Pay-for-Delay with Settlement Externalities*

Emil Palikot[†] and Matias Pietola[‡]

May 31, 2019

Abstract

Motivated by recent antitrust cases in the pharmaceutical industry, this article studies the interplay between pay-for-delay settlements, licensing deals and litigation. Our analysis highlights the externalities that they generate: pay-for-delay settlements reduce competition which encourages entry; licensing and litigation make entering less profitable. Faced with multiple entrants, the incumbent exploits these externalities by offering licensing deals to some entrants or by pursuing litigation in order to decrease the cost of delaying contracts offered to others. The number of delayed entrants increases with patent strength. Entrants without pay-for-delay settlements pursue litigation for patents of intermediate strength; otherwise, they receive licensing deals.

*We are grateful to Patrick Rey, Yassine Lefouili, Marc Ivaldi, Doh-Shin Jeon, Bruno Jullien, Klaus Kultti, Massimo Motta, Jorge Padilla, Martin Peitz, Carl Shapiro, Juuso Välimäki, and participants at the 15th IIOC, the 12th CRESSE conference, and 44th EARIE conference, and seminars in the Aalto University and the Toulouse School of Economics, for helpful comments. Matias Pietola gratefully acknowledges financial support from the Yrjö Jahnsson Foundation and the Finnish Cultural Foundation. Emil Palikot gratefully acknowledges support from the European Research Council under the Grant Agreement no. 340903. The working paper version of this article won the AdC Competition Policy Award 2018.

[†]Toulouse School of Economics (TSE), University of Toulouse 1 Capitole, Toulouse, FRANCE, emil.palikot@gmail.com.

[‡]Toulouse School of Economics (TSE), University of Toulouse 1 Capitole, Toulouse, FRANCE, matias.pietola@iki.fi.

1 Introduction

A patent grants its owner the right to exclude potential competitors from the market for a limited period. When it expires, the market opens for entry and competition lowers prices to the benefit of consumers. By determining the scope and duration of patents, the intellectual property system aims to balance incentives to innovate against ex-post consumer surplus. In practice, however, patent offices have limited resources and grant questionable patents, which can be challenged in court.[1] As litigation is costly, the parties often settle out of court.

Patent settlements in the pharmaceutical industry have recently caught the attention of antitrust authorities around the world, with the so-called "pay-for-delay" agreements spurring a particularly heated debate. These are settlements between an incumbent patent holder and a potential entrant: in exchange for financial compensation, the entrant agrees not to challenge the validity of the patent and to stay out of the market for a certain period. Such deals fall at the intersection of antitrust and intellectual property policies. The European Commission considers pay-for-delay agreements to be anti-competitive and has imposed significant fines on companies involved, most notably in *Servier*[2] and *Lundbeck*.[3] The US Supreme Court has instead adopted a rule of reason approach, as in *Actavis*.[4]

However, patent disputes need not result in pay-for-delay settlements. Over the last few years, these agreements have constituted only around 3-12% of all patent settlements in the pharmaceutical sector in the European Economic Area.[5] Licensing agreements, where an entrant buys a license from the incumbent and enters without

[1] Farrell and Shapiro (2008) provide a detailed analysis of the economics of "weak patents".

[2] European Commission decision, 9.7.2014, C(2014) 4955.

[3] European Commission decision, 19.6.2013, C(2013) 3803.

[4] *FTC v. Actavis, Inc.*, 133 S. Ct. 2223, 570 U.S. 136, 186 L. Ed. 2d 343 (2013).

[5] See The Pharmaceutical Sector Inquiry by the European Commission. From 2008, the European Commission has annually monitored patent settlements made by pharmaceutical companies in the EEA area.

delay, are more frequent. Although pay-for-delay agreements and licensing deals often coexist (with different entrants), the anticompetitive effects of the former have been analyzed in isolation. To our knowledge, the question of why incumbents offer different deals to (often similar) entrants has not yet been addressed in the economic literature. This article aims to fill the gap.

We develop a setting with one incumbent and multiple identical, potential entrants. The incumbent owns a patent of uncertain validity and enjoys a legal monopoly unless a court declares the patent invalid. Each entrant can either litigate, wait for the market to open (i.e., for the patent to expire or to be invalidated following litigation by another entrant), or settle with the incumbent. A settlement deal includes a financial transfer and an entry date.

Our analysis highlights the role of the externalities of settlement agreements. An entrant accepting a settlement with a late entry date imposes a positive externality on all other entrants: the expected profit attainable through litigation increases (due to reduced competition), which enhances the bargaining position of entrants negotiating licensing agreements or pay-for-delay deals. In particular, to monopolize the market (i.e., to delay the entry of all potential competitors), the incumbent would have to compensate every entrant for forgoing the duopoly profit that it could achieve through litigation. When there are many entrants, this cost exceeds the gain from maintaining the monopoly position.

Instead of delaying all entrants, to some of them the incumbent may license the patent or take them to court and face the risk of invalidation. Both strategies lower the cost of delaying other entrants, as they face increased competition if they reject their settlement offers and enter through litigation. Licensing and litigation are substitutes for the incumbent and never coincide. Litigation is costly but leads to higher profits if the court upholds the patent.

We show that the number of delayed entrants is higher when litigation costs are high, and the patent is strong. The incumbent may adopt a divide-and-conquer strategy, where it pays some entrants to delay their entry and either grants licenses to the others or fights them in court. Patents of intermediate strength are litigated, whereas sufficiently weak and strong patents are licensed. Furthermore, litigation is more likely to occur when litigation costs are low. In the extreme case, where litigation is costless, it is always the equilibrium outcome provided that competition is not too intense.

Our results contrast with predictions from single-entrant models. When facing only one entrant, the incumbent always delays entry; this strategy maximizes the industry profit (the market remains a monopoly), which is then shared between the parties according to their bargaining powers (Shapiro, 2003). However, with multiple entrants, the cost of offering pay-for-delay settlements to all of them may well exceed the gain from monopolization.

We provide an extension of our model where we allow the incumbent to offer settlements contingent on the validity of the patent. That is, an entrant agreeing to a pay-for-delay deal can nevertheless enter the market if the patent is invalidated, following litigation by another entrant. Therefore, by accepting a settlement, an entrant does not forgo the profit from entry in the case of patent invalidation. In addition, the payoff of a successful litigator decreases, as all competitors enter after the patent is invalidated. Hence, the payments required to delay entry are lower. As a result, the incumbent will delay all entrants in equilibrium, regardless of the strength of the patent and the costs of litigation. Thus, settlements contingent on patent validity hamper competition, rather than promote it.[6]

The possibility of entering into pay-for-delay agreements may also encourage the incumbent to license the patent, which benefits consumers. Because of this interdepen-

[6] In *Servier*, parties subject to a pay-for-delay agreement conditioned on patent validity argued that such a contract is pro-competitive. We show that the exact opposite is true.

dency, the analysis of the welfare consequences of prohibiting pay-for-delay settlements should not look at them in isolation, but account for the decreased incentives to license the patent. To make this point, we provide numerical examples illustrating the possibility that a ban on pay-for-delay settlements may decrease consumer welfare.

Related literature The current economic literature on patent settlements has two branches, one for pay-for-delay agreements (Shapiro, 2003; Meunier and Padilla, 2015; Elhauge and Krueger, 2012; Edlin et al., 2015; Manganelli, 2014; Gratz, 2012) and another for licensing deals (Farrell and Shapiro, 2008; Lemley and Shapiro, 2005; Kamien and Tauman, 1986; Katz and Shapiro, 1987; Amir et al., 2014). To our knowledge, licensing and pay-for-delay agreements have not previously been studied together. In this article, we show that this obfuscates an important economic mechanism, triggered by the settlement externalities between entrants. Because of these externalities, the incumbent may offer different agreements to otherwise similar entrants. This observation relates our work to the literature on contracting with externalities (Segal, 1999, 2003).

