Pay-for-Delay with Settlement Externalities*

Emil Palikot†and Matias Pietola‡

June 16, 2021

Abstract

Pay-for-delay patent settlements, where a potential entrant withdraws its challenge of the incumbent's patent and stays out of the market, in exchange for financial compensation, cost patients and taxpayers billions of dollars in higher pharmaceutical prices. In a market with one incumbent and several entrants, the possibility of conditioning settlements on litigation outcomes against other entrants results in the exclusion of all entrants from the market. When conditional contracts are infeasible, the incumbent either licenses or fights entry in court: the resulting competition benefiting consumers. Prohibiting pay-for-delay settlements increases litigation and may harm consumers by undermining licensing incentives.

1 Introduction

According to the Federal Trade Commission, pay-for-delay patent settlements cost American consumers and taxpayers alone $3.5 billion a year in higher drug prices.[1] In these settlements, a potential entrant agrees not to challenge the validity of the incumbent's patent and to stay out of the market, in exchange for financial compensation. These deals fall at the intersection of antitrust and intellectual property policies. The European Commission considers them to be anti-competitive and has imposed significant fines on companies involved.[2] The Supreme Court of the United States and the Court of

*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.

†Stanford Graduate School of Business, Stanford University, palikot@stanford.edu.

‡The Finnish Competition and Consumer Authority, matias.pietola@kkv.fi.

[1] Pay-For-Delay: How Drug Company Pay-Offs Cost Consumers Billions: A Federal Trade Commission Staff Study. Available at https://www.ftc.gov/reports.

[2] The Commission's view has been supported by the EU General Court in *Servier* (Case T691/14, *Servier SAS v European Commission*, 2018) and *Lundbeck* (Case T-472/13 *H. Lundbeck A/S v European Commission*, 2016).

Justice of the European Union have instead adopted more nuanced views on the subject.[3]

In this article, we show that patent litigation and licensing are strategically tied to pay-for-delay settlements. Over the last few years, around 3-12% of all patent settlements in the pharmaceutical sector in the European Economic Area have been pay-for-delay deals.[4] Settlements, where the entrant buys a license from the incumbent and starts manufacturing the product, are more frequent. Furthermore, licensing contracts are often proposed by the same incumbent that offers pay-for-delay agreements to other, seemingly similar entrants. Yet, these different deals have been analyzed in isolation. To the best of our knowledge, the question of why pay-for-delay and licensing agreements coexist has not yet been addressed in the economic literature. This article aims to fill the gap.

We develop a model of contracting between 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. the patent to expire or to be overturned following litigation by another entrant), or settle with the incumbent. We consider two types of contracting environments: a simple one, where a settlement deal includes a financial transfer and an entry decision, and an enriched setting in which contracts can be conditioned on the validity of the patent. We briefly review antitrust cases involving pay-for-delay settlements and show that both simple pay-for-delay contracts, as well as ones conditioned on the validity of the patent, are used in practice.

Our analysis shows that the contracting environment in which firms operate is central to the equilibrium outcome. In the rich environment, where entry exclusion can be conditioned on the validity of the patent, the incumbent always purchases the exit of all entrants, regardless of the strength of its patent.[5] In contrast, when firms are restricted to use simple contracts, a complex outcome emerges with some entrants entering the market, either through licensing or successful litigation, and others accepting pay-for-delay deals and staying out.

From the industry's perspective, eliminating mutual competition and redistributing the incumbent's monopoly profit between the firms (using transfers) is jointly optimal. However, excluding a potential competitor from the market imposes a positive externality on the other entrants: their expected profit attainable through litigation increases, which makes it costlier to purchase their exit and

[3]See *FTC v. Actavis, Inc.*, 133 S. Ct. 2223, 570 U.S. 136, 186 L. Ed. 2d 343 (2013). See also Case C-307/18, *Generics (UK) Ltd and others v Competition and Markets Authority*, 2020.

[4]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 European Economic Area (EEA).

[5]In the framework we propose, the strength of the patent reflects the likelihood of the court invalidating the patent if an entrant starts a litigation.

also impacts the licensing fees they are willing to pay. Conditional pay-for-delay agreements eliminate these externalities by allowing for entry if the patent is overturned, making it less expensive for the incumbent to purchase the exit of all entrants.

By contrast, to monopolize the market with simple contracts, the incumbent would have to compensate every entrant for the foregone duopoly profit that it could achieve through litigation while all rival entrants stay out. When there are many entrants, this cost exceeds the incumbent's gain from maintaining the monopoly position. The incumbent instead allows some entry to the market either through licensing or litigation which reduces payments needed to exclude other entrants.

We find that patents of intermediate strength are litigated, whereas sufficiently weak or strong patents are not taken to court. The number of excluded entrants is higher when the patent is strong and litigation is more likely to occur when litigation costs are low. In the extreme case, where litigation is costless, it is always an equilibrium outcome, provided that competition is not very intense.

Our results contrast with predictions from single-entrant models. When facing only one challenger, the incumbent always excludes the entrant by sharing part of its monopoly profit (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 derive two policy implications from our analysis. First, disallowing conditional clauses improves consumer welfare by making it more difficult for the industry to achieve the monopoly outcome. We also argue that total surplus is likely to increase, even if some patent disputes are taken to court. This is because litigation is played in equilibrium only if the litigation costs are small, relative to the deadweight loss created by the lack of competition in monopoly. Second, the possibility of entering into simple pay-for-delay agreements encourages the incumbent to license the patent, which benefits consumers. Banning pay-for-delay agreements then reduces licensing and increases litigation, and may even reduce expected consumer surplus.

We provide several extensions and robustness checks. First, we relax the assumption of perfectly correlated litigation outcomes and assume litigation outcomes to be independent across entrants.[6] Despite this change, licensing and pay-for-delay agreements still coexist in equilibrium of the game with simple contracts. Second, we analyze the game when settlement offers are secret (litigation is still observed). We find that litigation is more likely in the equilibrium of the game with simple contracts. Finally, we allow parties to renegotiate pay-for-delay agreements after the patent has been

[6] This extension can be thought of as a model of patent infringement, as opposed to our baseline model of patent invalidation.

invalidated. This makes litigation less profitable for the entrants and leads to more entry exclusion in equilibrium. In all three extensions, we still find that contracting externalities often prevent the parties from achieving a jointly efficient outcome under simple contracts. Accommodating entry to the market enables the incumbent to decrease reverse payments needed to keep other entrants out. This basic logic underlying the results of the baseline model is largely robust to the changes of the model assumptions considered in extensions.

Related literature The current economic literature on patent settlements has two branches, one for pay-for-delay agreements (Shapiro, 2003; Elhauge and Krueger, 2012; Gratz, 2012; Edlin et al., 2015; Padilla and Meunier, 2015; Manganelli, 2019, 2021) and another for licensing uncertain patents (Crampes and Langinier, 2002; Lemley and Shapiro, 2005; Farrell and Shapiro, 2008; Encaoua and Lefouili, 2009; Amir et al., 2014). To the best of 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 Padilla and Meunier (2015). They focus on the credibility of a litigation threat when 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 this phenomenon, even though all entrants are symmetric and have the ability to litigate.[7] 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.

[7] Padilla and Meunier (2015) assume that only one of the entrants can start the litigation.

The case of multiple entrants is also studied in Marxen and Montez (2020), albeit they consider a different environment, without pay-for-delay agreements. In their model, firms can sign early entry agreements, which lead to entry just before the patent expiry date. By signing an early entry agreement with one entrant, the incumbent can ensure that others stay out, and make monopoly profit for longer.

The strategy of allowing the entry of an authorized generic producer (either by an entrant or by developing an in-house generic version of the drug) to deter entry by other entrants is also studied by Bokhari et al. (2020). They show that when the first-mover advantage (due to higher willingness-to-pay for the first generic relative to other generics) is high enough, the threat to launch an authorized generic is credible; as a consequence, subsequent entrants may optimally choose to stay out of the market. Lemus and Temnyalov (2020) also study the incentives to introduce a follow-on product by the incumbent. They focus on the case when the follow-on products are protected by an indisputable patent and cannibalize the original drug. In this setting, they show that the possibility of introducing a follow-on product mitigates losses due to invalidation of the first patent and generally decreases reverse payments associated with pay-for-delay settlements.

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 (i.e., the agreed entry date) with 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. Ecer et al. (2020) show that the prohibition of pay-for-delay contracts impacts parties' payoffs proportionally to their initial bargaining power.

Use of divide-and-conquer strategies to exploit a coordination failure by the plaintiffs have been previously studied by Daughety and Reinganum (2002), Che and Spier (2008), and Posner et al. (2010). The principal (typically the defendant) offers beneficial treatment to some of the agents (often plaintiffs) who decide to settle. As a result, other plaintiffs drop their lawsuits either because the fixed costs of the litigation are now too high or the probability of winning the case has decreased enough. In our model, the incumbent also exploits a coordination failure of the entrants, but the context and the modeling approach differs from the previous literature. We do not study coalition-building, and the probability of winning against the incumbent does not depend on litigation by other entrants.

Finally, although our analysis draws heavily from the seminal works of Segal (1999, 2003) on contracting with externalities, there are several important differences. First, in our model the agents (entrants) have two outside options: they can either decide to litigate or they can decide to wait. Second, in the same model, we allow the principal (incumbent) to offer contracts with both positive (pay-for-delay) and negative (licensing) externalities to other entrants. Therefore, we study how an incumbent should optimally balance-off the two types of contracts. Finally, we explicitly consider and compare the equilibrium outcome under a rich contractual space, which includes settlements conditional on the validity of the patent and the narrower contracting space when such contracts are not feasible. We show that parties can resolve the coordination failure by introducing the conditional contracts, and we also study the consequences of doing so for consumer welfare.

The article is organized as follows. The next section is devoted to case studies in the pharmaceutical industry. We focus on the contracting environment, where pay-for-delay agreements are signed, and briefly summarize the antitrust response to them. Section 3 introduces the baseline model. In Section 4, we characterize the equilibrium of the model. Section 5 presents welfare analysis and discusses policy implications. In Section 6, we present extensions and robustness checks and Section 7 concludes.

2 Pay-for-delay in the pharmaceutical industry

The last decade saw several high-profile antitrust cases investigating pay-for-delay settlements. Documents presented in support of the decisions taken in these cases shed light on the practice of settling patent disputes and help to justify the modeling assumptions of our baseline analysis. Specifically, we explain why we consider a simultaneous game instead of a sequential one, why we assume perfectly correlated litigation outcomes and commonly known "strength" of the patent, and why we focus on publicly observable contract offers. Finally, we discuss different types of pay-for-delay agreements.

Two factors are the main drivers of competition between pharmaceutical companies: 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, generic producers can simultaneously contest the incumbent's monopoly. Generic entry typically decreases prices significantly and the patented drug loses market share.[8] The incumbent thus has a strong incentive to extend its monopoly period. A common practice is to apply for a new patent, which instead of protecting the main chemical compound, might protect the manufacturing process instead. Although these secondary patents are sometimes very weak and can be successfully challenged in court, they are presumed to be valid until a court of law overturns them. As such, parties to a patent dispute often decide to save time and litigation resources, and settle out of court.

The expiry date of a certain irrefutable patent is the natural entry date for generics. Hence, when we do observe sequential entry into the market, it is rather an outcome of strategic interaction, than an underlying feature of the environment itself. A contracting game that is often played close to the patent's expiry date determines which generic producers enter the market and at what date.[9]

The competitive effects of entry concern not only the incumbent but also the entrants themselves. In particular, the incentives to challenge the incumbent's patent depend on the number of entrants already in the market and whether more entries will follow, should the patent be invalidated. A legal counsel of Niche, one of the generic producers in *Servier*, explains this concern:

> "In the view of Niche, it was in the interests of neither party to engage in litigation on the validity and infringement of Servier's patents in open court. If Niche were successful in revoking Servier's patents, this would obviously be damaging for Servier. However, it would also not be particularly advantageous for Niche, given that it would open the way for other generic entrants into the market. Niche did not want to 'win the battle but lose the war.' "[10]

This strategic interdependence between entrants motivates our focus on contracting between the incumbent and multiple entrants. Furthermore, we incorporate free-riding on litigation effort in our baseline model by assuming perfectly correlated litigation outcomes. When one entrant litigates, the others have the option to wait and enter the market if the patent is overturned by the court. As an extension, we also look at the other polar case of independent litigation outcomes. Free-riding is no

[8]See, Pharmaceutical Sector Inquiry by European Commission. For example, when Servier, a French pharmaceutical company, lost one of its key patents related to perindopril, a medication used to treat high blood pressure, in the UK, generic entry decreased prices by almost 90 percent from Servier's original price.

[9]A notable exception to the simultaneous contestability is the US, where the so-called Hatch-Waxman Act provides 180 days of exclusivity to the first generic producer to challenge the incumbent. However, in other parts of the world, including the EU, there is no such law, and the incumbent must simultaneously settle with all challengers to avoid litigation.

[10]See paragraph 493 of the European Commission's decision in *Servier*. This issue is also highlighted by the Hatch-Waxman Act, which aims to restore the incentives to challenge the incumbent by shielding against follow-up entry into the market.

longer possible when litigation outcomes are independent. However, market-entry still impacts other entrants' incentives to challenge the patent. This extension can be thought of as a model of patent infringement, and our baseline as a model of patent invalidation.

