,1\} \notin \Theta$. In particular, $C + Nc = 0$ implies $1 \in \Theta$ and $0 \in \Theta$. Now take any $\theta, \theta' \in [0,1]$ such that $\triangle V(\theta) \geq 0$ and $\triangle V(\theta') \geq 0$. Then by concavity of $\triangle V(\cdot)$ we have that

$$\triangle V(t\theta + (1-t)\theta') \geq t\triangle V(\theta) + (1-t)\triangle V(\theta') \geq 0$$

for all $t \in [0,1]$.

Proof of Lemma 6 The proof is similar to the ones for Lemma 4 and Lemma 5.

Proof of Proposition 7 The proof follows similar steps as the ones for Proposition 4 and 5. First, differentiating

$$p_i(t_i) = c + \int_{t_i}^{1} \pi (1+N-s(t))\,\mathrm{d}t - (1-\theta)(1-l)\pi(1+N)$$

with respect to t_i we obtain

$$\frac{\mathrm{d}p_i(t_i)}{\mathrm{d}t_i} = -\pi(1+N-s(t))$$

and

$$p_i(1) = c - (1-\theta)(1-l)\pi(1+N)$$

As such, the total payment can be expressed as the number of settlements times the payment associated with full entry delay, subtracted with changes in payments due to early entry dates:

$$\begin{aligned}\sum_{i=1}^{N} p_i(t_i) =& p_i(1) N - \int_0^1 (N - s(t))\,\mathrm{d}p_i(t) \\ =& \int_0^1 (N - s(t))\,\pi(1+N-s(t))\,\mathrm{d}t \\ & - (1-\theta)(1-l)N\pi(1+N) + Nc\end{aligned}$$

Plug this into the incumbent's payoff to obtain:

$$\int_0^1 \left[\Pi(1+N-s(t)) + (N-s(t))\,\pi(1+N-s(t))\right]\mathrm{d}t$$

less a constant term. Industry profit is maximized under monopoly, so $s(t) = N$ for all $t \leq 1$ is optimal. This gives the incumbent a payoff

$$\Pi(1) + Nc - (1-\theta)(1-l)N\pi(1+N)$$

Let us now consider equilibria with litigation. Letting

$$p_i(t_i) = \int_{t_i}^{\max\{t_i,l\}} \pi(1+s(0)-s(t))\,\mathrm{d}t + \theta\int_{\max\{t_i,l\}}^{1} \pi(1+s(0)-s(t))\,\mathrm{d}t$$

denote the upper bound and differentiating it with respect to t_i we obtain

$$\frac{\mathrm{d}p_i(t_i)}{\mathrm{d}t_i} = \begin{cases} -\pi(1+s(0)-s(t)) & \text{if } t_i < l \\ -\theta\pi(1+s(0)-s(t)) & \text{if } t_i > l \end{cases}$$

and

$$p_i(1) = 0.$$

As such, the total payment can be expressed as the number of settlements times the payment associated with full entry delay, subtracted with changes in payments due to early entry dates:

$$\begin{aligned}\sum_{i=1}^{N} p_i(t_i) =& p_i(1)\, s(0) - \int_0^1 (s(0) - s(t))\, \mathrm{d}p_i(t) \\ =& \int_0^l (s(0) - s(t))\, \pi(1 + s(0) - s(t))\, \mathrm{d}t \\ &+ \theta \int_l^1 (s(0) - s(t))\, \pi(1 + s(0) - s(t))\, \mathrm{d}t\end{aligned}$$

We plug this into the incumbent's payoff to see impact of the incumbent's choice of entry dates on her profits. During the litigation period incumbent has a possibility to obtain entire industry profit:

$$\int_0^l \left[\Pi(1 + s(0) - s(t)) + (s(0) - s(t))\, \pi(1 + s(0) - s(t))\right] \mathrm{d}t$$

Industry profit is maximized under monopoly, thus the incumbent has no incentive to allow for entry, so $s(t) = s(0)$ for all $t \leq l$. After the litigation period, incumbent's profit depends on the outcome of litigation:

$$\begin{aligned}&\theta \int_l^1 \left[\Pi(1 + s(0) - s(t)) + (s(0) - s(t))\, \pi(1 + s(0) - s(t))\right] \mathrm{d}t \\ +& (1 - \theta) \int_l^1 \left[\Pi(1 + N - s(t))\right] \mathrm{d}t \\ &- C\end{aligned}$$

The expected instantaneous profit depends on time only through entry dates, so the problem is linear. If incumbent finds profitable to delay entry until any $t > l$ she will also want to delay to $t + \epsilon > t$ and so on until patent expiry date. Consequently, the incumbent's problem reduces to selecting an entry schedule

$$s(t) = \begin{cases} s(0) & \text{for all } t \in [0, l] \\ s & \text{for all } t \in (l, 1] \end{cases}$$

that maximizes its expected payoff

$$-C + l\Pi(1) + \theta(1-l)\Pi(1+s(0)-s) + \theta(s(0)-s)\pi(1+s(0)-s) \\ + (1-\theta)(1-l)(\Pi(1+N-s))$$

subject to $s \leq s(0) < N$. Since industry profit is decreasing in the number of firms in the market, $s(0) = s$ is optimal, where s is chosen optimally to maximize $\Pi(1+N-s)$ subject to $s < N$ and thus $s = N-1$. The incumbent thus obtains

$$-C + l\Pi(1) + \theta(1-l)\Pi(1) + (1-\theta)(1-l)\Pi(2)$$

Comparing the incumbent's payoffs then concludes the proof.