Shapiro (2003) introduces the canonical model of pay-for-delay. He considers a framework with a single entrant who may challenge the incumbent patent holder. The two parties have the opportunity to settle and avoid going to court. Generally, they will conclude a pay-for-delay settlement which extends the monopoly period, and divide the resulting high profits. Our approach extends this seminal work by allowing for multiple entrants. This modification reveals the connection between pay-for-delay agreements, licensing, and litigation. In this way, we show that entry may occur in sequence, even though agreements are reached simultaneously with identical entrants.

Pay-for-delay settlements in an environment with multiple entrants have been previously studied by Meunier and Padilla (2015). They focus on the credibility of a litigation threat when a successful litigation by one firm opens the market for all entrants. Thus, entrants who do not pursue litigation effectively free-ride on the litigator. Our results also show the phenomenon of free-riding on litigation efforts, even though all entrants are symmetric and have the ability to litigate (Meunier and Padilla (2015) assume that only one of the entrants can start the litigation). Due to this logic, some patents are too strong to be challenged. We explicitly derive the threshold of patent strength above which entrants will not challenge the patent; however, our focus is on the cases where the threat of litigation is credible.

Shapiro (2003) proposes a general rule for evaluating patent settlements: allowing for settlements should not leave the consumers worse off compared to prohibiting them. Therefore, welfare analysis simplifies to comparing the duration of exclusion resulting from the settlement (the agreed entry date) to the expected entry date when there is no settlement (i.e., entry can occur through expiry or the invalidation of the patent). When the reverse payment is higher than the litigation cost incurred by the incumbent, exclusion due to 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 illegal, regardless of the probability of the patent being invalid. We argue that this logic may fail when there is more than one entrant. We show that the fact of observing a high payment from the incumbent to an entrant does not necessarily imply that consumers are harmed.

Divide-and-conquer strategies have previously been studied in different contexts. Posner et al. (2010) show how a defendant can optimally exploit coordination failure between several plaintiffs. Other related works include Daughety and Reinganum (2002) and Che and Spier (2008). Typically, in these articles some of the plaintiffs are offered beneficial treatments and decide to settle with the defendant, which leads to the others dropping their lawsuits. In our case, the incumbent also exploits a coordination failure of the entrants, but the context and the modeling approach differs from the previous literature.

The article is organized as follows. The next section is devoted to case studies in the pharmaceutical industry. We discuss the context of pay-for-delay settlements and the antitrust response to them. Section 3 introduces the model. Section 4 presents an extension of our model, by allowing the incumbent to offer contracts contingent on the validity of the patent. Section 5 shows how a ban on pay-for-delay agreements changes the equilibrium outcome, and provides some policy implications, and Section 6 concludes. The Appendix contains all proofs.

2 Pay-for-delay cases in the pharmaceutical industry

The competitive landscape in the pharmaceutical sector is, to a large extent, shaped by two factors: the patent protection of newly developed drugs, and entry by producers of generic bio-equivalent medicines. Once the primary patent protecting the main chemical compound has expired, the producers of generics can simultaneously contest the originator's monopolistic position. From that date on, (usually) several generic producers race to be the first in the market. Often, despite the expiry of the main patent, the legal situation is, however, unclear: the originator has applied for secondary patent protection and sometimes holds patents on alternative methods of manufacturing the medicine. This legal uncertainty can result in a dispute over the validity of one of the patents that, in many cases, ends in a court of law. In this section, we briefly review three important cases involving pay-for-delay settlements; the first two are the European *Servier* and *Lundbeck* cases, and the last is a US case: *Actavis*.

The *Lundbeck* case considers several agreements between a Danish pharmaceutical company, Lundbeck, and several producers of generic drugs. In the 1970s and 1980s, Lundbeck developed the antidepressant drug Citalopram that was first marketed in the

1990s. The medicine was very successful and became the main product of Lundbeck; for example, it constituted 80-90% of the company's revenues in 2002. At the time of the settlements, during 2002 and 2003, patents related to the chemical compound and the original process had expired. In principle, the market was therefore 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. In order to limit the extent of market entry by generic producers, Lundbeck implemented a so-called "generic strategy" that involved different kinds of agreements with several entrants. Overall, the "generic strategy" was a mixture of pay-for-delay settlements, takeovers, licensing, accommodation, and even an 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. In Iceland, Lundbeck allowed a market entry without litigation. In June 2013, the European Commission ruled that the part of the "generic strategy" consisting of settlements delaying market entry was a violation of the European antitrust law, as its objective was to reduce potential competition. Lundbeck received a fine of EUR 93.8 million, and the generic producers involved were fined EUR 52.2 million.

The *Servier* case involves a French pharmaceutical manufacturer, Servier, and generic producers of Perindopril, a medicine for treating high blood pressure developed by Servier in the 1980s. Perindopril became Servier's most successful product with annual global sales exceeding USD 1 billion in 2006 and 2007, 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 started designing and implementing "the generic strategy" from the late 1990s. This strategy included acquiring new patents and resulted in five settlement agreements with different generic producers between 2005 and 2007. Four of these agreements were pay-for-delay settlements, whereas the fifth was a licensing deal. Servier considered litigation and licensing as alternative strategies to accommodate one of the challengers. In July 2014, the European Commission imposed fines totaling EUR 427 million on Servier and other pharmaceutical companies involved in the pay-for-delay settlements.

The Servier and Lundbeck decisions were both appealed before the EU General Court, which gave its judgment on the Servier case in December 2018. For the most part, the court upheld the Commission's decision, except for the Krka (a manufacturer of a generic version of Perindopril) settlement, for which the fine was annulled. The court made a clear distinction between side-deals in general and licensing agreements between Servier and Krka. Most interestingly, the court stated that licensing agreements should not be considered as suspicious side-deals, but as an appropriate means of settling a patent dispute: a licensing agreement stems from the parties' mutual understanding about the validity of the patent, making the settlement possible.[7]

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. This legislation aims to promote the entry of generics by guaranteeing the first of them a duopolistic position. When a producer of generics files an Abbreviated New Drug Application (ANDA) to the Food and Drug Administration (FDA), the Hatch-Waxman Act requires the declaration of a relationship to a patent mentioned in the Orange Book.[8] If a generic producer states that the relevant patents are no longer valid, or that they are not infringed, the certification is granted. Importantly, the Hatch-Waxman Act provides 180 days of exclusivity to the first entrant: no other producer of generics 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.

[7] Case T-691/14, Servier v. Commission, EU:T:2018:922, paras 943-963.

[8] A list of approved drug products together with a catalog of patents related to each of them.

Solvay was granted a new patent for AndroGel in 2003. Later, Actavis filed an ANDA to the FDA and stated that Solvay's new patent was invalid and thus the generic version produced by Actavis could 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 was 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 a decision; (2) the patent holder was likely to have enough power to implement antitrust harm in practice; (3) the antitrust action was likely to prove more feasible administratively than the Eleventh Circuit believed: a large, unexplained payment from the patent holder to the generic producer could provide a workable surrogate for a patent's weakness; and (4) the parties could have made another type of settlement agreement by allowing the generic producer to enter the market before the patent expired, without the need for a reverse payment.

Several valuable lessons can be drawn from the case law on pay-for-delay agreements. First of all, even a weak patent can be useful for the incumbent, because of high litigation costs and free-riding between the generic companies in their litigation efforts. Patent invalidation opens the market for everybody, not merely for the generic producer who took the litigation effort and, incurred an often 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.[9] The incentives to settle litigation are particularly pronounced when other generic producers wait to enter the market.