The free-riding problem makes it more tempting to accept a settlement from the incumbent. However, to reach a mutually beneficial settlement, parties need to have a similar assessment of the strength of the patent. Pharmaceutical firms, frequently, use laboratory tests and seek third-party advice to ensure that their assessments reflect the actual probability of patent invalidation.[11] With this evidence in mind, we think that assuming commonly known patent strength is a good approximation of reality.

Pharmaceutical companies tend to have good market intelligence. Internal documents presented in *Servier* show that Lupin, a producer of a generic version of the Servier product, knew exactly what type of deals the other generics had obtained from Servier: Krka, another generic producer, had received a licensing deal and the others were being excluded.[12] In our baseline analysis, we thus assume that contract offers are publicly observable. However, we do include an extension where we study secret offers with publicly observable litigation.

Pay-for-delay contracts can be divided into two broad categories: first, conditional pay-for-delay agreements, where the generic producer can nevertheless, enter if the incumbent's patent is revoked following litigation by another generic; and second simple agreements, where there is no such option.

Interestingly, in *Servier* a simple pay-for-delay agreement with Teva (an entrant) was amended to include the possibility of entry should another generic producer successfully challenge Servier's patent:

> "It should be stressed - as noted by Servier and Teva (see below) - that Amendment No 1 did not allow for an early – let alone immediate – entry for Teva. Servier and Teva agreed to tie Teva's entry date to the resolution of the UK proceedings between Apotex and Servier (or the expiry/revocation of the '947 patent)."[13]

However, such a condition is not always included. In the other landmark case, Lundbeck used simple pay-for-delay agreements to purchase the exit of a generic producer Merck. In the UK, Lundbeck had entered into pay-for-delay agreements with two generic producers (Merck and Arrow), paying each of them an amount to approximate what they would make being the only generic in the

[11] See paragraph 709 of the European Commission's decision in *Servier* and paragraph 522 of its decision in *Lundbeck*.

[12] See paragraph 1023 of the European Commission's decision in *Servier*.

[13] See paragraphs 772 and 774 of the European Commission's decision in *Servier*.

market: in other words, what they would get by rejecting the offer and entering the market, while the other firm is excluded.[14]

An alternative strategy aimed at excluding a rival is to acquire the company and discontinue the effort of bringing the generic to the market. In *Lundbeck*, generic producer VIS was acquired by Lundbeck and its application for a market authorization was later removed.[15] When launching a specific generic drug is the sole objective of a company, a pay-for-delay contract (a simple one, by the nature of an acquisition) closely resembles such a "killer acquisition".[16]

3 Baseline model

Consider a model in which one firm, "the incumbent," owns an uncertain patent and can contract with n other symmetric firms – "the entrants." The incumbent's trade with each entrant i is denoted by x_i in $\{0, 1, L, W\}$, where 0 means that the entrant stays out of the market, 1 denotes entry, L litigation, and W waiting.[17] Litigation is costly and results in entry with probability $1 - \theta$, where θ reflects the strength of the patent. Waiting costs nothing and leads to entry with the same probability if another entrant litigates and to no entry otherwise. Litigation outcomes are assumed to be perfectly correlated. We analyze the other polar case of independent litigation outcomes in Section 6.

For any trade profile, $\mathbf{x} = (x_1, \ldots, x_n)$, we denote by $q_i(\mathbf{x})$ the random variable that indicates whether or not i enters the market. The number of active firms, including the incumbent, is given by $Q(\mathbf{x}) \triangleq 1 + \sum_i q_i(\mathbf{x})$ with expected value $\mathbb{E}[Q(\mathbf{x})]$. Let k by a realization of $Q(\mathbf{x})$; given k the incumbent makes a profit $\Pi(k)$ and each active entrant makes a profit $\pi(k)$, whereas the entrants that stay out make zero profit. The profits are all positive and total industry profit decreases in k.

As firms compete in the same market, contracting externalities between them arise. Entrant i's payoff is its expected profit $\mathbb{E}[\pi(Q(\mathbf{x}))\, q_i(\mathbf{x})]$ net of the payment t_i and cost of action $c(x_i)$, which equals the entrant's litigation cost c if $x_i = L$ and zero otherwise. The incumbent's payoff is the sum of its expected profit $\mathbb{E}[\Pi(Q(\mathbf{x}))]$ and total payment $\sum_i t_i$ net of its cost of action $C(\mathbf{x})$, which equals the incumbent's litigation cost C if at least one of the entrants litigates and otherwise it is zero.[18]

[14] See paragraphs 239, 267, and 398 of the European Commission's decision in *Lundbeck*.

[15] See paragraph 220 of the European Commission's decision in *Lundbeck*.

[16] The phenomenon of the so-called killer acquisitions is well documented in the pharmaceutical sector, see for example Cunningham et al. (2021).

[17] Our results generalize to a dynamic setting in which a contract may include any entry date from an interval starting at date zero and ending at patent expiration, and where litigation is time-consuming. In equilibrium, any agreed entry date is either the patent expiration date (corresponding to exit) or the date when the court would give its judgement (corresponding to entry). We allow for discounting. Details of this analysis are available in Appendix A.

[18] Assuming that the incumbent incurs its litigation cost only once is innocuous, as one entrant at most will litigate against the incumbent in equilibrium (even if the incumbent would incur its litigation cost many times).

We say that the incumbent sells entry or licenses the patent if $x_i = 1$ and $t_i > 0$. On the contrary, the incumbent buys exit or signs a pay-for-delay agreement if $x_i = 0$ with a "reverse payment" $t_i < 0$. Furthermore, $x_i = W$ with $t_i < 0$ defines conditional pay-for-delay agreements, where the incumbent pays the entrant to stay out of the market on the condition that the patent stays valid. Thus, if another entrant successfully litigates against the incumbent, i enters the market. The entrants have two outside options: $x_i \in \{L, W\}$ with zero payment.

By the assumption that the total industry profit is decreasing in the number of firms in the market, the expected producer surplus, $PS(\mathbf{x})$, which equals the expected industry profit net of the total cost of action, is maximized in monopoly with no litigation: $\max_{\mathbf{x}} PS(\mathbf{x}) = \Pi(1)$. Any combination of waiting and pay-for-delay agreements achieves this. Indeed, if all entrants either stay out or wait for a rival to challenge the patent, there is no entry and the market is monopolized by the incumbent.

We are interested in departures from the monopoly benchmark due to contracting externalities between the firms. To study them, we consider the following two-stage game:

- **Stage 1**: The incumbent commits to a set $\{x_i, t_i\}_{i=1}^{n}$ of publicly observable bilateral contract offers to the entrants.

- **Stage 2**: The entrants simultaneously decide either to accept their respective offers, reject the offer and litigate, or reject the offer and wait.

After the second stage, a court decides about the validity of the patent if there is litigation, and payoffs are realized. As in Segal (1999), we study the incumbent's preferred subgame perfect Nash equilibrium of the game.

The incumbent can always offer L or W with a zero payment, so without loss of generality we can restrict attention to equilibria in which every entrant accepts its offer. A strategy profile constitutes a second-stage Nash equilibrium if and only if the following participation constraint is satisfied:

$$\mathbb{E}\left[\pi\left(Q\left(\mathbf{x}\right)\right) q_i\left(\mathbf{x}\right)\right] - c\left(x_i\right) - t_i \geq r_i\left(\mathbf{x}_{-i}\right) \text{ for every } i. \tag{1}$$

On the right-hand side of the inequality (1) is the reservation utility or rent of an entrant. This amounts to what the entrant would obtain by rejecting the offer and choosing the better of the two available outside options:

$$r_i\left(\mathbf{x}_{-i}\right) \triangleq \max_{x_i \in \{L, W\}} \mathbb{E}\left[\pi\left(Q\left(x_i, \mathbf{x}_{-i}\right)\right) q_i\left(x_i, \mathbf{x}_{-i}\right)\right] - c\left(x_i\right).$$

In the incumbent's preferred subgame perfect equilibrium, all participation constraints must bind, as otherwise the incumbent could make more profit by raising some payments. Deriving transfers from the binding participation constraints and substituting them into the incumbent's payoff, we obtain

$$\mathbb{E}\left[\Pi\left(Q\left(\mathbf{x}\right)\right)\right]+\sum_{i}t_{i}-C\left(\mathbf{x}\right)=PS\left(\mathbf{x}\right)-\sum_{i}r_{i}\left(\mathbf{x}_{-i}\right).$$

The incumbent maximizes the producer surplus net of the rents it has to leave to the entrants.

Notice that there is no credible litigation threat for a sufficiently strong patent. Suppose one entrant challenges patent validity in court, whereas all other entrants wait and free ride on this effort. The challenger wins with probability $1-\theta$ and faces tough competition. Its expected profit is:

$$\underline{c}\left(\theta\right)\triangleq\left(1-\theta\right)\pi\left(1+n\right).$$

If the litigation cost c is larger than $\underline{c}\left(\theta\right)$ (i.e., the patent is strong enough), the challenger is weakly better off when it drops the case. We obtain the following result:

Proposition 1. *If there is no credible litigation threat, i.e. $c\geq\underline{c}\left(\theta\right)$, then by offering the outside option $x_i^*=W$ with $t_i^*=0$ to each i, the incumbent monopolizes the market at no cost.*

Proof. Waiting $\left(W,\ldots,W\right)$ maximizes $PS\left(\mathbf{x}\right)$ at $\Pi\left(1\right)$, which entirely goes to the incumbent, as $c\geq\underline{c}\left(\theta\right)$ implies $r_i\left(W,\ldots,W\right)=0$ for each i (waiting is the better outside option). □

Proposition 1 is reassuring: everything else fixed, stronger patents are less likely to be challenged in court and result in a monopoly more often. The concept of the patent system is to reward a true innovator with a legal monopoly, and a strong patent that keeps entry at bay ensures it. However, "strong" in this context also relates to the number of entrants and their cost of litigation: with high litigation cost or many potential entrants, even a weak patent can lead to a monopoly.

Remark 1. In the two-stage offer game we implicitly assume that the incumbent is able to commit to pursue litigation even if ex post, after the second stage of the game (at which the entrants decide whether to accept or reject the offers), it would have an incentive to acknowledge patent revocation to save on its own litigation cost. Lack of commitment to litigate would increase the entrants' reservation utilities and lead to more entry in equilibrium.[19]

[19] The incumbent is less likely to have a credible litigation threat when a high number of entrants have signed pay-for-delay agreements (it has less to gain by going to court).

4 Main results under a credible litigation threat

The outcome of the game becomes more nuanced when the litigation threat is credible, i.e., $c < \underline{c}(\theta)$, which we assume throughout this section. We will analyze two contracting environments: first, the unrestricted environment where conditional pay-for-delay agreements (as defined above) are feasible; and second, the restricted environment where they are not feasible and parties use simple contracts. We analyze both frameworks for two reasons. First, both types of agreements appear in practice.[20] Second, we are interested in the implications of conditional pay-for-delay agreements for market entry and total welfare, and this requires analyzing the case when such contracts are not allowed.[21]

Conditional pay-for-delay agreements

When the litigation threat is credible, the entrants have positive litigation payoffs even if all rival entrants wait and free-ride on the litigation effort. Thus, the free-riding problem between the entrants no longer protects the incumbent against entry to the market. Nevertheless, the incumbent is able to profitably maintain its monopoly position, by using conditional pay-for-delay agreements:

Proposition 2. *If conditional pay-for-delay agreements are feasible, then the incumbent monopolizes the market by offering:*

$$x_i^* = W \text{ with } t_i^* = -(1-\theta)\,\pi(1+n) + c < 0 \text{ to each } i.$$

Proof. First, $(W, \ldots, W)$ maximizes $PS(\mathbf{x})$ at $\Pi(1)$. Furthermore, for any $\mathbf{x}$ we have

$$r_i(W, \ldots, W) = \underline{c}(\theta) - c \le r_i(\mathbf{x}_{-i}) \text{ for each } i.$$

Thus, $x_i = W$ with $t_i = -r_i(W, \ldots, W)$ for each i maximizes incumbent's payoff. □

When all entrants accept a conditional pay-for-delay agreement, a deviation to litigation entails paying the litigation cost and facing severe competition if the court invalidates the patent (whereby all other entrants can now enter). As a result, the reservation utilities of the entrants, and hence the reverse payments needed to keep them out, are small. This makes it less expensive for the incumbent to monopolize the market by purchasing the exit of all entrants.[22]

[20] See Section 2.

[21] In fact, it has been argued that delaying entry only on the condition that the patent remains valid is less harmful, because rivals are free to enter the market in the event that a court declares the patent invalid. See *Servier*, for example.

[22] An increase in the number of potential entrants actually makes the incumbent better off if the entrants' joint profit is decreasing in n.

Simple contracts

In light of Proposition 2, it is natural to ask what would happen if conditional pay-for-delay agreements were not used, for example, due to legal reasons or incomplete contracting. We divide the analysis of simple contracts into two cases, depending on whether there is litigation or not.