Proof of Proposition 8 Suppose $\theta \in \Theta$. Then Proposition 6 implies that there is never licensing and the change in consumer surplus from allowing settlement cannot be positive, entry happens at the same time, but number of entrants is decreased. Suppose $\theta \notin \Theta$. Then Proposition 7 implies that there is no litigation and settlements strictly improve consumer welfare when

$$(1-l)\left[(1-\theta)\underbrace{(\sigma(n)-\sigma(n-s_1))}_{\geq 0} + \theta\underbrace{(\sigma(0)+\sigma(n-s_1))}_{\leq 0}\right] \tag{29}$$

is strictly positive.

Example 2 We consider an economy with three single-product firms and a competitive numeraire sector. There is a continuum of consumers characterized by the same utility function. The utility function is separable and linear in the numeraire good, so there are no income effects, which allows us to perform a partial equilibrium analysis. In our specification the Cournot equilibrium is unique.

In monopoly there is only one firm, therefore there is no parameter γ, inverse demand is given by:

$$p = \alpha - \beta q \text{ for monopoly}$$

When there are two or three firms, differentiation starts to matter,

$$p_i = \alpha - \beta q_i - \gamma q_j \text{ for duopoly}$$

and

$$p_i = \alpha - \beta q_i - \gamma\left(q_j + q_k\right) \text{ for triopoly}$$

Each firm maximizes price times quantity. The Nash equilibrium of the Cournot game(s) is unique. The equilibrium quantities and profits are (set $\Pi = \pi$)

$$q(1) = \frac{\alpha}{2\beta}, \pi(1) = \frac{\alpha^2}{4\beta} \text{ for monopoly}$$

$$q(2) = \frac{\alpha}{\gamma^2 + 2\beta^2}, \pi(2) = \frac{\alpha^2\left(2\beta^2 + \gamma^2 - \beta - \gamma\right)}{\left(2\beta^2 + \gamma^2\right)^2} \text{ for duopoly}$$

$$q(3) = \frac{\alpha(\beta - \gamma)}{2\left(\beta^2 - \gamma\right)}, \pi(3) = \frac{\alpha^2(\beta - \gamma)\left(\beta^2 - 2\gamma + 3\beta\gamma - 2\gamma^2\right)}{4\left(\beta^2 - \gamma\right)^2} \text{ for triopoly}$$

The consumer surpluses are

$$\sigma(1) = \frac{\alpha^2}{8\beta} \text{ for monopoly}$$

$$\sigma(2) = \frac{\alpha^2}{2\beta^2 + \gamma^2} \text{ for duopoly}$$

$$\sigma(3) = \frac{\alpha^2\left(2\beta^3 - 5\beta^2\gamma + [4 - 5\gamma]\gamma^2 + 4\beta\gamma[2\gamma - 1]\right)}{4\left(\beta^2 - \gamma\right)} \text{ for triopoly}$$

Now, using these expressions for consumer surplus, we obtain

$$\tilde{\theta} \equiv \frac{2\beta^3 - 5\beta^2\gamma + [4 - 5\gamma]\gamma^2 + 4\beta\gamma[2\gamma - 1] - \frac{\beta^2 - \gamma}{2\beta^2 + \gamma^2}}{2\beta^3 - 5\beta^2\gamma + [4 - 5\gamma]\gamma^2 + 4\beta\gamma[2\gamma - 1] - \frac{\beta^2 - \gamma}{2\beta}}$$

Proof of Proposition 9 Firstly, suppose that $\hat{s_1}$ is

$$\arg\max_{s_1}\{\theta\Pi(0) + (1 - \theta)\Pi(n - s_1) - s_1(1 + \tau)\pi(n - s_1 + 1)\}$$

and s_1' is

$$\arg\max_{s_1}\{\theta\Pi(0) + (1 - \theta)\Pi(n - s_1) - s_1\pi(n - s_1 + 1)\}$$

, we will claim that $s' \geq \hat{s}$. Because $\hat{s}$ is a maximizer of the first problem we know that:

$$\Pi(n - \hat{s_1}) - \hat{s_1}(1 + \tau)\pi(1 + n - \hat{s_1}) \geq \Pi(n - s_1') - d'(1 + \tau)\pi(1 + n - s_1')$$

Adding and subtracting terms we can see that:

$$0 \geq \Pi(1+n-\hat{s_1})-\hat{s_1}\pi(1+n-\hat{s_1})-\Pi(n-s_1')+s_1'\pi(1+n-s_1') \geq \tau\left[\hat{s_1}\pi(1+n-\hat{s_1})-s_1'\pi(1+n-s_1')\right]$$

,which can only be true if $s_1' \geq \hat{s_1}$. In an analogous way we can prove that number of delayed entrants decreases in equilibrium with litigation.