Second, 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 a 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 undertake laboratory tests and seek third-party advice.[10]

Third, the entry game starts after the expiry of a certain patent, and this date is common knowledge. Firms, wishing to enter the market arrive 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, whereas the others are delayed. So-called "generic strategies" are typically combinations of pay-for-delay agreements, licensing deals, litigation, and takeovers.

3 Model

Consider a market with one incumbent firm, I, and $N \geq 2$ symmetric entrants, E_1, E_2, ..., E_N. I owns a patent and enjoys a legal monopoly until the patent expires, unless at least one E_i, for $i \in \{1, 2, \ldots, N\}$, litigates and a court declares the patent invalid. Litigation costs $c \geq 0$ to E_i and $C \geq 0$ to I. It is common knowledge that if any E_i litigates, the court will declare the patent invalid with probability $1 - \theta$ and uphold the patent otherwise; θ thus reflects the strength of the patent. We assume that, if several entrants litigate, the litigation outcomes are perfectly correlated.[11]

The patent starts on date $t = 0$, and time runs continuously until the patent expires

[9] See paragraph 493 of the Servier decision by the European Commission.

[10] See paragraph 709 of the Servier decision and 522 of the Lundbeck decision.

[11] In practice, courts often bundle similar cases.

at time $t = 1$, after which free entry drives all profits down to zero. At date zero, the firms play a two-stage negotiation game:

- **Stage 1**: I offers each E_i a settlement deal, which includes an entry date $t_i \in [0, 1]$ and a payment $p_i \in \mathbb{R}$ from E_i to I. The offers are public and are observed by all firms.

- **Stage 2**: Each E_i either accepts the settlement offer, rejects the offer and litigates, or rejects the offer and waits for the market to open.

Litigation is time-consuming: it takes the court $1 - l$ time to make a decision; thus, after the court rules on patent validity there is $l > 0$ time left before the initial patent expiry date. If the court declares the patent invalid, those entrants who have rejected the settlement offer (either litigate or wait) enter the market, whereas those who have signed a settlement deal, are bound by it.[12] If instead the court upholds the patent, entry is possible only through a settlement. This is also the case during the litigation period.[13]

At any point in time, if n firms are active in the market (I and $n - 1$ entrants), I makes a profit $\Pi(n)$ and each active E_i makes a profit $\pi(n)$, whereas the entrants that stay out make zero profit.[14] We assume that the profits are all positive and that total industry profit decreases in n. All firms are risk-neutral, and their payoffs are the sums of payments and profits. The equilibrium concept is subgame perfect Nash equilibrium in pure strategies and we use backward induction to solve the game.

[12] This is in line with the legal principle of *pacta sunt servanda*. It is the ex-ante view of the strength of the patent that matters for reaching a settlement before the court, not the validity of the patent resolved ex-post. Later on we consider entry delay contingent on patent validity.

[13] We thus do not allow for entry at risk, which would open the question about the appropriate damage rule if the patent is valid.

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

We make two key assumptions. First, we assume that the monopoly profit is not too high:

Assumption 1. *There exists* $x \in \{0, \ldots, N-1\}$ *such that*

$$\Pi (1 + N - x) - x\pi (2 + N - x) > \Pi (1) - N\pi (2) .$$

The purpose of this assumption is to rule out the trivial case when I always finds it profitable to delay all entry until patent expiration, even if the patent is invalid with certainty and litigation is costless. The assumption is satisfied by the symmetric Cournot quantity-setting game with a linear demand, for example, and generally satisfied for N large enough.

Second, we assume that the litigation threat is credible:

Assumption 2. *The expected payoff from litigation is at least zero:*

$$(1 - \theta) l\pi (1 + N) - c \geq 0.$$

This assumption is necessary to make the game interesting, as I would otherwise simply make unacceptable settlement offers to all entrants and monopolize the market at no cost.

Although, I can offer settlements with any entry date $t_i \in [0, 1]$ and payment $p_i \in \mathbb{R}$, only two types of settlement agreements are signed in equilibrium (see the Appendix for the formal proof):

- Licensing agreements: E_i enters at date $t_i = 1 - l$ and pays a licensing fee $p_i > 0$;

- Pay-for-delay agreements: E_i delays entry until patent expiry $t_i = 1$ and receives a reverse payment $-p_i > 0$.

There are two reasons for this. First, entry before the court decides on the validity of the patent is not possible without a settlement. Thus, as the industry profits are highest under monopoly, I never finds it profitable to license during the litigation period. Second, we do not allow for contracts with exit, implying that payoffs are time-independent. Hence, if licensing is profitable for I at some point, it is already profitable at the moment when the litigation period ends.

To characterize the equilibrium of the negotiation game, it is useful to define two functions: $X : [0,1] \longrightarrow \{0,1,2,\ldots,N\}$ and $f : [0,1] \longrightarrow \mathbb{R}$, where $X(\theta)$ is a function of the strength of the patent, and it maximizes:

$$\begin{aligned} g(x;\theta) &:= \theta\left[(N-x)\,\pi(1+N-x) + x\pi(2+N-x)\right] \\ &\quad + \Pi(1+N-x) - x\pi(2+N-x)\,. \end{aligned} \tag{1}$$

As it turns out, $X(\theta)$ determines the number of delayed entrants in equilibrium.

$$f(\theta) := \theta\Pi(1) + (1-\theta)\,g(X(0);0) - g(X(\theta);\theta)\,, \tag{2}$$

is the difference in maximum profits I can attain by pursuing litigation and by settling with all entrants, net of total effective litigation costs λ, where:

$$\lambda := \frac{C+Nc}{l}. \tag{3}$$

$f(\theta) = \lambda$ makes I indifferent between N settlements and litigation. We are now ready to state our first result, which characterizes the equilibrium of the negotiation game, fixing the strength of the patent, and varying λ.

Proposition 1. *There exists an essentially unique equilibrium of the negotiation game:*

- *If litigation costs are high, $\lambda > f(\theta)$, then there is no litigation: $X(\theta)$ entrants are delayed and $N - X(\theta)$ buy a license;*

- *If litigation costs are low, $\lambda < f(\theta)$, then there is litigation: one entrant litigates, $X(0)$ entrants are delayed and $N - 1 - X(0)$ wait.*

Furthermore, $X(\cdot)$ is weakly increasing and satisfies $X(1) = N$.

As the litigation outcomes are perfectly correlated, at most one entrant litigates in equilibrium (due to symmetry, the identity of the entrant does not matter for I). Furthermore, as the litigation threat is credible, waiting is never a best response if no rival entrant litigates. Thus, in any equilibrium, either one and only one entrant litigates, or all entrants settle.

Whether litigation happens in the equilibrium depends on how high the litigation costs are relative to profits. For high litigation costs, $\lambda > f(\theta)$, all entrants settle. To delay x entrants, I must compensate each of them for payoff they would obtain by rejecting the settlement deal and going to court instead. The cost of entry delay is therefore:

$$x \cdot \underbrace{[(1-\theta)\, l\pi\,(2 + N - x) - c]}_{\text{reverse payment}}. \tag{4}$$

Furthermore, the highest possible licensing fee each licensee is willing to pay corresponds to the difference between the profit they make in the market and their payoff from litigation. Thus, the licensing revenue is:

$$(N - x) \cdot \underbrace{[\theta l\pi\,(1 + N - x) + c]}_{\text{licensing fee}}. \tag{5}$$

Reverse payments and licensing fees are increasing in the number of delayed entrants. Each entrant that accepts a delaying settlement imposes a positive externality on all other entrants: by agreeing to stay out of the market, a delayed entrant increases the expected profits from litigation to all other entrants, which in consequence increases settlement payoffs needed to compensate for withdrawing from litigation. Licensing fees are increasing in the number of delayed entrants, because profits in the market are higher when more entrants are delayed. Licensing agreements have an opposite effect: each entrant that accepts a licensing agreement imposes a negative externality on all other entrants; there is increased competition following a successful litigation. By accepting a settlement and forgoing litigation each entrants saves on the litigation cost; however, I accounts for it in settlement payments; thus, effectively I extracts all savings from forgone litigation costs.