Without litigation there is no uncertainty regarding entry to the market. The number of market participants is $Q(\mathbf{x}) = k$ with probability one, comprising of the incumbent and $k-1$ licensees. All other $n+1-k$ entrants sign pay-for-delay agreements. Using the participation constraints, we can solve for the licensing fees $t_i = \theta\pi(k) + c$ and reverse payments $-t_i = (1-\theta)\pi(k+1) - c$, that yield the entrants their reservation litigation payoffs.[23] Summing all of the payments, we obtain the incumbent's payoff without litigation:

$$\Pi(k) + \sum_i t_i = f(k;\theta) + nc,$$

where

$$f(k;\theta) \triangleq \Pi(k) + \theta(k-1)\pi(k) - (1-\theta)(n+1-k)\pi(k+1).$$

Note that, $f(k;1)$ equals the industry profit. We denote $k(\theta) \triangleq \arg\max_k f(k;\theta)$ as the optimal level of competition from the incumbent's point of view, satisfying $k(1) = 1$. For n large enough, $k(0) > 1$ necessarily holds (licensing is profitable), because buying the exit of all entrants becomes too costly.

Suppose now that, instead of reaching a settlement with everyone, the incumbent pursues litigation with one of the entrants; this entrant now has to pay the litigation cost. Then, any rival entrant who rejects its offer will wait and free-ride on the litigation effort. The payoff from waiting pins down reverse payments $-t_i = (1-\theta)\pi(k+1)$, where k now (with a slight abuse of notation) denotes the number of active firms in the market *if* the patent is declared invalid. As the reverse payments depend on the level of competition only when the patent is overturned, there is no point in licensing: the incumbent can reduce the reverse payments by the same amount through waiting as through licensing, without losing any profit if the patent is upheld by the court.

It thus follows that k includes the incumbent, the challenger and the entrants who wait. As there are no licensees, $Q(\mathbf{x}) = 1$ with probability θ and $Q(\mathbf{x}) = k$ otherwise. By summing the payments,

[23]Observe that the reverse payments depend on $k+1$ instead of k, because an entrant rejecting a pay-for-delay agreement out of equilibrium adds to the k market participants ($k-1$ licensees and the incumbent).

we obtain the incumbent's litigation payoff

$$\mathbb{E}\left[\Pi\left(Q\left(\mathbf{x}\right)\right)\right]+\sum_{i}t_{i}-C\left(\mathbf{x}\right)=\theta\Pi\left(1\right)+\left(1-\theta\right)f\left(k;0\right)-\mathcal{C}.$$

Importantly, a stronger patent improves the incumbent's chance of monopolizing the market *and* reduces the reverse payments paid to exclude entrants. As a result, the optimal amount of competition, which we denote by $k^{*}\triangleq\arg\max_{k>1}f\left(k;0\right)$, is independent of the strength of the patent.[24]

The incumbent's payoffs from settling with all entrants are the same as when pursuing litigation if and only if the expected gain from litigation:

$$G\left(\theta\right)\triangleq\theta\Pi\left(1\right)+\left(1-\theta\right)f\left(k^{*};0\right)-f\left(k\left(\theta\right);\theta\right),$$

equals the total litigation cost $\mathcal{C}=C+nc$. If the total cost of litigation is higher than the expected gain, the incumbent prefers to avoid litigation and uses licensing to reduce the cost of excluding entry. In the opposite case it is profitable to test patent validity in court.[25]

In particular, it is possible that the expected gain from litigation is negative, so that the incumbent prefers to avoid litigation even if the total litigation cost is zero. Interestingly, this is the case when licensing the patent is not profitable: $k\left(0\right)=1$. We have the following result for this special case:

Proposition 3. *Suppose that conditional pay-for-delay agreements are infeasible and $k\left(0\right)=1$. Then, the incumbent monopolizes the market by offering:*

$$x_{i}^{*}=0\textit{ with }t_{i}^{*}=-\left(1-\theta\right)\pi\left(2\right)+c<0\textit{ for each }i.$$

Proof. If $k\left(0\right)=1$, then for any θ we have $k\left(\theta\right)=1$ and thus $f\left(k\left(\theta\right);\theta\right)=f\left(1;\theta\right)$. For any θ we may then write $f\left(1;\theta\right)=\theta\Pi\left(1\right)+\left(1-\theta\right)f\left(k\left(0\right);0\right)$ and therefore

$$G\left(\theta\right)=\left(1-\theta\right)\left[f\left(k^{*};0\right)-f\left(k\left(0\right);0\right)\right]<0,$$

where $f\left(k\left(0\right);0\right)>f\left(k^{*};0\right)$ by definition. □

Proposition 3 shows that when the monopoly profit is high enough (i.e., when licensing is never

[24]Note that $k>1$, because one of the entrants litigates.

[25]The incumbent only cares about the total litigation cost, because without litigation it is able to extract the litigation cost from each entrant, whereas under litigation it only incurs its own litigation cost once.

profitable) then the incumbent always avoids litigation and purchases the exit of all entrants by using simple pay-for-delay agreements. In this case, even with a null patent, it is better to pay n times the duopoly profit than face competition.

However, this price becomes too high for n large enough. In fact, for many modes of competition, including our example below, licensing a null patent is already profitable for $n = 2$. Then, as Proposition 4 shows, there is litigation for patents of intermediate strength.

Proposition 4. *Suppose that conditional pay-for-delay agreements are infeasible and $k(0) > 1$. Then, for any $\mathcal{C} < \max G$ there exist thresholds of patent strength, $\underline{\theta}(\mathcal{C}) < \overline{\theta}(\mathcal{C})$, where $\underline{\theta}(\mathcal{C})$ satisfies $\underline{\theta}(0) = 0$ and increases in $\mathcal{C}$, and $\overline{\theta}(\mathcal{C})$ satisfies $\overline{\theta}(0) = 1$ and decreases in $\mathcal{C}$, such that:*

- *Patents of intermediate strength, $\theta \in [\underline{\theta}(\mathcal{C}), \overline{\theta}(\mathcal{C})]$, are litigated: $n + 1 - k^*$ entrants are excluded, one entrant litigates and $k^* - 2$ wait;*
- *Patents that are sufficiently strong, $\theta > \overline{\theta}(\mathcal{C})$, or weak, $\theta < \underline{\theta}(\mathcal{C})$, are not taken to court: $n + 1 - k(\theta)$ entrants are excluded and $k(\theta) - 1$ buy a license.*

Furthermore, $k(\cdot)$ is weakly decreasing, and satisfies $k(0) = k^*$ and $k(1) = 1$.

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

$$\frac{\partial f(k;\theta)}{\partial \theta} = (k-1)\,\pi(k) + (n+1-k)\,\pi(k+1) > 0.$$

Therefore, $f(k(\theta);\theta)$ is a strictly increasing, convex, piecewise linear function of θ. This then implies that $G(\theta)$ is a concave, piecewise linear function of θ. Furthermore, as $k^* = k(0)$ by $k(0) > 1$, it satisfies $G(0) = G(1) = 0$, so that for any $\mathcal{C} < \max G$, we can define thresholds of patent strength, $\underline{\theta}(\mathcal{C}) < \overline{\theta}(\mathcal{C})$, where $\underline{\theta}(\mathcal{C})$ satisfies $\underline{\theta}(0) = 0$ and increases in $\mathcal{C}$, and $\overline{\theta}(\mathcal{C})$ satisfies $\overline{\theta}(0) = 1$ and decreases in $\mathcal{C}$, such that $G(\theta) \geq \mathcal{C}$ if and only if $\theta \in [\underline{\theta}(\mathcal{C}), \overline{\theta}(\mathcal{C})]$.

It remains to show that $k(\theta)$ is weakly decreasing in θ and satisfies $k(1) = 1$. For the latter part, note that $f(k;1)$ equals the industry profit, being decreasing in k and maximized at $k = 1$. Furthermore, we can write $f(k;\theta) = \theta f(k;1) + (1-\theta) f(k;0)$, where $f(k;0)$ is maximized at $k(0) > 1$. Thus, $k(\theta)$ must be weakly decreasing in θ. □

With simple contracts, an entrant accepting a pay-for-delay agreement remains bound by it even if another entrant litigates and a court declares the patent invalid. A pay-for-delay agreement with one entrant thus creates a positive externality on the others who face reduced competition. This positive externality increases the cost of excluding entry considerably. To monopolize the market, the incumbent must compensate each entrant with the expected duopoly litigation payoff it would obtain by rejecting the offer and going to court while all rival entrants stay out of the market. For a sufficiently large number of entrants, the cost will exceed the gain from monopolization.

There are two ways the incumbent can decrease the cost of excluding entry. It can either start licensing the patent to some of the entrants or take the patent dispute to court, facing the risk of invalidation. Going to court is costly but gives the incumbent 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. In such a case, the cost of litigation relative to the reverse payments is high so the incumbent prefers to avoid litigation. If instead the patent is weak, the incumbent's chance of monopolizing the market through litigation is small, so the entrants have a strong outside option. Therefore, for a weak enough patent, the incumbent offers licensing deals to save on the costs of litigation.

Litigation occurs in equilibrium for patents of intermediate strength and is less likely when total litigation costs are high. To build intuition for this result, it is useful to look at the extremes of the patent strength. For a patent of zero strength, the incumbent will undoubtedly lose in court, which results in the same amount of competition as licensing: the number of excluded entrants is the same. Thus, following a small increase in patent strength, the expected gain from litigation is positive: there is a small chance of monopolizing the market, whereas licensing accommodates entry for sure. Recall that the industry profit is decreasing in the number of firms in the market, so the impact of a marginal increase in the probability of monopolizing the market on the incumbent's payoff is higher than the gain through licensing fees.

If instead the patent is indisputable, licensing is not profitable and the expected gain from litigation is again zero: buying the exit of all the entrants costs nothing, and litigation results in a monopoly for sure. Starting from this extreme, a small decrease in patent strength makes litigation relatively more profitable, because the increase in the cost of excluding all entrants is higher than the profit lost by the incumbent in the unlikely event that the court declares the patent invalid.

In both cases, as the strength of the patent moves away from the extreme, litigation becomes more attractive to the incumbent. It will pursue this strategy as soon as the relative advantage of litigation over settling with all entrants, is higher than total litigation costs.

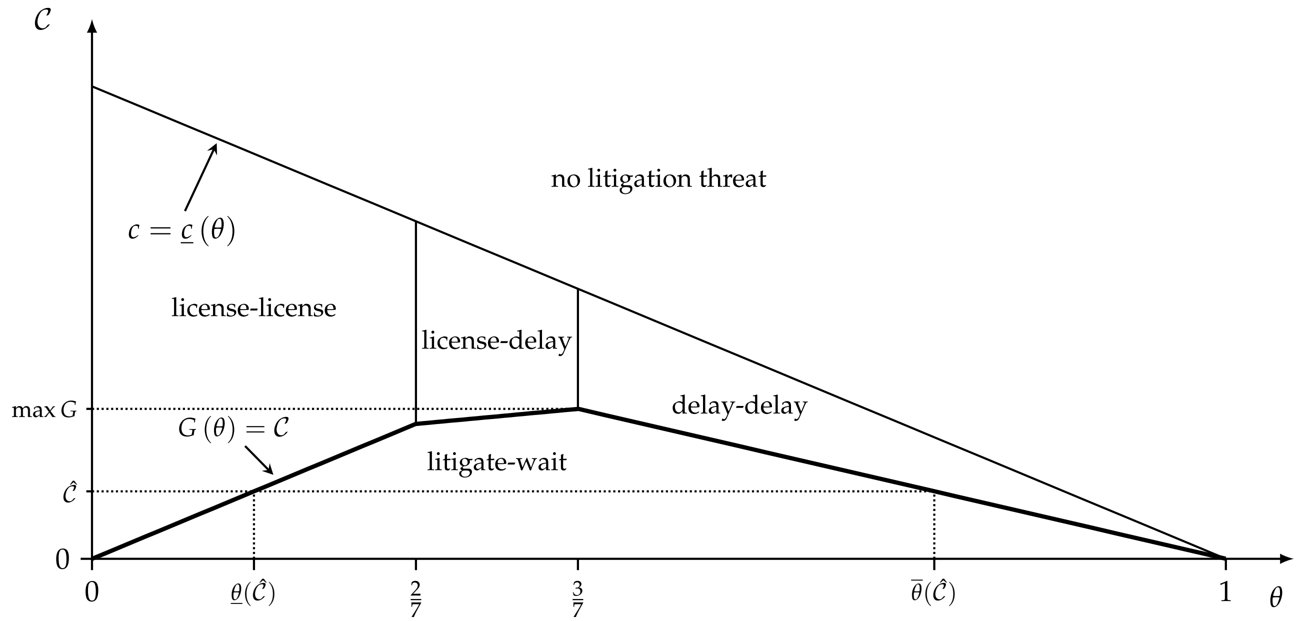

Figure 1: Equilibrium of the game with simple contracts for two entrants, as a function of the strength of the patent, θ, and the total cost of litigation, $\mathcal{C}$. Profits are as in the symmetric Cournot quantity-setting game with zero marginal costs and consumer valuations uniformly drawn from the unit interval.

Example 1. We illustrate the equilibrium outcome of the game with a textbook Cournot quantity setting game with two entrants, zero marginal costs and consumer valuations that are uniformly distributed on the unit interval. Figure 1 depicts the outcome as a function of the strength of the patent and the costs of litigation.