In addition to the payments from the settlements, from the market I makes the profit:

$$\underbrace{(1-l)\,\Pi\,(1)}_{\text{litigation period}} + \underbrace{l\Pi\,(1+N-x)}_{\text{own profit under licensing}}. \tag{6}$$

Hence, putting everything together, I's payoff from delaying x entrants and licensing the patent to the rest writes:

$$(1-l)\,\Pi\,(1) + lg\,(x;\theta) + Nc, \tag{7}$$

which is maximized at $X\,(\theta)$.

The number of delayed entrants is weakly increasing in the strength of the patent. First, the cost of entry delay is decreasing in the strength of the patent. I has to compensate delayed entrants for withdrawing from litigation; the stronger the patent, the lower the expected profit from starting a litigation, and hence, the lower the necessary reverse payments. Second, industry profits are decreasing in the number of firms in the market. Thus, if the cost of entry delay decreases sufficiently, I always has an incentive to delay another entrant.

For low litigation costs, $\lambda < f(\theta)$, there is litigation in the equilibrium. In this case, to delay x entrants, I must compensate each of them for the payoff they would obtain by waiting and free-riding on the litigation costs taken by the entrant who went to court. The cost of entry delay writes:

$$x \cdot \underbrace{[(1-\theta)\, l\pi\, (2+N-x)]}_{\text{reverse payment}}. \tag{8}$$

In principle, entrants who are not delayed could either obtain a license or wait for the market to open through litigation. However, when litigation is already ongoing, I is always better off by making no deal than licensing the patent. To see this, suppose I offers a licensing deal to $y \leq N - x - 1$ entrants, yielding the licensing revenue:

$$y \cdot \underbrace{[\theta l\pi\, (1+y)]}_{\text{licensing fee}}. \tag{9}$$

Furthermore, from the market I would make the expected profit:

$$\underbrace{(1-l)\,\Pi\,(1)}_{\text{litigation period}} + \theta \cdot \underbrace{l\Pi\,(1+y)}_{\text{own profit if wins in court}} + (1-\theta) \cdot \underbrace{l\Pi\,(1+N-x)}_{\text{own profit if loses in court}}. \tag{10}$$

The number of licensees only affects the licensing revenue and I's own profit. As the industry profit is decreasing in the number of firms, the optimal number of licensees is zero: $y = 0$, which allows I to maintain a monopoly position if the court upholds the patent.

Generally, I takes entrants to court or licenses the patent in order to decrease the cost of entry delay. The cost of delaying entry equals compensation to all delayed entrants for withdrawing from litigation. Litigation brings them profits only if the court invalidates the patent. Therefore, the fact that licensing leads to entry for sure (unlike litigation) does not reduce the cost of delay, but decreases I profit for sure, rather than with probability $1-\theta$. Hence, if litigation is already underway, I will not offer any licensing contracts.

However, when there is litigation I may still find it profitable to delay some of the entrants. Considering the payments and profits together, I's payoff from litigation is:

$$(1-l+\theta l)\,\Pi(1)+(1-\theta)\,lg(x;0)-C, \tag{11}$$

which is maximized at $X(0)$.[15]

Finally, the difference between the highest profits I can obtain with and without litigation, net of the effective litigation costs, is:

$$f(\theta)=\theta\Pi(1)+(1-\theta)\,g(X(0);0)-g(X(\theta);\theta)\,. \tag{12}$$

Therefore, the difference between $f(\theta)$ and costs λ determines the equilibrium. If litigation is costly, $\lambda > f(\theta)$, the settlement payoff exceeds the payoff from litigation, and if the reverse inequality holds, I prefers to litigate.

By observing that $g(X(\theta);\theta)$ is convex, being the upper envelope of affine functions, we deduce that $f(\theta)$ is concave and continuous. Furthermore, it clearly satisfies $f(0)=f(1)=0$. Hence, we obtain our next result, which characterizes the equilibrium of the game, fixing the costs of litigation, and varying θ.

Proposition 2. *For any $\lambda < \max f$ there exist thresholds of patent strength, $\underline{\theta}(\lambda) < \overline{\theta}(\lambda)$, where $\underline{\theta}(\lambda)$ satisfies $\underline{\theta}(0)=0$ and increases in λ, and $\overline{\theta}(\lambda)$ satisfies $\overline{\theta}(0)=1$ and decreases in λ, such that, in equilibrium:*

- *Patents of intermediate strength, $\theta \in \left[\underline{\theta}(\lambda), \overline{\theta}(\lambda)\right]$, are litigated;*

[15] Note that $X(0) \leq N-1$ by Assumption 1.

- *Patents that are sufficiently strong, $\theta > \overline{\theta}(\lambda)$, or weak, $\theta < \underline{\theta}(\lambda)$, are not taken to court: all entrants settle.*

Going to court is costly but gives I a chance of monopolizing the market without paying the entrants. If the patent is strong the entrants, who are likely to lose in court are willing to accept pay-for-delay agreements with small reverse payments. I then prefers to avoid costly litigation and to delay entrants. If instead the patent is weak, I's chance of monopolizing the market through litigation is small, and the entrants have a strong bargaining position. Therefore, for a weak enough patent, I offers licensing deals to save on the costs of litigation.

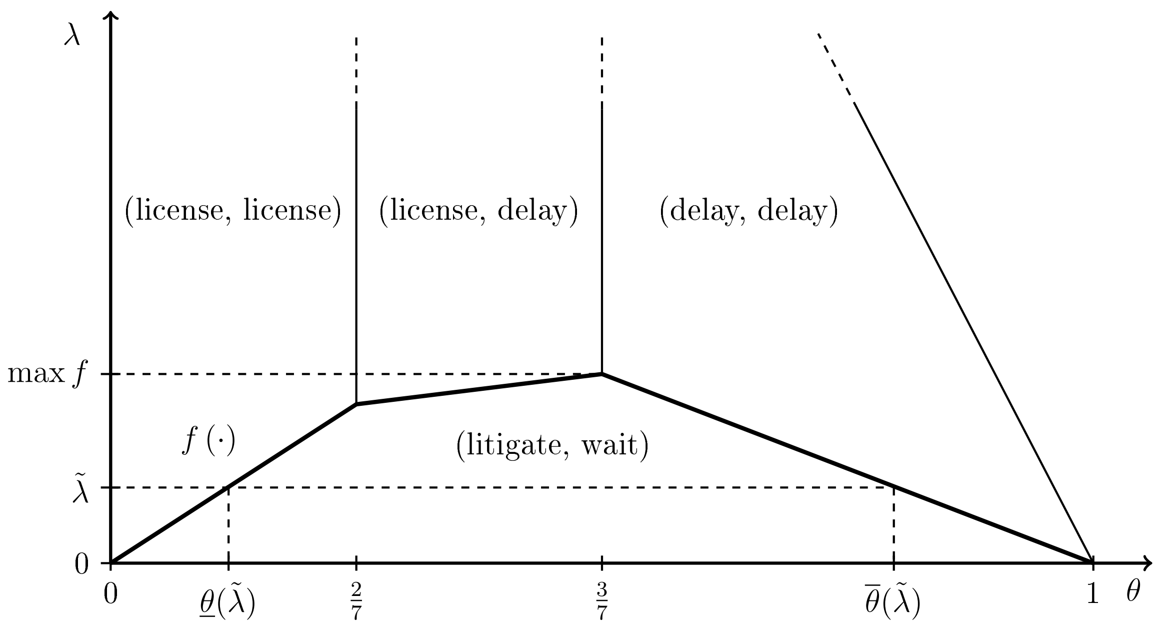

Figure 1: Equilibrium of the negotiation game with two entrants, as a function of the strength of the patent, θ, and the total cost of litigation, λ. Profits as in the symmetric Cournot quantity-setting game with an inverse demand $p = 1 - Q$. Details of this example are in the Appendix.

For patents of intermediate strength, I has a real chance of monopolizing the market through litigation, whereas delaying entry is not cheap. It will then take its chances and pursue litigation, unless the litigation costs are too high, in which case it is better to play "divide-and-conquer" by offering a licensing deal to some of the entrants while delaying the other ones.

Figure 1 uses a parametrized Cournot quantity setting game with two entrants to depict the outcome of the negotiation game as a function of the strength of the patent and the costs of litigation. For any level of costs λ we can determine the interval $\left[\underline{\theta}(\lambda), \overline{\theta}(\lambda)\right]$ of patent strength, where there is litigation in the equilibrium. For zero litigation costs, going to court is always the profit-maximizing strategy for the incumbent, and for sufficiently high litigation costs, there is never litigation.

Remark 1. Identical entrants have different equilibrium payoffs when I treats them differently. If there is litigation in equilibrium, the entrants who wait obtain the highest payoff, $(1-\theta)\, l\pi\,(1+N-X(0))$, whereas the litigator gets the same but has to pay for the litigation cost. The delayed entrants get reverse payments equal to $(1-\theta)\, l\pi\,(2+N-X(0))$, because I has to compensate them for the payoff they *would* obtain by rejecting the settlement deal. Depending on how costly litigation is, the delayed entrants are better off than the litigator or vice versa. If instead there is no litigation in equilibrium, each delayed entrant obtains $(1-\theta)\, l\pi\,(2+N-X(\theta))-c$, whereas licensees receive $(1-\theta)\, l\pi\,(1+N-X(\theta))-c$. Thus, licensees are always better off than delayed entrants.

4 Conditional settlements

The legal principle of *pacta sunt servanda* states that a contract should stay in power despite an *expected* change of environment. Following this principle, our analysis to this point has assumed that a settlement stays in force even when a court declares the patent invalid. Indeed, when parties negotiate over a settlement to a patent dispute, they have a belief about the strength of the patent, and this influences the terms of the settlement. If the patent is later declared invalid by a court, following litigation by a third party, the parties to the contract have been aware of this risk when agreeing to the settlement. It is the ex-ante view of the strength of the patent that matters in reaching a settlement, not the validity of the patent resolved ex-post.

Even so, the parties could explicitly formulate a settlement agreement conditional on the validity of the patent. Allowing for such deals completely changes the equilibrium of the game:

Proposition 3. *The negotiation game with conditional settlement terms has a unique equilibrium, where all entrants are delayed until patent expiration, regardless of the strength of the patent and the costs of litigation.*

Conditioning pay-for-delay settlements on the validity of the patent decreases the payoff from litigation. If the court invalidates the patent, the litigating entrant has to compete with all delayed entrants; thus the incentives to start the litigation are lower. As a result, the cost of entry delay reduces to:

$$x \cdot \underbrace{[(1-\theta)\, l\pi\, (1+N) - c]}_{\text{reverse payment (conditional deals)}} \leq x \cdot \underbrace{[(1-\theta)\, l\pi\, (1+N-x) - c]}_{\text{reverse payment (regular deals)}}. \tag{13}$$

We have argued that the key mechanism forcing I to allow for some entry is the positive externality imposed by an entrant accepting a delaying settlement on other entrants' payoffs from litigation. Pay-for-delay settlements conditioned on the validity of the patent do not create such an externality, because they are no longer binding when a successful litigator enters the market.

In *Servier*, for example, firms argued that such conditional pay-for-delay agreements are less harmful to consumers because the delayed entrant can enter the market if the patent is declared invalid. This logic, however, is misleading. By using conditional settlements, the incumbent is able to reduce the cost of entry delay to the extent that it is always profitable to delay all entrants until patent expiration. We thus have a clear policy implication: conditional settlement terms hamper competition by reducing the number of firms in the market, and thus from an antitrust policy perspective, they should not be allowed.

5 Banning pay-for-delay agreements

In this section, we look at the implications of a ban on pay-for-delay agreements on the equilibrium of the negotiation game. One could argue that from the perspective of the patent system, the outcome, as illustrated by Figure 1, is reassuring. First, strong patents are not challenged in court and result in a monopoly, which is the point of issuing patents: a patent grants its owner the right to exclude rivals from the market, thus, it would be a waste of resources if firms were litigating over the validity of *ex-ante* strong patents. Second, weak patents are licensed; therefore, they do not prevent entry to the market, which benefits consumers. Third, patents of intermediate strength are litigated, correcting the legal uncertainty created by the imperfect screening of patents.

A ban on pay-for-delay agreements alters this picture significantly. To see how the equilibrium of the negotiation game changes, define:

$$f_0(\theta) := \theta \left[\Pi(1) - \Pi(1+N) - N\pi(1+N)\right], \tag{14}$$

which is simply the function $f(\theta)$ under the constraint $X = 0$.

Proposition 4. *Suppose pay-for-delay agreements are banned. Then, there exists an essentially unique equilibrium of the game:*

- *If litigation is costly, $\lambda > f_0(\theta)$, then there is no litigation: all entrants buy a license;*

- *If litigation is not costly, $\lambda < f_0(\theta)$, then there is litigation: one entrant litigates while the others wait.*

Furthermore, for any $\lambda < f_0(1)$, there exists a threshold of patent strength, $\theta_0(\lambda)$, such that there is litigation if the patent is sufficiently strong, $\theta > \theta_0(\lambda)$, and no litigation otherwise.

When pay-for-delay agreements are not allowed, the only way I can avoid litigation is to license the patent to all entrants. As a consequence, the space for litigation increases, because this is the only way for the incumbent to monopolize the market. For strong patents, signing pay-for-delay agreements would save on the costs of litigation, while still resulting in the same market outcome with a high probability. Furthermore, the scope of licensing increases, because offering all entrants a license is the only way the incumbent can avoid litigation, if litigation is too expensive. This means that even strong patents are licensed, when litigation costs are high. Figure 2 illustrates the outcome of the negotiation game when pay-for-delay agreements are illegal and thus outside the incumbent's toolkit.

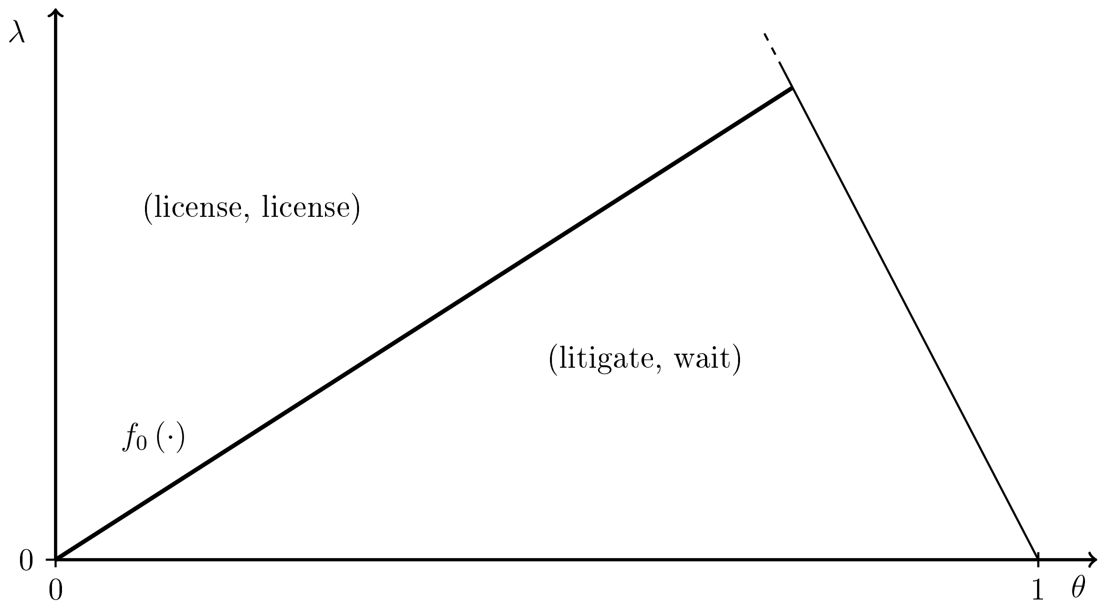

Figure 2: The equilibrium of the negotiation game with two entrants when pay-for-delay agreements are banned, as a function of the strength of the patent, θ, and the total cost of litigation, λ.