In this example, the profits are $\Pi(1) = \frac{1}{4}$, $\Pi(2) = \pi(2) = \frac{1}{9}$ and $\Pi(3) = \pi(3) = \frac{1}{16}$. The optimal settlement strategy of the incumbent is then given by

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

Both entrants buy a license if the patent is weak and sign pay-for-delay agreements if the patent is strong. For patents of intermediate strength, one entrant gets a license and the other is excluded. Furthermore, under litigation $k^* = k(0) = 3$, including the incumbent, the challenger and the freerider. The expected gain from litigation is:

$$G(\theta) = \frac{1}{144} \cdot \begin{cases} 9\theta & \text{if } \theta \leq \frac{2}{7}, \\ 2(1+\theta) & \text{if } \theta \in \left[\frac{2}{7}, \frac{3}{7}\right], \\ 5(1-\theta) & \text{if } \theta \geq \frac{3}{7}, \end{cases}$$

which is drawn in Figure 1. Notice that this is a piecewise linear and concave function, which satisfies $G(0) = G(1) = 0$, and attains its maximum at $\frac{5}{252}$. Therefore, for any $\mathcal{C} < \frac{5}{252}$, there exists a nonempty interval of patent strength, for which litigation takes place in equilibrium.

Remark 2. The incumbent implements a divide-and-conquer strategy, where identical entrants receive different equilibrium payoffs. If there is litigation in equilibrium, the entrants who wait (the free-riders) obtain the highest payoff, $(1-\theta)\,\pi(k(0))$, whereas the challenger gets the same profit but has to pay for the litigation cost.[26] The excluded entrants get reverse payments equal to $(1-\theta)\,\pi(k(0)+1)$, because the incumbent has to compensate them for the payoff that they would obtain by rejecting the pay-for-delay agreement. Depending on how costly the litigation is, the excluded entrants are better or worse off than the challenger. If instead there is no litigation in equilibrium, each excluded entrant obtains $(1-\theta)\,\pi(k(\theta)+1) - c$, whereas licensees get $(1-\theta)\,\pi(k(\theta)) - c$. Thus, licensees are always better off than the entrants with pay-for-delay agreements.

In the next section, we study the welfare implications of pay-for-delay agreements. Before this it is useful to summarize our main findings; first, the free-riding problem between the entrants protects the incumbent's monopoly when the patent is sufficiently strong. Increasing the number of potential entrants or their cost of litigation exacerbates this problem. As a result, there may be no credible litigation threat even for "weak" patents. Second, under a credible litigation threat, conditional pay-for-delay agreements enable the incumbent to keep its monopoly position at a low price. By contrast, when conditional deals are not feasible, the incumbent must pay high reverse payments to exclude all entrants, which is often unprofitable. It is then better to license weak patents and take uncertain patents to court. Only sufficiently strong patents lead to a monopoly.

[26] Our model is agnostic as to how the entrants determine which one pays the litigation cost. The literature studying volunteer's dilemma provides some suggestions on how the litigator might emerge: one possibility is that a lot determines who is the unlucky challenger (communication between entrants is needed for this to be possible). See Diekmann (1985) for a detailed discussion of this class of games. Additionally, experimental studies suggests that individuals do volunteer in similar scenarios - see Goeree et al. (2017) for an example.

5 Welfare analysis and policy implications

In this section we analyze how pay-for-delay settlements affect the expected consumer surplus, defined as $CS(\mathbf{x}) \triangleq \mathbb{E}[cs(Q(\mathbf{x}))]$, where $cs(k)$ denotes the *ex post* consumer surplus. We assume that $cs(k)$ increases in k – the realized number of active firms in the market.[27] We also discuss the implications of pay-for-delay for total surplus, $CS(\mathbf{x}) + PS(\mathbf{x})$, which accounts for the costs of litigation.

We proceed in two steps. First, we examine the consequences of a ban on conditional pay-for-delay agreements only, allowing for simple pay-for-delay contracts. Second, we look at the effects of also prohibiting simple pay-for-delay agreements. Throughout the analysis, we assume that the litigation threat is credible, as otherwise pay-for-delay agreements are not signed in equilibrium and the welfare effects from prohibiting them are zero.

Prohibiting conditional pay-for-delay agreements

Proposition 2 shows that conditioning pay-for-delay contracts on the validity of the patent reduces the reverse payments to such an extent that the incumbent finds it profitable to buy the exit of all entrants, even if there are many of them and the patent is weak. Proposition 5 states the ensuing policy implication.

Proposition 5. *Prohibiting only conditional pay-for-delay agreements weakly increases the expected consumer surplus. Consumers are always strictly better off from such a policy if the patent is weak enough and n is sufficiently large.*

Proof. By Propositions 2, 3 and 4, the change in expected consumer surplus is $cs(k(\theta)) - cs(1)$ if $G(\theta) < \mathcal{C}$ and $(1-\theta)[cs(k(0)) - cs(1)]$ if instead $G(\theta) > \mathcal{C}$. In both cases the change is weakly positive, as $cs(k)$ is increasing in k and $k(\theta) \geq 1$ for any θ. For the last part, recall that, for n large enough, $k(0) > 1$ necessarily holds. □

Consumers are strictly better off under a ban on conditional pay-for-delay settlements when the equilibrium outcome under simple pay-for-delay contracts involves some entry to the market. Proposition 4 shows that the incumbent will accommodate entry into the market or litigate for sufficiently weak patents and a large number of entrants. Both litigation and licensing increase the expected consumer surplus compared to the monopoly outcome, which is the outcome with conditional pay-for-delay agreements.

[27] In our numerical example $cs(1) = \frac{1}{8}$, $cs(2) = \frac{2}{9}$ and $cs(3) = \frac{9}{32}$.

The effect on total surplus is slightly more complex because conditional pay-for-delay agreements save on litigation costs whenever there is litigation without them. However, these savings are likely outweighed by the expected deadweight loss caused by lack of competition. Recall that litigation occurs only when litigation costs are small compared to the incumbent's chance of monopolizing the market. In particular, with our numerical example the difference between the deadweight loss and the incumbent's private gain from litigation is always positive:

$$DWL - G(\theta) = \frac{1}{288} \cdot \begin{cases} 9(35 - 17\theta) & \text{if } \theta \leq \frac{2}{7}, \\ 311 - 139\theta & \text{if } \theta \in \left[\frac{2}{7}, \frac{3}{7}\right], \\ 5(61 - 25\theta) & \text{if } \theta \geq \frac{3}{7}, \end{cases}$$

where

$$DWL = (1 - \theta)\left[\Pi(3) + 2\pi(3) + cs(3) - \Pi(1) - cs(1)\right] = \frac{5}{32}(7 - 3\theta).$$

This implies that the total welfare effect of prohibiting conditional pay-for-delay settlements is positive because the incumbent's gain from litigation must be higher than the total litigation cost for litigation to take place.

In *Servier*, the parties argued that it is pro-competitive to accommodate the excluded entrant if the patent is declared invalid following litigation by another entrant.[28] As our analysis shows, such logic is misleading, because the threat of rival entry reduces the cost of excluding the entrants, eliminating all competition in equilibrium.

Prohibiting simple pay-for-delay agreements

An entrant agreeing to a pay-for-delay contract commits not to enter the market. As a result, the expected number of firms in the market decreases, which lowers consumer surplus and total welfare. However, as shown in Proposition 4, such pay-for-delay contracts are strategically tied to licensing agreements signed with other entrants. Hence, the welfare effects of a prohibition of pay-for-delay contracts are more involved than they might appear at a first glance.

When the incumbent cannot conclude pay-for-delay contracts, either all entrants will receive a licensing agreement or one of them will litigate while the others wait. The gain from litigation is

[28] See, e.g., paragraph 1558 the European Commission decision in *Servier*.

proportional to the strength of the patent:

$$\bar{G}(\theta) \triangleq \theta \left[\Pi(1) - \Pi(n+1) - n\pi(n+1)\right]. \tag{2}$$

By comparing the gain described by equation (2) to $G(\theta)$, we obtain the following result:

Proposition 6. *Prohibiting simple pay-for-delay agreements increases the scope of litigation. There is litigation in equilibrium if $\bar{G}(\theta) > \mathcal{C}$, in which case one entrant litigates and all the others wait. If the reverse inequality holds, every entrant buys a license from the incumbent.*

Proof. Let us show that $\bar{G}(\theta) \geq G(\theta)$ for every θ. If $k(0) = 1$, then $G(\theta) \leq 0$, so $\bar{G}(\theta) \geq G(\theta)$ holds. If instead $k(0) > 1$, then $k(0) = k^* \leq n+1$ and we have

$$G'(0) = \Pi(1) - \Pi(k^*) - (k^* - 1)\pi(k^*) \leq \bar{G}'(0).$$

This implies $\bar{G}(\theta) \geq G(\theta)$ for every θ, because G is concave, $\bar{G}$ is linear and $G(0) = \bar{G}(0) = 0$. Finally, notice that $\bar{G}(1) > G(1) = 0$. $\square$

Prohibiting pay-for-delay agreements has the unintended consequence of increasing litigation because winning in court is the only way for the incumbent to monopolize the market. The alternative strategy is licensing the patent to all entrants, which leads to intense competition. Strong patents end up challenged in court and sufficiently weak patents are licensed. The scope of licensing increases in the costs of litigation.

Figure 2 uses our numerical example, for which $\bar{G}(\theta) = \frac{\theta}{16}$, to describe the outcome of the contracting game when pay-for-delay agreements are prohibited. The gray area shows the increase in the scope of litigation compared to the baseline with simple pay-for-delay contracts. Recall that, for patents of intermediate strength, $\theta \in \left(\frac{2}{7}, \frac{3}{7}\right)$, and high enough litigation costs, without the ban, the incumbent would license the patent to one of the entrants and exclude the other. In this case, banning pay-for-delay agreements can make consumers strictly worse off by making it relatively more profitable for the incumbent to pursue litigation instead of licensing.[29]

Proposition 7. *Banning simple pay-for-delay agreements may decrease the expected consumer surplus.*

[29] However, if licensing and pay-for-delay do not coexist in equilibrium, consumer surplus unambiguously increases.

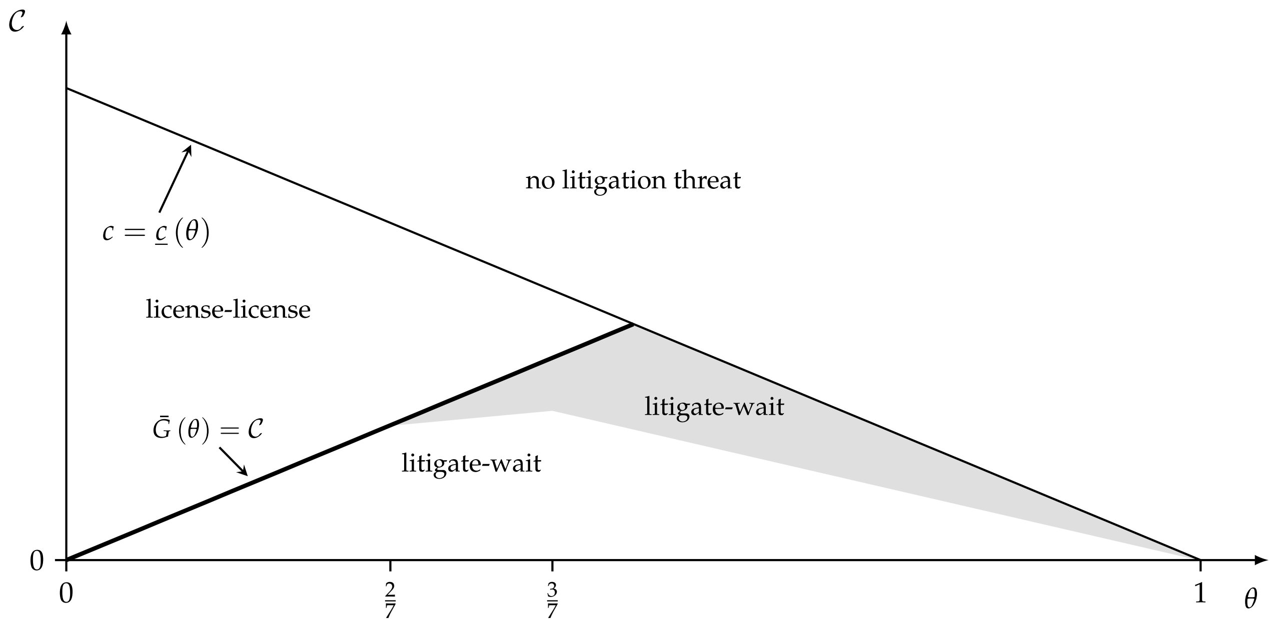

Figure 2: The equilibrium of the contracting game with two entrants when pay-for-delay agreements are prohibited, as a function of the strength of the patent, θ, and the total cost of litigation, $\mathcal{C}$. The gray area marks the increase in the scope of litigation.