Consumer welfare in a market governed by a patent depends on the date of entry: licensing agreements result in earlier entry compared to litigation, and litigation leads to (expected) more competition than pay-for-delay. As we argue in this article, the possibility of entering into pay-for-delay agreements may encourage the incumbent to license the patent. Therefore, in some cases consumers might be worse-off from a ban on pay-for-delay settlements; this is when the outcome under the *laissez faire* rule involves licensing, and under the ban no licensing happens in the equilibrium. However, if all entrants are delayed, prohibiting pay-for-delay settlements would improve consumer welfare. Figure 3 illustrates how a ban on pay-for-delay agreements influences the consumer surplus ex-post in our parametrized example.

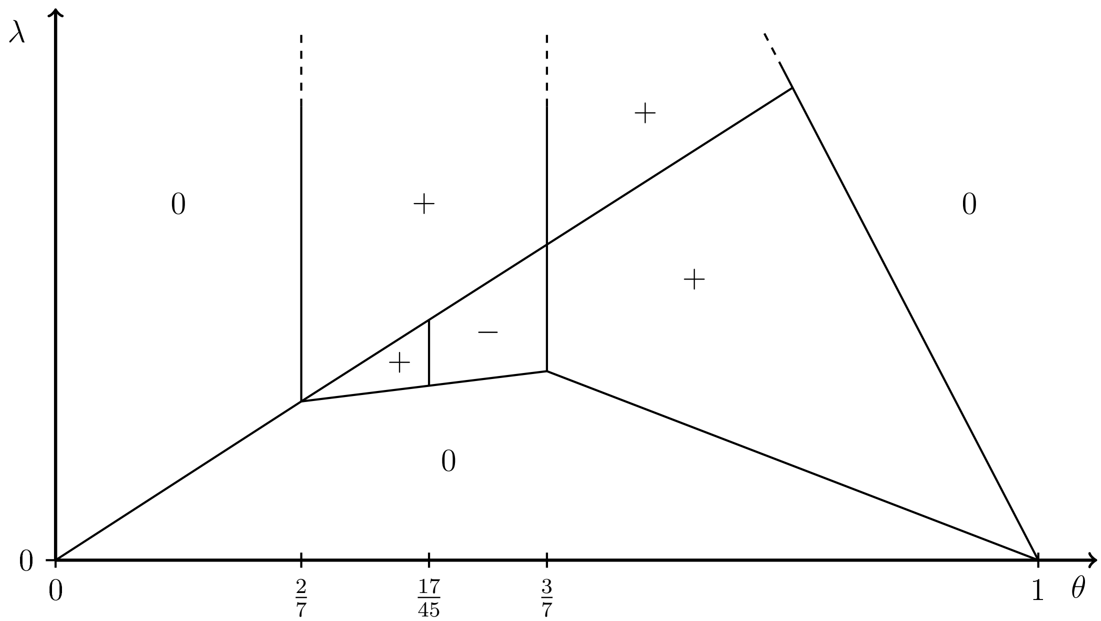

Figure 3: The change in expected consumer surplus (ex-post, not considering the incentives to innovate) following a ban on pay-for-delay agreements.

In the area with the negative sign the incumbent stops licensing the patent and goes to court instead, although the expected consumer surplus from litigation is lower than the one in duopoly, which would prevail if pay-for-delay agreements were allowed. This highlights a more general point: due to the interdependency between pay-for-delay and licensing incentives, the analysis of the welfare consequences of prohibiting pay-for-delay settlements should not look at them in isolation, but account for the decreased incentives to license the patent.

6 Concluding remarks

In a market covered by a patent, consumer welfare depends largely on the date of entry of competing firms. Therefore, settlements of patent litigation aimed at delaying or accelerating entry have a severe economic impact. We propose a model of patent settlements in an environment with multiple entrants. This approach allows us to study an important phenomenon which is new to the literature on pay-for-delay: settlement externalities. A settlement defining the date of entry of an entrant impacts incentives to enter the market by other firms. If a settlement leads to a delayed entry it incentivizes all other potential entrants to challenge the patent holder and enter the market; on the contrary, if a settlement is a licensing agreement, it discourages entry by others.

Economic literature studying pay-for-delay agreements has to date, focused either on the entry of a single firm (Shapiro, 2003) or on sequential entry (Gratz, 2012). It has been shown that an incumbent patent holder can agree with a potential entrant to delay its entry with a reverse payment (Shapiro, 2003). Allowing for multiple *ex-ante* identical entrants highlights an important externality of such an agreement, as it makes entering the market more attractive to all other potential entrants. In order to preserve its monopoly position, the incumbent has to offer expected duopoly profits to all potential entrants, as these are attainable by rejecting the settlement offer and pursuing litigation instead. For sufficiently many firms the patent holder will allow some entry.

We show that more pay-for-delay settlements are concluded when the patent is strong and the litigation costs are high. When delaying all entrants becomes too expensive, the incumbent will implement a more complicated strategy: divide-and-conquer. To decrease the cost of pay-for-delay settlements, the incumbent will allow some entry to the market, either through licensing or litigation. Thus, in the equilibrium, entrants receive different payoffs despite being identical. Furthermore, we show that sequential entry to the market can be an outcome of a negotiation game between an incumbent and several entrants arriving simultaneously. We, also, find that litigation occurs for patents of intermediate strength.

This article contributes to the debate on the economic consequences of pay-for-delay agreements. First, we prove that settlements which are conditional on patent validity should not be allowed if a policy-maker's goal is to promote entry to the market. When a settlement is only binding as long as the patent is valid, the positive externality caused by pay-for-delay settlements to the payoff from litigation disappears. Under conditional settlements, a firm entering the market through litigation faces competition from all entrants who have accepted delaying settlements, and achieves lower profits. As a consequence, reverse payments associated with pay-for-delay agreements are much lower, and the incumbent delays all entrants.

Second, by analyzing how the equilibrium of the game changes when pay-for-delay settlements are prohibited, we contribute to the discussion on the welfare consequences of pay-for-delay settlements. Previous literature studying models with one entrant provides useful, information-light rules to guide antitrust enforcement with respect to pay-for-delay settlements (Shapiro, 2003; Elhauge and Krueger, 2012). Unfortunately, in an environment with multiple entrants, the settlement date of entry or the size of payment are not sufficient statistics to determine the welfare consequences of concluded settlements. The interdependency of settlements due to the externalities discussed in this article means that such threshold rules might deem welfare-improving conduct as anti-competitive. In our setting, the *laissez-faire* approach to pay-for-delay settlements might improve consumer welfare when sufficiently many firms receive licenses.

A particular feature of the pharmaceutical market in the US is the Hatch-Waxman Act. This legislation aims to promote the entry of generics by guaranteeing the first entrant a duopoly position. In light of our results, such a policy should not be effective. Once exclusivity to the first entrant is granted, the incumbent will delay its entry; there are no settlement externalities because the other entrants are excluded from the market by law. However, a comprehensive study of such legislation should account for incentives to innovate by generics, which was outside the scope of our analysis.

We have explicitly avoided introducing asymmetric information. Beliefs about the strength of the patent are critical factors in agreeing on a settlement. Settlement offers in a game of incomplete information signal the strength of the patent, and the incumbent would have to account for the impact of offers on the beliefs. In a setting with multiple entrants, this could lead to complex settlement strategies, thus, careful analysis of such a signaling game is an interesting avenue for future research.

Finally, another valuable extension would be to introduce entry-at-risk. For example, in *Servier*, one of the entrants decided to launch its product before the resolution of the patent dispute. The incentive to "enter at risk" would be shaped by the possibility of obtaining an injunction and by the damage rule applied by the court. Determining conditions under which entry-at-risk could occur, and its welfare consequences is an important extension to the current analysis.

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.

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 Proposition 1

To prove Proposition 1 we will proceed in the following steps: first, we show that in the equilibrium of the continuation game, either one entrant litigates or all entrant settle. Second, we derive optimal settlement offers and I's payoff when all entrants settle. Third, we solve the game when one entrant litigates. Fourth, the condition for the equilibrium with litigation to arise is shown. Finally, we prove the monotonicity result.

Two types of equilibria: with and without litigation

In any equilibrium of the continuation game, either one and only one entrant litigates or all entrants settle.

Due to perfectly correlated litigation outcomes, if some E_j litigates in the continuation game, the best response of E_i is to either to accept the settlement offer or to reject it and wait, because waiting saves on the litigation cost but otherwise gives the same payoff as litigation. If instead none of the rival entrants litigates, the best response of E_i is to either to accept the settlement offer or to reject it and litigate because the litigation payoff is positive by Assumption 2, whereas the payoff from waiting is zero.

We can thus cover all potential equilibria by first analyzing the ones where all entrants settle and then the ones where one entrant litigates, and the others either wait or settle.

It is convenient to denote $s(t)$ as the number of entrants who settle with an entry date t or later. At any date t, we can then categorize the entrants into three groups:

- $s(t)$ delayed entrants,

- $s(0) - s(t)$ entrants who settle and enter before t,
- $N - s(0)$ entrants who reject the settlement offer and either litigate or wait.

Equilibria without litigation

Suppose first that all entrants settle in equilibrium. Then, by definition, in the continuation game each E_i has a best response to settle given that all rival entrants settle. Thus, the payoff E_i obtains from accepting the settlement offer,

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

is no less than the payoff E_i *would* obtain by rejecting the offer and going to court instead:

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

As it is not profitable for I to leave E_i strictly better off from the settlement, the equilibrium payment $p_i(t_i)$ satisfies

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

Note that, given the offers $\{p_i(t_i), t_i\}$ there exists a unique continuation equilibrium where all entrants accept the settlement offer. In any other equilibrium, as shown above, one entrant would litigate, whereas the others would either wait or settle. However, the litigator would have a best response to accept the settlement deal instead.

Importantly, the pay-off from the settlement relative to the pay-off from litigation weakly increases in the number of entrants who reject the settlement offer and wait instead. Thus, if it is optimal for an entrant to accept the settlement when all other entrants accept, it is also optimal to accept, when some entrants reject the settlement.

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

Furthermore,

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

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

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

The payoff of I is the sum of the profits it makes in the market and the total payment. During the litigation period I has the possibility to obtain the entire industry profit:

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

Industry profit is maximized under monopoly, thus I has no incentive to allow for entry, so $s(t) = N$ for all $t \leq 1-l$. After the litigation period, I's profit depends on the outcome of the litigation. Adding the payments we have:

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

Notice that the expected instantaneous profit depends on time only through entry dates, so the problem is linear. If the incumbent finds it profitable to delay entry until any $t > 1-l$, it will also want to delay to $t+\epsilon > t$ and so on until patent expiry date. Consequently,

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

where $X\left(\theta\right)$ maximizes

$$\begin{aligned}
g\left(x;\theta\right) = \theta &\left[\left(N-x\right)\pi\left(1+N-x\right) + x\pi\left(2+N-x\right)\right] \\
&+ \Pi\left(1+N-x\right) - x\pi\left(2+N-x\right).
\end{aligned}$$

subject to $x \leq N$. The payoff in equilibrium without litigation is then

$$S\left(\theta\right) \coloneqq \left(1-l\right)\Pi\left(1\right) + lg\left(X\left(\theta\right);\theta\right) + Nc.$$

Equilibria with litigation

Suppose now that one entrant litigates, $s\left(0\right) \leq N-1$ entrants settle and $N-1-s\left(0\right)$ entrants wait in equilibrium. By offering an entrant a negative payoff from the settlement (for example a strictly positive payment with delay until patent expiration),

I can be sure that the entrant rejects the deal and either litigates or waits in the continuation equilibrium. Therefore, it suffices to consider the best responses of those who settle. As one entrant litigates, by accepting the settlement offer, E_i gets

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

which is no less than the payoff E_i *would* obtain by rejecting the offer and waiting instead:

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

Again, as it is not profitable for I to leave E_i strictly better off from the settlement, the equilibrium payment $p_i\left(t_i\right)$ satisfies

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

Observe that, by the same logic as in the previous subsection, the offers $\{p_i\left(t_i\right), t_i\}$ imply that there exists a unique continuation equilibrium where $s\left(0\right)$ entrants accept the settlement offer. If some of them were to reject and wait instead, the payoffs from accepting the settlement deal can only increase relative to the payoffs from waiting.

Differentiating $p_i\left(t_i\right)$ with respect to t_i we obtain

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

and

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

As earlier, we may write:

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

During the litigation period I obtains the entire industry profit:

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

which is maximized under monopoly: $s\left(0\right)=s\left(t\right)$ for all $t\leq 1-l$. After the litigation period, incumbent's profit depends on the outcome of litigation:

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

Again, if the incumbent finds profitable to delay entry until any $t>1-l$, 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

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

where Y and X maximize

$$\theta\left[\Pi(1+y)+y\pi(1+y)\right]+(1-\theta)\underbrace{\left[\Pi(1+N-x)-x\pi(2+N-x)\right]}_{=g(x;0)}$$

subject to $x\leq x+y\leq N-1$. As industry profit is decreasing in the number of firms in the market, $Y=0$. Furthermore, X maximizes $g(x;0)$ and by Assumption 1 $X=X(0)\leq N-1$. Overall, the payoff in equilibrium with litigation is thus

$$L(\theta):=(1-l+\theta l)\,\Pi(1)+(1-\theta)\,lg(X(0);0)-C.$$

Comparison

The difference between I's payoffs in equilibrium without and with litigation can be written as

$$\frac{L(\theta)-S(\theta)}{l}=\underbrace{\theta\Pi(1)+(1-\theta)\,g(X(0);0)-g(X(\theta);\theta)}_{:=f(\theta)}-\underbrace{\frac{C+Nc}{l}}_{:=\lambda}.$$

It thus follows that there exists litigation in equilibrium if $\lambda<f(\theta)$ and no litigation if the reverse inequality holds.

Monotonicity

It remains to show that $X(\cdot)$ is weakly increasing and satisfies $X(1) = N$. For the last part, note that

$$g(x;1) = \Pi(1+N-x) + (N-x)\pi(1+N-x)$$

is the industry profit, which is maximized at $x = N$. Furthermore, we can rewrite $g(x;\theta)$ as

$$\begin{aligned} g(x;\theta) = {} & \theta\left[\Pi(1+N-x) + (N-x)\pi(1+N-x)\right] \\ & + (1-\theta)\left[\Pi(1+N-x) - x\pi(2+N-x)\right], \end{aligned}$$

where the term multiplying θ is the industry profit, which is increasing in x, and the term multiplying $(1-\theta)$ is maximized at $X(0) \leq N-1 < N = X(1)$. Thus, clearly, $X(\theta) = \arg\max_x g(x;\theta)$ is weakly increasing in θ.