Proof. Our numerical illustration provides an example. For any $\theta \in \left(\frac{17}{45}, \frac{3}{7}\right)$ we have

$$\bar{G}(\theta) - G(\theta) = \frac{7\theta - 2}{144} > 0.$$

For any $\mathcal{C} \in (G(\theta), \bar{G}(\theta))$ we then have that, without the ban, the incumbent licenses the patent to one entrant and excludes the other one, resulting in consumer surplus $cs(2) = \frac{2}{9}$. By contrast, if the ban is implemented, the incumbent litigates against one entrant while the other one waits, which results in expected consumer surplus

$$CS(L, W) = \theta cs(1) + (1-\theta) cs(3) = \frac{9 - 5\theta}{32},$$

which is less than $\frac{2}{9}$ by $\theta > \frac{17}{45}$. $\square$

The possibility of entering into pay-for-delay agreements may encourage the incumbent to license the patent. Licensing benefits consumers, even though the objective is inherently anti-competitive: this reduces the cost of excluding the other entrants. Due to this strategic link, the incumbent may stop licensing and pursue litigation following a ban on pay-for-delay agreements. In this case, the expected consumer surplus decreases if the strength of the patent is sufficiently high so that the incumbent is likely to win in court. We can conclude that, due to the interdependence 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.

There are two main limitations to the welfare analysis presented in this section. First, we do not account for innovation and investment incentives, which are typically assumed to depend on the payoffs of the firms, and therefore on the possibility of concluding pay-for-delay agreements. Modelling innovation and investment is outside of the scope of this paper and we leave it to future research. Second, we do not consider the public cost of litigation in the total welfare analysis. Such cost might include the costs of the court considering the case, as well as the opportunity cost of not ruling on other cases. A ban on pay-for-delay agreements increases the scope for litigation, therefore, the analysis presented here is a lower bound of the negative impact of such a policy on welfare.

6 Extensions

We analyze three extensions of our baseline model. First, we consider the case when the litigation outcomes are independent. Second, we analyze the game when the incumbent's offers are observed only by the entrant receiving the offer and not by others. Finally, we allow parties to renegotiate their contracts after litigation. For simplicity, throughout the section we assume that there are only two entrants, as this is enough to highlight the key differences to our baseline results.

To aid this comparison, we define thresholds $\{\alpha, \beta\}$ of patent strength by $f(1;\alpha) = f(2;\alpha)$ and $f(2;\beta) = f(3;\beta)$. These thresholds are given by

$$\alpha = \frac{\Pi(2) + \pi(3) - \Pi(1) + 2\pi(2)}{\pi(2) - \pi(3)} \text{ and } \beta = \frac{\Pi(3) - \Pi(2) - \pi(3)}{\pi(2) - \pi(3)}.$$

Selling a license to one of the entrants while paying to exclude the other one is optimal if the patent is of moderate strength: $\beta < \theta < \alpha$. For strong patents, i.e. when $\theta > \alpha$ and $\theta > \frac{1}{2}(\alpha + \beta)$, the incumbent prefers to exclude both entrants; whereas, weak patents satisfying $\theta < \beta$ and $\theta < \frac{1}{2}(\alpha + \beta)$ are licensed to both entrants. Furthermore, when there is litigation, the non-challenger signs a pay-for-delay agreement if β is negative and waits otherwise.

The thresholds α and β are also useful to distinguish three different modes of competition:

- *High monopoly profit*: $\alpha < 0$ and $\alpha + \beta < 0$. Even with simple contracts both entrants are always excluded from the market. This corresponds to Proposition 3;

- *High duopoly profit*: $\alpha > 0 > \beta$. The incumbent always excludes at least one of the entrants. The other entrant is excluded if the patent is sufficiently strong and buys a license if the patent is weak. For patents of intermediate strength that entrant litigates;

- *High triopoly profit*: $\beta > 0$ and $\alpha + \beta > 0$. Both entrants are excluded if the patent is sufficiently strong and both buy a license if the patent is weak. For not-so-weak nor strong patents, one entrant is excluded and the other buys a license if $\alpha > \beta$. For patents of intermediate strength, one entrant litigates and the other waits.

Notice that our textbook Cournot quantity-setting example belongs to the third category with $\alpha = \frac{3}{7}$ and $\beta = \frac{2}{7}$. When convenient, we revisit this example to further illustrate our results.

Independent litigation outcomes

In the baseline scenario, we assume that litigation outcomes are perfectly correlated. We think that in the model of patent invalidation, this is a natural assumption. However, patent litigation can also take the form of an infringement claim. In such a case, the entrants are independently attempting to show that their technologies are not infringing upon the patent. This scenario is the focus of this section.

With independent litigation outcomes, there might be two lawsuits, and the market structure will depend on the outcomes of them both. The incumbent can win against both, win against one and lose against the other, or lose against both entrants. Furthermore, the entrants can no longer free-ride on each other's litigation efforts: waiting always results in zero payoff. From the entrants' perspective, this makes litigation more profitable, because there is a possibility of being the only winner against the incumbent. Each entrant now receives litigation profit of:

$$\hat{c}(\theta) = (1-\theta)\left[\theta\pi(2) + (1-\theta)\pi(3)\right].$$

Note, the litigation threat can be credible even if the entrants would find litigation too costly with perfectly correlated litigation outcomes. To compare our results, we focus on the case $c < \underline{c}(\theta)$, which corresponds to a credible litigation threat in the baseline model (we discuss the other case below). Interestingly, when we move from a perfect correlation to independent litigation outcomes, there is no longer litigation in equilibrium. However, the outcome without litigation stays the same:

Proposition 8. *Suppose that litigation outcomes are independent and $c < \underline{c}(\theta)$. Then, the equilibrium outcome corresponds to the equilibrium without litigation in the baseline model under simple contracts.*

Proof. See Appendix B. □

Hence, there is no litigation in the equilibrium, but the main observation that licensing and pay-for-delay settlements are interlinked carries over. When both entrants settle in equilibrium, their reservation utilities are determined by the litigation payoff each would obtain by rejecting the incumbent's deal and going to court while the other one settles. The degree of correlation between litigation outcomes plays no role, when there is only one lawsuit. Thus, the incumbent's payoff from settling with both entrants is exactly the same as in the baseline model.

However, the incumbent's payoff from litigation decreases. First, when it fights both entrants in court, the chance of monopolizing the market is reduced by $\theta(1-\theta)$ as compared to litigating against one entrant while the other one free-rides on this effort, which is a possibility under perfectly correlated lawsuits. Second, the incumbent must pay a higher reverse payment to exclude the non-challenger if only one entrant litigates, because the excluded entrant's reservation utility is now higher due to the chance of being the only winner against the incumbent.

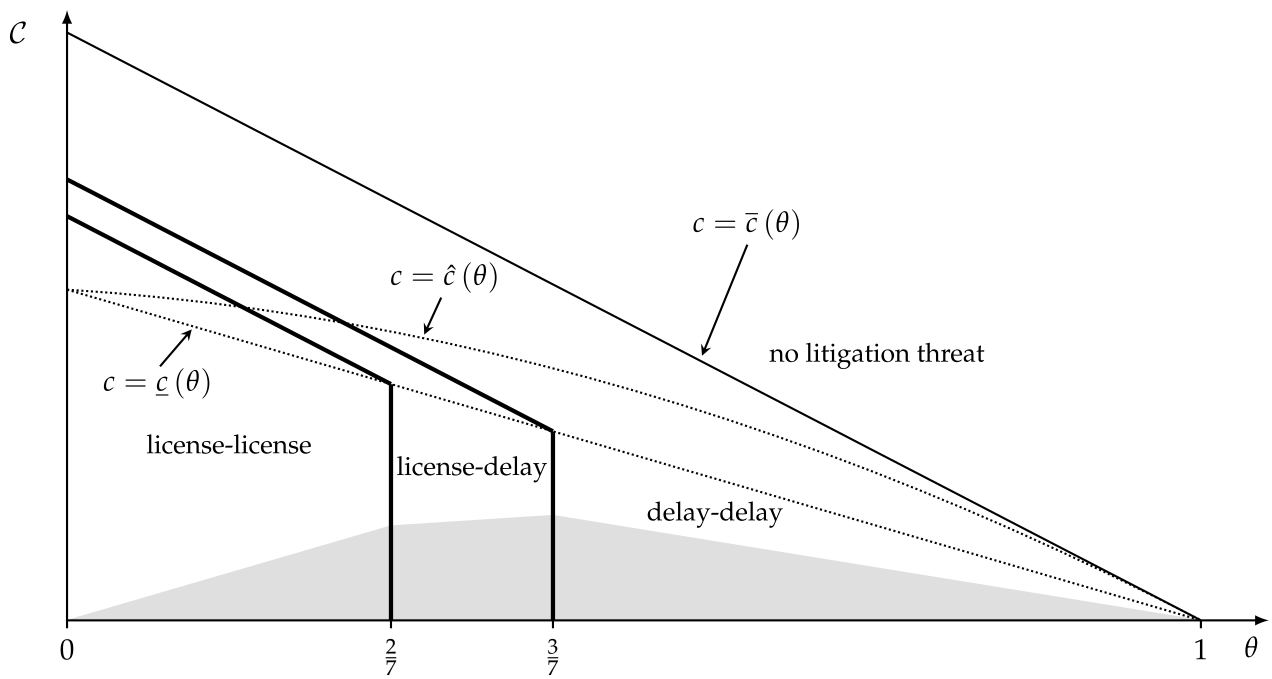

Figure 3: Equilibrium of the game with independent litigation outcomes, as a function of the strength of the patent, θ, and the total cost of litigation, $\mathcal{C}$. Profits are as in the symmetric Cournot quantity-setting game with zero marginal costs and consumer valuations uniformly drawn from the unit interval. The gray area marks the scope of litigation with perfectly correlated litigation outcomes.

Finally, let us briefly discuss the other case $c \geq \underline{c}(\theta)$, which corresponds to no credible litigation threat in the baseline model. With independent litigation outcomes, instead, the incumbent can monopolize the market with zero cost only if $c \geq \bar{c}(\theta)$, where $\bar{c}(\theta) = (1-\theta)\pi(2)$ is the expected profit a challenger obtains if the other entrant is excluded from the market. Otherwise, for $c < \hat{c}(\theta)$ litigation is profitable even if the other entrant litigates, and if the reverse inequality holds, it is still profitable if the other entrant is excluded. Using our numerical example, Figure 3 provides a complete characterization of the equilibrium with independent litigation outcomes.

Secret offers

In this section, we relax the assumption that settlement offers are publicly observable. Instead, we will assume that each entrant does not observe if the other entrant receives a pay-for-delay agreement, a license, or waits, but does however observe litigation. A registry of pending court cases is typically publicly available; hence, it is natural to consider litigation as observed by all firms.

Analyzing secret offers, we must take a stance on how cautious the entrants are towards unexpected (out-of-equilibrium) offers from the incumbent; particularly, when they do not observe litigation against the other entrant. In fact, the entrants have a reason to be suspicious, simply because the deal made with one entrant affects the incumbent's own profit as well as the profitability of contracting with the other entrant. Thus, upon receiving an unexpected offer from the incumbent, it is reasonable for each entrant to anticipate that also the other entrant received a changed offer.[30]

To formally model how beliefs change when receiving an out-of-equilibrium offer, we adopt the notion of "wary beliefs" introduced by McAfee and Schwartz (1994), with a slight modification due to observable litigation. When an entrant receives an offer from the incumbent, it believes the following: a) the incumbent expects it to accept the offer; b) given this offer, the incumbent offers the other entrant a settlement that is best for the incumbent, among all acceptable settlements to the other entrant, or litigates; and c) the other entrant reasons the same way.

Under these assumptions on the beliefs, the baseline analysis with conditional deals is robust to secret contracting.[31] Given that one entrant accepts a conditional pay-for-delay agreement, the incumbent maximizes the industry profit shared with the other entrant by offering the same deal. Furthermore, the incumbent obtains the highest possible share of the monopoly profit, as the entrant is indifferent between accepting the reverse payment or rejecting the deal.

However, for the same reason, the analysis of simple contracts will change: the incumbent wants

[30] A similar justification for wary beliefs is provided by Nocke and Rey (2018).

[31] This is true even with secret litigation, as with public offers there is no litigation in equilibrium.

to secretly stop licensing whenever the other entrant is excluded. Proposition 9 describes the equilibrium of the game with simple, secret offers.[32]

Proposition 9. *Suppose that litigation is publicly observable but offers are otherwise secret and the entrants hold wary beliefs. Then, if conditional pay-for-delay agreements are feasible, the equilibrium outcome is the same as with the public offers. With simple contracts, the gain from litigation increases to:*

$$\tilde{G}(\theta) = G(\theta) + [\pi(2) - \pi(3)] \cdot \begin{cases} \max\{0, \alpha - \theta\} & \text{if } \beta < 0 \\ \max\{0, \min\{\theta - \beta, \alpha - \theta\}\} & \text{otherwise.} \end{cases}$$

There is litigation in equilibrium if $\tilde{G}(\theta) > C$, in which case the non-challenger is excluded if $\beta < 0$ and waits otherwise, as with public offers. If the reverse inequality holds, both entrants buy a license if $\theta < (\alpha + \beta)/2$ and $\beta \geq 0$, and otherwise both of them are excluded.

Proof. See Appendix B. □

The intuition for the result is straightforward. When settlement offers are secret, the incumbent cannot credibly use licensing to reduce the cost of excluding the other entrant. It has an incentive to secretly deviate by offering a pay-for-delay agreement to the licensee (the monopoly profit is greater than the duopoly industry profit). Therefore, an entrant receiving a pay-for-delay offer will not believe that the other entrant is getting a licensing agreement. Thus, the only way for the incumbent to reduce the reverse payment is to pursue litigation, which explains why the incumbent litigates more often.