Proof of Proposition 2

Note that $g(x;\theta)$ defines a family of affine functions of θ parametrized by x. The epigraph of an affine function is a half-space and any intersection of half-spaces is a convex set. The value function $g(X(\theta);\theta)$ is therefore convex and piecewise linear. Notice that

$$\frac{\partial g(x;\theta)}{\partial\theta} = (N-x)\pi(1+N-x) + x\pi(2+N-x) > 0.$$

Therefore, $g(X(\theta);\theta)$ is a strictly increasing, convex, piecewise linear function of θ. But then

$$f(\theta) = \theta \underbrace{\Pi(1)}_{=g(X(1);1)} + (1-\theta)\, g(X(0);0) - g(X(\theta);\theta)$$

is concave, piecewise linear function of θ. Furthermore, it clearly satisfies $f(0) = f(1) = 0$, so that for any $\lambda < \max f$, we can define thresholds of patent strength, $\underline{\theta}(\lambda) < \overline{\theta}(\lambda)$, where $\underline{\theta}(\lambda)$ satisfies $\underline{\theta}(0) = 0$ and increases in λ, and $\overline{\theta}(\lambda)$ satisfies $\overline{\theta}(0) = 1$ and decreases in λ, such that there is litigation in equilibrium if and only if $\theta \in \left[\underline{\theta}(\lambda), \overline{\theta}(\lambda)\right]$.

Proof of Proposition 3

Suppose first that all entrants settle in equilibrium: $s(0) = N$. Then, for each E_i the equilibrium payment is given by

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

where

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

and

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

Plugging the total payment

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

into the incumbent's payoff we 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

$$S_c(\theta) \coloneqq \Pi(1) - N \cdot \left[(1-\theta)(1-l)\,\pi(1+N) - c\right].$$

Suppose now that at least some E_j litigates in equilibrium. Then, for each E_i who settles we have

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

Differentiating with respect to t_i we obtain

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

and

$$p_i(1) = 0.$$

The total payment is

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

During the litigation period the incumbent has a possibility to obtain the entire industry profit:

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

Industry profit is maximized under a monopoly, thus the incumbent has no incentive to allow for entry, so $s\left(t\right) = s\left(0\right)$ for all $t \leq 1 - l$. After the litigation period, the incumbent's profit depends on the outcome of litigation:

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

which is maximized by setting $s\left(t\right) = s\left(0\right)$ for all $t > 1 - l$. Thus, in equilibrium with litigation, I obtains

$$L_c\left(\theta\right) \coloneqq \left(1 - l\right) \Pi\left(1\right) + l\left[\theta \Pi\left(1\right) + \left(1 - \theta\right) \Pi\left(1 + N\right)\right] - C,$$

which is always less than

$$
\begin{aligned}
S_c(\theta) &= \Pi(1) - N \cdot [(1-\theta)\, l\pi(1+N) - c] \\
&= (1-l)\,\Pi(1) + l\,[\theta\Pi(1) + (1-\theta)(\Pi(1) - N\pi(1+N))] + Nc \\
&\geq (1-l)\,\Pi(1) + l\,[\theta\Pi(1) + (1-\theta)\,\Pi(1+N)] - C = L_c(\theta)\,,
\end{aligned}
$$

by the assumption that the industry profit is decreasing in the number of firms in the market.

Proof of Proposition 4

Suppose first that all entrant settle in equilibrium, with $t_i \in [0, 1-l]$ for each E_i. Following similar steps as in the proof of Proposition 1, we may calculate the the payoff of I and show that during the litigation period it is never profitable for I to accommodate entry to the market. Thus, $t_i = 1-l$ for each E_i, which yields I the payoff

$$S_0(\theta) \coloneqq (1-l)\,\Pi(1) + lg(0;\theta) + Nc,$$

where

$$g(0;\theta) = \Pi(1+N) + \theta N\pi(1+N)\,.$$

Suppose now that one entrant litigates, $s(0) \leq N-1$ entrants settle and $N-1-s(0)$ entrants wait in equilibrium. Again, following similar steps as in the proof of Proposition 1, we may show that it is never profitable for I to license the patent. Thus, $s(0) = 0$ is optimal and I's payoff from litigation is

$$L_0(\theta) \coloneqq (1-l+\theta l)\,\Pi(1) + (1-\theta)\,lg(0;0) - C,$$

with $g(0;0) = \Pi(1+N)$. Taking the difference between I's payoffs in equilibrium without and with litigation one gets

$$\frac{L_0(\theta) - S_0(\theta)}{l} = \underbrace{\theta\Pi(1) + (1-\theta)\,g(0;0) - g(0;\theta)}_{:=f_0(\theta)} - \underbrace{\frac{C+Nc}{l}}_{:=\lambda},$$

where

$$f_0(\theta) = \theta\left[\Pi(1) - \Pi(1+N) - N\pi(1+N)\right].$$

It thus follows that there exists litigation in equilibrium if $\lambda < f(\theta)$ and no litigation if the reverse inequality holds.

Numerical example

The instantaneous profit functions are determined by a textbook Cournot quantity-setting game with zero marginal costs and an 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 $\Pi(1) = \frac{1}{4}$, $\Pi(2) = \pi(2) = \frac{1}{9}$ and $\Pi(3) = \pi(3) = \frac{1}{16}$. We thus have

$$\begin{aligned}
g(0;\theta) &= \theta 2\pi(3) + \Pi(3) = \frac{2\theta+1}{16}, \\
g(1;\theta) &= \theta\left[\pi(2) + \pi(3)\right] + \Pi(2) - \pi(3) = \frac{25\theta+7}{144}, \\
g(2;\theta) &= \theta 2\pi(2) + \Pi(1) - 2\pi(2) = \frac{8\theta+1}{36},
\end{aligned}$$

so that

$$X(\theta) = \begin{cases} 0 & \text{if } \theta \leq \frac{2}{7}, \\ 1 & \text{if } \theta \in \left[\frac{2}{7}, \frac{3}{7}\right], \\ 2 & \text{if } \theta \geq \frac{3}{7}, \end{cases}$$

and

$$f(\theta) = \begin{cases} \frac{\theta}{16} & \text{if } \theta \leq \frac{2}{7}, \\ \frac{\theta+1}{72} & \text{if } \theta \in \left[\frac{2}{7}, \frac{3}{7}\right], \\ \frac{5(1-\theta)}{144} & \text{if } \theta \geq \frac{3}{7}. \end{cases}$$

In particular,

$$f\left(\frac{2}{7}\right) = \frac{1}{56},$$
$$f\left(\frac{3}{7}\right) = \frac{5}{252}.$$

Thus, applying Proposition 1, if $\lambda > f(\theta)$, both entrants obtain a license if $\theta \leq \frac{2}{7}$, one entrant obtains a license while the other one is delayed if $\theta \in \left[\frac{2}{7}, \frac{3}{7}\right]$ and both entrants are delayed if $\theta \geq \frac{3}{7}$. If instead $\lambda < f(\theta)$, one of the entrants litigates while the other one waits.

The instantaneous consumer surpluses can be calculated by integrating the demand function from zero to the equilibrium industry output:

$$CS(1) = \frac{1}{2}[q(1)]^2 = \frac{1}{8},$$
$$CS(2) = \frac{1}{2}[2q(2)]^2 = \frac{2}{9},$$
$$CS(3) = \frac{1}{2}[3q(3)]^2 = \frac{9}{32}.$$

The expected consumer surplus under litigation is

$$\theta CS(1) + (1-\theta) CS(3) = \theta\frac{1}{8} + (1-\theta)\frac{9}{32},$$

which equals $CS(2)$ at $\theta = \frac{17}{45}$. When θ is below this threshold the consumer surplus in (license, delay) is higher than in (litigate, wait).