In Figure 4 we use our numerical illustration to show how the contracting outcome with simple contracts changes when we move from public to secret offers (with observable litigation). The main observation is that licensing and pay-for-delay no longer coexist and that the scope of litigation increases. The incumbent can no longer use licensing to reduce the cost of excluding the other entrant, thus will resort to litigation more often.

Renegotiation

In this section, we allow the firms to renegotiate contracts: we are interested in analyzing the impact of renegotiation on contracting externalities and the equilibrium market entry. We allow the parties to renegotiate after a court decides whether or not the patent is valid. We assume that, after the patent has been invalidated, the parties can no longer enter into pay-for-delay agreements. Having

[32] With passive beliefs there is no litigation in equilibrium only if both entrants are excluded. See our Appendix online.

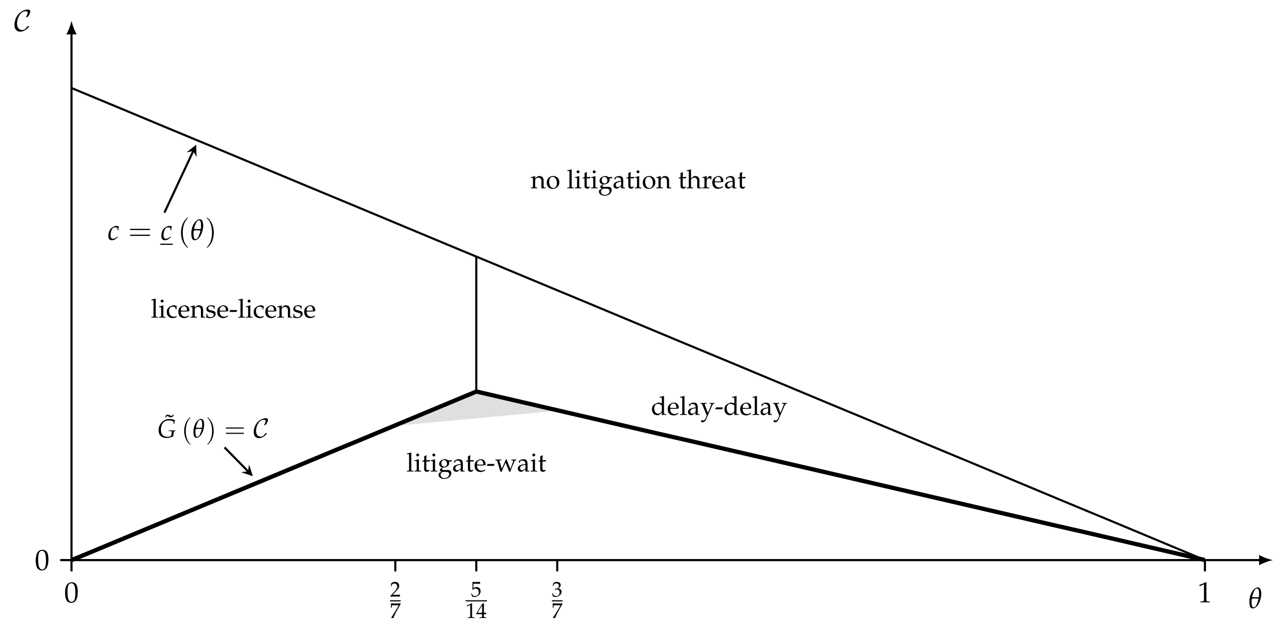

Figure 4: Equilibrium of the game with simple contracts and secret contracting for two entrants, as a function of the strength of the patent, θ, and the total cost of litigation, $\mathcal{C}$. Profits are as in the symmetric Cournot quantity-setting game with zero marginal costs and consumer valuations uniformly drawn from the unit interval. The gray area marks the (small) increase in the scope of litigation.

no patent, the incumbent cannot exclude actual competitors by paying for their exit without violating antitrust laws. We thus discard the possibility to exclude active entrants and consider a case, where renegotiation means accommodating an entrant that would otherwise stay out of the market (due to a pay-for-delay agreement signed before patent litigation).[33]

Specifically, at the renegotiation stage, the incumbent can make two types of offers to the entrants. It can either propose to follow the original agreement or offer an entry deal with a new payment. We consider a two-stage renegotiation game, where the incumbent first commits to publicly observable bilateral offers, and the entrants simultaneously decide either to accept their respective offers or reject the offer and invoke the original agreement. As in the baseline model, without loss of generality, we can focus on equilibria in which each entrant accepts the new offer.

In the incumbent's preferred subgame perfect equilibrium, the entrants must be indifferent between accepting the new offer and sticking to the original agreement. Thus, each entrant receives the utility it would obtain from the original deal. If the original deal allows for entry, the reservation

[33] Ex ante, an entrant is only a potential competitor and the incumbent has intellectual property rights, so the legal situation is less clear, albeit the European Commission considers pay-for-delay agreements as restrictions to competition by object. In our online Appendix, we relax this assumption and allow for pay-for-delay contracts at the renegotiation stage. We find that the space for pay-for-delay contracts increases, but the observation that litigation happens for patents of intermediate strength carries over.

utility thus equals the entrant's profit (accounting for renegotiation between the other entrant and the incumbent) net of the original payment. If the entrant stays out without renegotiation, its reservation utility equals the reverse payment initially agreed with the incumbent.[34]

At the renegotiation stage, if both entrants are out, the incumbent has no incentive to renegotiate; the monopoly outcome maximizes its payoff, which is the sum of industry profit and a constant term capturing the initial payments. If instead one entrant is active and the other one is out, the incumbent can change the original outcome by also accommodating the excluded entrant. This is profitable if their joint triopoly profit is higher than the incumbent's duopoly profit: $\Pi(3) + \pi(3) > \Pi(2)$ or equivalently $\beta > 0$. Finally, if both entrants are active, there is no scope for renegotiation, as the incumbent can only replicate the original outcome.

The firms anticipate renegotiation at the initial stage of contracting. If this results in successful litigation by one of the entrants and $\beta > 0$ holds, the possibility of renegotiation will have a direct impact on the profits made by the incumbent. Otherwise, the prospect of renegotiation only influences the reservation utilities of the entrants, indirectly affecting the incumbent's payoff. However, as Proposition 10 shows, this indirect effect plays an important role in equilibrium.

Proposition 10. *Suppose that the incumbent can renegotiate contracts by offering licensing deals. Then, the baseline analysis with simple contracts is renegotiation-proof, unless $\beta > 0$, in which case the incumbent excludes both entrants by offering them low reverse payments, corresponding to the outcome with conditional pay-for-delay agreements.*

Proof. See Appendix B. □

The possibility to renegotiate agreements signed before patent invalidation reduces equilibrium entry into the market. The underlying mechanism is similar to that for conditional pay-for-delay agreements: litigation becomes less profitable when the incumbent can let more entry to the market after the patent has been invalidated. As a result, entrants will accept lower reverse payments to forego litigation and stay out of the market.

To conclude, in all extensions considered here, we find that contracting externalities prevent the parties from achieving a jointly efficient outcome for some parameter values. Accommodating entry to the market enables the incumbent to decrease reverse payments needed to keep other entrants

[34] Alternatively, we could interpret the payments as sunk, in which case the entrants' reservation utility at the renegotiation stage would equal its profit if active in the market and zero otherwise. This would not affect the analysis, however.

out. This basic logic underlying the results of the baseline model largely carries over to the three environments analyzed in this section.

7 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 substantial economic impact. We propose a model of patent settlements in an environment with multiple entrants. This approach allows us to study an important phenomenon that is new to the literature on pay-for-delay: settlement externalities. A settlement defining the date of entry of an entrant impacts the incentives to enter the market for other firms. If a settlement excludes an entrant from the market, 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). Shapiro (2003) shows that in an environment with a single entrant, parties will agree to delay entry and share monopoly rent. Allowing for multiple entrants highlights an important externality of such an agreement: it makes entering the market more attractive to all other potential entrants. 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: that of 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 into the market can be an outcome of a contracting game between an incumbent and several entrants arriving simultaneously. We, also, find that litigation occurs for patents of intermediate strength.

This logic hinges on the contracting environment; that is, when the incumbent can condition the validity of pay-for-delay settlements on litigation outcomes with other entrants. Such contracts drastically decrease the cost of entry delay by undermining the incentives to deviate from an equilibrium in which all entrants are delayed. When conditional contracts are feasible, even a weak patent allows the incumbent to delay the entry of all entrants.

This article contributes to the debate on the economic consequences of pay-for-delay agreements. First, we show that settlements which are conditional on patent validity should not be allowed if the goal of the policy-maker is to promote entry to the market. Second, we show that a ban on simple pay-for-delay deals has more nuanced consequences than in an environment with just a single entrant (Shapiro, 2003; Elhauge and Krueger, 2012). Models with one entrant provide useful, information-light rules to guide antitrust enforcement concerning pay-for-delay settlements. 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, a ban on pay-for-delay settlements might decrease 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 this legislation should account for incentives to innovate by generics, which was outside the scope of our analysis.

We provide several extensions and robustness checks of the baseline model. We separately consider independent litigation outcomes, the possibility of renegotiating the settlements, and secret offers. These extensions impact the equilibrium outcome, however, the main logic of settlement externalities is present in all of them. We have avoided introducing asymmetric information even though the beliefs about the strength of the patent are critical factors in agreeing on a settlement. Thus we believe that a 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.[35]

[35]This could be done by allowing entry at risk during the litigation period in our continuous time model presented in

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.

Bokhari, F. A., Mariuzzo, F., and Polanski, A. (2020). Entry limiting agreements: First-mover advantage, authorized generics, and pay-for-delay deals. *Journal of Economics & Management Strategy*.

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.

Crampes, C. and Langinier, C. (2002). Litigation and settlement in patent infringement cases. *RAND Journal of Economics*, pages 258–274.

Cunningham, C., Ederer, F., and Ma, S. (2021). Killer acquisitions. *Journal of Political Economy*, 129(3):649–702.

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

Diekmann, A. (1985). Volunteer's dilemma. *Journal of conflict resolution*, 29(4):605–610.

Ecer, S., Montes, R., and Weiskopf, D. (2020). On the application of nash bargaining in reverse payment cases in the pharmaceutical industry. *European Journal of Law and Economics*, 50(1):133–147.

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.

Encaoua, D. and Lefouili, Y. (2009). Licensing 'weak' patents. *The Journal of Industrial Economics*, 57(3):492–525.

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

Goeree, J. K., Holt, C. A., and Smith, A. M. (2017). An experimental examination of the volunteer's dilemma. *Games and Economic Behavior*, 102:303–315.

Appendix A.

Gratz, L. (2012). Economic analysis of pay-for-delay settlements and their legal ruling. *Available at SSRN 1979699*.

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

Lemus, J. and Temnyalov, E. (2020). Pay-for-delay with follow-on products. *Review of Industrial Organization*, pages 1–18.

Manganelli, A.-G. (2019). Reverse payments and risk of bankruptcy under private information. *International Journal of the Economics of Business*, 26(3):411–429.

Manganelli, A.-G. (2021). Reverse payments, patent strength, and asymmetric information. *Health Economics*, 30(1):20–35.

Marxen, A. and Montez, J. (2020). Licensing at the patent cliff and market entry. *Available at CEPR Discussion Paper No. DP14276*.

McAfee, R. P. and Schwartz, M. (1994). Opportunism in multilateral vertical contracting: Nondiscrimination, exclusivity, and uniformity. *The American Economic Review*, pages 210–230.

Nocke, V. and Rey, P. (2018). Exclusive dealing and vertical integration in interlocking relationships. *Journal of Economic Theory*, 177:183–221.

Padilla, J. and Meunier, V. (2015). Should reverse payment patent settlements be prohibited per se? *Available at SSRN 2604071*.

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

A Continuous time with a litigation period

In this Appendix, we consider a setting where the patent runs from time 0 to time 1, litigation is time-consuming, and parties can agree on intermediate entry dates. The incumbent's trade with entrant i is denoted by x_i in $[0,1] \cup \{L, W\}$, where $x_i \in [0,1]$ denotes the date of entry, L litigation, and W waiting. Litigation results in entry at date $\lambda \in [0,1]$ with probability $1-\theta$ and date one otherwise. Waiting results in the same entry date as litigation if another entrant litigates and to date one otherwise. Litigation outcomes are perfectly correlated.

For any trade profile, $\mathbf{x} = (x_1, \dots, x_n)$, we denote by $q_i(\mathbf{x}; \tau)$ the random variable that indicates whether i is active in the market at date τ. We define $Q(\mathbf{x}; \tau) = 1 + \sum_i q_i(\mathbf{x}; \tau)$ and let $\mathbb{E}[Q(\mathbf{x}; \tau)]$ be the expected number of firms in the market at that date. For a given realization k of $Q(\mathbf{x}; \tau)$, the incumbent makes an instantaneous profit $\Pi(k)$ and each active entrant makes a profit $\pi(k)$, whereas the entrants that are out make zero profit. As in the baseline, the profits are all positive and total industry profit decreases in k. One unit of money at time 0 is worth $e^{-\delta\tau}$ units at time τ for all firms.

Using the linearity of expectations, entrant i's expected overall profit satisfies

$$\mathbb{E}\left[\int_0^1 e^{-\delta\tau} \pi\left(Q(\mathbf{x}; \tau)\right) q_i(\mathbf{x}; \tau) \, \mathrm{d}\tau\right] = \int_0^1 e^{-\delta\tau} \mathbb{E}\left[\pi\left(Q(\mathbf{x}; \tau)\right) q_i(\mathbf{x}; \tau)\right] \mathrm{d}\tau.$$

Entrant's payoff is the profit from equation above net of the payment t_i and cost of action $c(x_i)$.

Similarly, the incumbent's payoff is the sum of its overall expected profit and the total payment $\sum_i t_i$ net of its cost of action $C(\mathbf{x})$. The costs are defined as in the baseline model, and the incumbent's expected profit is

$$\mathbb{E}\left[\int_0^1 e^{-\delta\tau} \Pi\left(Q(\mathbf{x}; \tau)\right) \mathrm{d}\tau\right] = \int_0^1 e^{-\delta\tau} \mathbb{E}\left[\Pi\left(Q(\mathbf{x}; \tau)\right)\right] \mathrm{d}\tau.$$

The incumbent sells entry or licenses the patent if $x_i \leq \lambda$ and $t_i > 0$. On the contrary, we say that the incumbent buys exit or signs a pay-for-delay agreement if $x_i > \lambda$ with a "reverse payment" $t_i < 0$. As in the baseline model, $x_i = W$ with $t_i < 0$ defines conditional pay-for-delay agreements, and the entrants have two outside options: $x_i \in \{L, W\}$ with zero payment.

The expected producer surplus, $PS(\mathbf{x})$, equals the overall expected industry profit net of the total cost of action. This can be conveniently expressed as

$$PS(\mathbf{x}) = \int_0^1 PS(\mathbf{x};\tau)\,\mathrm{d}\tau,$$

where $PS(\mathbf{x};\tau)$ is the instantaneous expected producer surplus:

$$PS(\mathbf{x};\tau) = e^{-\delta\tau}\mathbb{E}\left[\Pi\left(Q(\mathbf{x};\tau)\right) + \left(Q(\mathbf{x};\tau) - 1\right)\pi\left(Q(\mathbf{x};\tau)\right)\right] - C(\mathbf{x}) - \sum_i c(x_i).$$

Note that costs of action do not depend on τ so we can drop the integrals. As in the baseline model, $PS(\mathbf{x};\tau)$ for any τ and therefore $PS(\mathbf{x})$ is maximized in monopoly with no litigation, which can be achieved by any combination of waiting and pay-for-delay agreements with full entry delay.

The two-stage game, as defined in the baseline model, is played at date zero. Without loss of generality, we can restrict attention to equilibria in which every entrant accepts its offer, because the incumbent can always offer $\{L, W\}$ with a zero payment. We have a second-stage Nash equilibrium if and only if for each i the following participation constraint is satisfied:

$$\int_0^1 e^{-\delta\tau}\mathbb{E}\left[\pi\left(Q(\mathbf{x};\tau)\right)q_i(\mathbf{x};\tau)\right]\mathrm{d}\tau - c(x_i) - t_i \geq r_i(\mathbf{x}_{-i}),$$

where

$$r_i(\mathbf{x}_{-i}) = \max_{x_i\in\{L,W\}} \int_\lambda^1 e^{-\delta\tau}\mathbb{E}\left[\pi\left(Q(x_i,\mathbf{x}_{-i};\tau)\right)q_i(x_i,\mathbf{x}_{-i};\tau)\right]\mathrm{d}\tau - c(x_i).$$

In the incumbent's preferred subgame perfect equilibrium, all participation constraints must bind. Substituting transfers from the binding constraints into the incumbent's payoff, we obtain the incumbent's objective

$$\int_0^1 e^{-\delta\tau}\mathbb{E}\left[\Pi\left(Q(\mathbf{x};\tau)\right)\right]\mathrm{d}\tau + \sum_i t_i - C(\mathbf{x}) = PS(\mathbf{x}) - \sum_i r_i(\mathbf{x}_{-i}).$$

We directly obtain the counterparts of Propositions 1 and 2 by observing that $(W,\ldots,W)$ maximizes $PS(\mathbf{x})$ and for each i

$$r_i(W,\ldots,W) = \max\{\underline{c}(\theta;\lambda) - c, 0\} \leq r_i(\mathbf{x}_{-i}),$$

where

$$\underline{c}(\theta;\lambda) = \int_\lambda^1 e^{-\delta\tau}\mathrm{d}\tau\cdot(1-\theta)\,\pi(1+n).$$

In particular, $\underline{c}(\theta;0)$ with $\delta = 0$ corresponds to the definition of $\underline{c}(\theta)$ in the baseline model.

To show that also Propositions 3 and 4 carry over, from now on let us restrict attention to simple contracts and assume $c < \underline{c}(\theta;\lambda)$. First, under this assumption, the entrants always have positive litigation payoffs, which implies that each entrant i's best outside option is to wait if no other entrant litigates and to litigate otherwise, regardless of the number of firms in the market at given point in time. Thus, we can take the maximum in $r_i(\mathbf{x}_{-i})$ inside the integral and write

$$r_i(\mathbf{x}_{-i}) = \int_{\lambda}^{1} r_i(\mathbf{x}_{-i};\tau)\,\mathrm{d}\tau,$$

where

$$r_i(\mathbf{x}_{-i};\tau) = \max_{x_i \in \{L,W\}} e^{-\delta\tau}\mathbb{E}\left[\pi\left(Q\left(x_i,\mathbf{x}_{-i};\tau\right)\right) q_i\left(x_i,\mathbf{x}_{-i};\tau\right)\right] - \frac{1}{1-\lambda}c(x_i).$$

The incumbent's objective can be separated into two terms: the payoff during the litigation period and the payoff afterwards. We have:

$$PS(\mathbf{x}) - \sum_i r_i(\mathbf{x}_{-i}) = \int_0^{\lambda} PS(\mathbf{x};\tau)\,\mathrm{d}\tau + \int_{\lambda}^{1}\left[PS(\mathbf{x};\tau) - \sum_i r_i(\mathbf{x}_{-i};\tau)\right]\mathrm{d}\tau.$$

As the incumbent gets the entire producer surplus during the litigation period, it is never optimal to allow entry before λ. Thus, we can restrict attention to entry dates inside the interval $[\lambda, 1]$.

When there is no litigation, there is no uncertainty regarding entry to the market. We let $Q(\mathbf{x};\tau) = k(\tau)$ denote the number of active firms in the market at date $\tau \in [\lambda, 1]$, comprising of $k(\tau) - 1$ entrants with $x_i \leq \tau$ and the incumbent. The instantaneous reservation utility of entrant i at date $\tau \in [\lambda, 1]$ is given by

$$r_i(\mathbf{x}_{-i};\tau) = -c + e^{-\delta\tau}(1-\theta)\cdot\begin{cases}\pi(k(\tau)) & \text{if } x_i \leq \tau,\\ \pi(k(\tau)+1) & \text{if } x_i > \tau.\end{cases}$$

By subtracting the sum of these rents from the instantaneous producer surplus, we obtain

$$PS(\mathbf{x};\tau) - \sum_i r_i(\mathbf{x}_{-i};\tau) = e^{-\delta\tau} f(k(\tau);\theta) + nc,$$

which is maximized by $k(\theta) = \arg\max_k f(k;\theta)$ independently of τ. Hence, $k(\theta) - 1$ entrants sign licensing deals with $x_i = \lambda$ and $n + 1 - k(\theta)$ are excluded with full entry delay.

Suppose now that one of the entrants litigates. Then, any rival entrant who rejects its offer will wait and free-ride on the litigation effort. The payoff from waiting thus pins down reservation utilities:

$$r_i\left(\mathbf{x}_{-i};\tau\right)=e^{-\delta\tau}\left(1-\theta\right)\cdot\begin{cases}\pi\left(k_0\left(\tau\right)\right) & \text{if } x_i\leq\tau,\\ \pi\left(k_0\left(\tau\right)+1\right) & \text{if } x_i>\tau,\end{cases}$$

where $k_0(\tau)$ denotes the number of active firms in the market at date $\tau \geq \lambda$ if the patent is declared invalid. By subtracting the sum of these rents from the expected instantaneous producer surplus, we obtain

$$PS\left(\mathbf{x};\tau\right)-\sum_i r_i\left(\mathbf{x}_{-i};\tau\right)=\theta e^{-\delta\tau}\left[\Pi\left(k_1\left(\tau\right)\right)+\left(k_1\left(\tau\right)-1\right)\pi\left(k_1\left(\tau\right)\right)\right]+\left(1-\theta\right)e^{-\delta\tau}f\left(k_0\left(\tau\right);0\right)-C,$$

where $k_1(\tau)$ denotes the number of active firms in the market if the patent is upheld by the court. As the industry profit is maximized in monopoly, it is optimal for the incumbent to choose $k_1(\tau) = 1$ and $k_0(\tau) = k^*$, where $k^* = \arg\max_k f(k;0)$ independently of τ. Thus, one entrant litigates, $k^* - 2$ entrants wait and $n - k^* + 1$ sign pay-for-delay agreements with full entry delay.

The incumbent's payoffs when it settles with all entrants and when it pursues litigation coincide if and only if the expected gain from litigation, $\int_\lambda^1 e^{-\delta\tau}\mathrm{d}\tau \cdot G(\theta)$, where $G(\theta)$ is the same as in the baseline model, equals the total litigation cost $C + nc$. By denoting

$$\mathcal{C}=\frac{C+nc}{\int_\lambda^1 e^{-\delta\tau}\mathrm{d}\tau},$$

we can then apply Propositions 3 and 4 directly.

B Proofs

Proof of Proposition 8

Each entrant's reservation utility is given by its litigation payoff, which is strictly positive by $c < \underline{c}(\theta)$, whereas the payoff from waiting is zero with independent litigation outcomes. Let us first show that the incumbent is always worse off by litigating against both entrants as opposed to litigating only against one of them. By litigating against both entrants, the incumbent obtains $f_{LL}(\theta) - C$, where

$$f_{LL}\left(\theta\right)=\theta^2\left[\Pi\left(1\right)+\Pi\left(3\right)-2\Pi\left(2\right)\right]+2\theta\left[\Pi\left(2\right)-\Pi\left(3\right)\right]+\Pi\left(3\right).$$

The payoff from litigating only against i is the maximum of two payoffs: the incumbent can either license to j or exclude it. The payoff from these strategies is

$$-\left[\hat{c}(\theta)-c\right]-C+\begin{cases}\theta\left[\Pi(2)+\pi(2)\right]+(1-\theta)\left[\Pi(3)+\pi(3)\right] & \text{if } x_j=1,\\ \theta\Pi(1)+(1-\theta)\Pi(2) & \text{if } x_j=0,\end{cases}$$

where $\hat{c}(\theta)-c$ is j's reservation utility. The maximum of these payoffs is $f_L(\theta)+c-C$, where

$$\begin{aligned}f_L(\theta)&=\left[\pi(2)-\pi(3)\right]\\&+\begin{cases}\theta\left[\Pi(2)-\Pi(3)+\pi(3)\right]+\Pi(3) & \text{if } \theta(1+\beta-\alpha)\leq\beta,\\ \theta\left[\Pi(1)-\Pi(2)+2\pi(3)-\pi(2)\right]+\Pi(2)-\pi(3) & \text{if } \theta(1+\beta-\alpha)\geq\beta,\end{cases}\end{aligned}$$

and $\{\alpha,\beta\}$ are defined as in the text:

$$\alpha=\frac{\Pi(2)+\pi(3)-\Pi(1)+2\pi(2)}{\pi(2)-\pi(3)} \text{ and } \beta=\frac{\Pi(3)-\Pi(2)-\pi(3)}{\pi(2)-\pi(3)}.$$

Notice that $1+\beta-\alpha>\beta$ by the assumption that the industry profit is decreasing in the number of firms in the market. When θ is low, the incumbent sells a license to j, and otherwise it is excluded. We have:

$$0\leq f_L(\theta)-f_{LL}(\theta)=\begin{cases}\theta\left[\beta-\theta(1+\beta-\alpha)\right] & \text{if } \theta(1+\beta-\alpha)\leq\beta,\\ (1-\theta)\left[\theta(1+\beta-\alpha)-\beta\right] & \text{if } \theta(1+\beta-\alpha)\geq\beta.\end{cases}$$

Let us now show that $f(k(\theta);\theta)\geq f_L(\theta)$ for all θ, which is sufficient for $f(k(\theta);\theta)+2c\geq f_L(\theta)-C+c$. With two entrants:

$$f(k(\theta);\theta)=\begin{cases}\theta 2\pi(3)+\Pi(3) & \text{if } \theta\leq\beta \text{ and } \theta\leq\frac{1}{2}(\alpha+\beta),\\ \theta\left[\pi(2)+\pi(3)\right]+\Pi(2)-\pi(3) & \text{if } \theta\in[\beta,\alpha],\\ \theta 2\pi(2)+\Pi(1)-2\pi(2) & \text{if } \theta\geq\alpha \text{ and } \theta\geq\frac{1}{2}(\alpha+\beta).\end{cases}$$

Assuming first $\theta(1+\beta-\alpha) \leq \beta$, we have

$$f(k(\theta);\theta) - f_L(\theta) = [\pi(2) - \pi(3)] \cdot \begin{cases} \theta(\beta-\theta) & \text{if } \theta \leq \beta \text{ and } \theta \leq \frac{1}{2}(\alpha+\beta), \\ (1-\theta)(\theta-\beta) & \text{if } \theta \in [\beta,\alpha], \\ 2\theta - (\alpha+\beta) + \theta(\beta-\theta) & \text{if } \theta \geq \alpha \text{ and } \theta \geq \frac{1}{2}(\alpha+\beta), \end{cases}$$

which is clearly positive if $\theta \leq \alpha$ or $\theta \leq \frac{1}{2}(\alpha+\beta)$. Furthermore, $\theta > \alpha$ and $\theta > \frac{1}{2}(\alpha+\beta)$ is equivalent to $2\theta > \alpha+\beta$ if $\alpha \leq \beta$. In this case, $\theta(1+\beta-\alpha) \leq \beta$ together with $\alpha \leq \beta$ implies $\beta > \theta$, which together with $2\theta > \alpha+\beta$ implies

$$2\theta - (\alpha+\beta) + \theta(\beta-\theta) > 0.$$

If instead $\alpha > \beta$, then $\theta > \alpha$ and $\theta > \frac{1}{2}(\alpha+\beta)$ is equivalent to $\theta > \alpha$, which together with $\alpha > \beta$ implies

$$2\theta - (\alpha+\beta) + \theta(\beta-\theta) = \theta - \alpha + (1-\theta)(\theta-\beta) > 0.$$

Suppose now that $\theta(1+\beta-\alpha) \geq \beta$. Then, we have

$$f(k(\theta);\theta) - f_L(\theta) = [\pi(2) - \pi(3)] \cdot \begin{cases} \beta - \theta + \theta(\alpha-\theta) & \text{if } \theta \leq \beta \text{ and } \theta \leq \frac{1}{2}(\alpha+\beta), \\ \theta(\alpha-\theta) & \text{if } \theta \in [\beta,\alpha], \\ (1-\theta)(\theta-\alpha) & \text{if } \theta \geq \alpha \text{ and } \theta \geq \frac{1}{2}(\alpha+\beta), \end{cases}$$

which is obviously positive if $\theta \geq \beta$ or $\theta \geq \frac{1}{2}(\alpha+\beta)$. Similarly as above, $\theta < \beta$ and $\theta < \frac{1}{2}(\alpha+\beta)$ is equivalent to $2\theta < \alpha+\beta$ if $\alpha \leq \beta$. These imply

$$\beta - \theta + \theta(\alpha-\theta) \geq (1-\theta)(\beta-\theta) > 0.$$

Furthermore, $\theta < \beta$ and $\theta < \frac{1}{2}(\alpha+\beta)$ is equivalent to $\theta < \beta$ if $\alpha > \beta$. But then, $\theta(1+\beta-\alpha) \geq \beta$ together with $\alpha > \beta$ implies $\theta > \beta$, a contradiction to $\theta < \beta$. We thus conclude that $f(k(\theta);\theta) \geq f_L(\theta)$ for all θ. Therefore, the incumbent's payoff from the optimal settlement strategy is always higher than the payoff from litigation.

Finally, let us consider the other case, when $c \geq \underline{c}(\theta)$ holds. Then, if $c < \hat{c}(\theta)$, the incumbent's payoff from litigating against both entrants stays the same. However, its payoff from litigating againt one entrant reduces, because now it has to offer money to i for it to litigate while j buys a license. Therefore, $f_L(\theta) + c - C$ gives an upper bound for the incumbent's litigation payoff. By contrast, $f(k(\theta); \theta) + 2c$ now gives a lower bound for the incumbent's payoff from no litigation, because $c \geq \underline{c}(\theta)$ implies that the reservation utility of an entrant is zero when the other one buys a license. Hence, the incumbent is able to extract a larger part of the industry profit when settling with both, so the fact that there is no litigation when $c < \underline{c}(\theta)$ holds also implies that there is no litigation when $\underline{c}(\theta) \leq c < \hat{c}(\theta)$ is true. The same logic applies when $c > \hat{c}(\theta)$, because this only reduces the incumbent's payoff from litigating against both entrants, as it must pay them to go to court.

Proof of Proposition 9

If conditional contracts are allowed, the equilibrium outcome under secret contracts coincides with the equilibrium under public offers. Upon receiving a conditional pay-for-delay agreement entrant i correctly believes that entrant j got the same deal, as this maximizes the incumbent's joint payoff with j, assuming that i accepts the offer. The equilibrium outcome is thus consistent with wary beliefs.

It remains to consider the case of simple contracts. Given $x_i \in \{0, 1, W\}$ the contract offer to x_j maximizing the incumbent's joint profit with j is given by:

$$BR_j(x_i) = \begin{cases} 0 & \text{if } x_i \in \{0, W\}, \\ 0 & \text{if } x_i = 1 \text{ and } \beta \leq 0, \\ 1 & \text{if } x_i = 1 \text{ and } \beta \geq 0. \end{cases}$$

Thus, licensing to both entrants is consistent with wary beliefs only if $\beta \geq 0$, whereas excluding both entrants is always supportable as a continuation equilibrium of the game. In particular, licensing to one and excluding the other one is never an equilibrium; conditioned on one entrant being delayed, the incumbent always has an incentive to secretly delay the other one. Thus, in equilibrium without litigation the incumbent offers a license to both entrants if $\beta \geq 0$ and $f(3; \theta) > f(1; \theta)$, the latter inequality being equivalent to $\theta < (\alpha + \beta)/2$. Otherwise, it excludes both entrants.

As litigation is observable, in equilibrium with litigation only the beliefs of the challenger are relevant, and

$$BR_j(L) = \begin{cases} 0 & \text{if } \beta \leq 0, \\ W & \text{if } \beta \geq 0, \end{cases}$$

corresponding to the outcome with public offers. The gain from litigation is:

$$\tilde{G}(\theta) = \theta\Pi(1) + (1-\theta) f(k^*;0) - \begin{cases} f(1;\theta) & \text{if } \beta < 0, \\ \max\{f(1;\theta), f(3;\theta)\} & \text{if } \beta \geq 0. \end{cases}$$

By subtracting and adding $f(k(\theta);\theta)$, this can be written in terms of $G(\theta)$ in the following way:

$$\tilde{G}(\theta) = G(\theta) + \begin{cases} f(k(\theta);\theta) - f(1;\theta) & \text{if } \beta < 0, \\ f(k(\theta);\theta) - \max\{f(1;\theta), f(3;\theta)\} & \text{if } \beta \geq 0. \end{cases}$$

First, if $k(0) = 1$, then $k(\theta) = 1$ and therefore $\tilde{G}(\theta) = G(\theta)$ for all θ, so Proposition 3 applies. Furthermore, if $k(0) = 3$ and $\alpha \leq 0$, then $f(k(\theta);\theta) = \max\{f(1;\theta), f(3;\theta)\}$ and again $\tilde{G}(\theta) = G(\theta)$ for all θ, so that Proposition 4 applies. We are then left with two cases: either $k(0) = 2$ or $k(0) = 3$ together with $\alpha > 0$ hold. In the latter case, $k(0) = 3$ implies $\beta \geq 0$, which together with $\alpha > 0$ implies

$$f(k(\theta);\theta) - \max\{f(1;\theta), f(3;\theta)\} = [\pi(2) - \pi(3)] \max\{0, \min\{\theta - \beta, \alpha - \theta\}\}.$$

Finally, suppose that $k(0) = 2$. Then $\beta < 0$ and we have

$$f(k(\theta);\theta) - f(1;\theta) = [\pi(2) - \pi(3)] \max\{0, \alpha - \theta\}.$$

We obtain the expression for $\tilde{G}(\theta)$ by observing that in the first two cases $\alpha \leq 0$ holds.

Proof of Proposition 10

After litigation, the incumbent has an incentive to renegotiate if and only if one of the entrants is active and $\beta > 0$ holds. In this case, it accommodates the originally excluded entrant, instead of sticking to the status quo, where that entrant stays out. But then, the incumbent can initially monopolize the market by offering both entrants pay-for-delay agreements with $-t_i = (1-\theta)\pi(3) - c$. Each accepts the offer, anticipating that, if it rejects the deal and litigates, the incumbent renegotiates with the other entrant to accommodate it.

Online Appendix

Secret contracting with passive beliefs

In this section, we analyze secret contracting with passive beliefs (and publicly observable litigation) for the case of two entrants. In this setting, when receiving an unexpected offer from the incumbent, each entrant "passively" believes that the other entrant has nevertheless received the equilibrium offer. We have the following result:

Proposition 11. *Suppose that litigation is publicly observed, whereas the contracts are otherwise private and the entrants hold passive beliefs. Then, there is litigation in equilibrium if $G(\theta) > \lambda$. If instead $G(\theta) < \lambda$, then both entrants are excluded from the market if $k(\theta) = 1$ and $\Pi(1) \geq \Pi(3) + 2\pi(2)$ holds. Otherwise, there exists no equilibrium in pure strategies.*

Proof. As litigation is observable, we can apply the baseline analysis directly to conclude that there is litigation in equilibrium if $G(\theta) > \lambda$. If the reverse inequality is true, in equilibrium both entrants must settle. To see that licensing and pay-for-delay do not coexist in equilibrium, note that the incumbent always has an incentive to secretly deviate by offering the licensee a pay-for-delay agreement with a high reverse payment, to share the monopoly profit instead of the duopoly industry profit. The licensee will accept the modified offer, because it believes that the rival will still be excluded from the market. Similarly, licensing to both entrants cannot constitute an equilibrium, because the incumbent can profitably deviate by secretly offering both of them a pay-for-delay agreement with a low reverse payment. Each will accept, believing that the rival buys a license. The deviation is profitable for the incumbent, because the monopoly profit is greater than the triopoly industry profit. Thus, the only candidate equilibrium without litigation is the one where both entrants sign pay-for-delay agreements. But if $\Pi(1) < \Pi(3) + 2\pi(2)$ holds, the incumbent again has a profitable deviation: to offer each entrant a licensing agreement with a high licensing fee. Each will accept, believing that the other entrant stays out. □

Unrestricted renegotiation

In this section, we enrich the space of contracts that can be offered in the renegotiation stage. Now we assume that the renegotiated agreement includes a simple entry or exit decision and an updated payment from the entrant to the incumbent. As before, we consider a two-stage renegotiation game, where the incumbent first commits to a new set of publicly observable bilateral contract offers to the entrants, and the entrants then simultaneously decide either to accept their respective offers or reject the offer and invoke the original agreement.

At the renegotiation stage, if both entrants are out, the incumbent has no incentive to renegotiate: the monopoly outcome maximizes its payoff, which is the sum of the industry profit and the initial payments. By contrast, if only one of the entrants is out, the incumbent always renegotiates. Perhaps surprisingly, it prefers to accommodate the excluded entrant and purchase the exit of the active one, instead of sticking to the status quo. In the equilibrium, when the rival entrant, that was formerly excluded, enters the market, the formerly active entrant is willing to accept the triopoly profit to stay out, whereas the inactive entrant is ready to pay the entire duopoly profit to enter the market. The incumbent's renegotiation payoff is therefore $\Pi(2) + \pi(2) - \pi(3)$ as opposed to $\Pi(2)$ it would obtain without renegotiation (net of the initial payments).[36]

When only one entrant is out, the incumbent could also renegotiate to accommodate or exclude both entrants. However, the excluded entrant is only willing to pay the triopoly profit if the other entrant stays active. Thus, the incumbent's renegotiation payoff from accommodating both is only $\Pi(3) + \pi(3)$, which is suboptimal, because the industry profit is decreasing in the number of firms in the market. Excluding both entrants is instead profitable if the monopoly profit is sufficiently high:

$$\Pi(1) - \pi(2) > \Pi(2) + \pi(2) - \pi(3). \tag{3}$$

In this case, the incumbent pays the duopoly profit to exclude also the active entrant

Finally, if both entrants are in the market, the incumbent has the same incentives as in the initial stage of contracting with a null patent. Applying our baseline analysis, the incumbent renegotiates to optimally exclude $3 - k(0)$ entrants, where $k(0)$ depends on the mode of competition:

$$k(0) = \arg\max_{k=1,2,3} f(k;0).$$

In the equilibrium, the firms anticipate renegotiation at the initial stage of contracting. If there is no litigation in equilibrium, the prospect of renegotiation only influences the rents the incumbent must leave to the entrants (the reservation utility of an entrant is given by its litigation payoff accounting

[36]With any number of entrants, we thus know that, if the number of originally excluded entrants is greater than the number of accommodated entrants after renegotiation, it must be the case that every accommodated entrant was originally excluded. And in the opposite case, every originally excluded entrant is accommodated.

for renegotiation after the court ruling). But if there is litigation, renegotiation also has a direct impact on the profits made by the incumbent.

We consider three cases $k(0) = 1$, 2, or 3 and discuss them one by one. When $k(0) = 1$, the monopoly profit is so high that the incumbent always renegotiates to exclude both entrants unless they are already out. To save on litigation costs, the incumbent offers each entrant a pay-for-delay agreement with a high reverse payment $-t_i = (1-\theta)\,\pi(2) - c$ and the entrants accept these offers. This outcome corresponds to Proposition 3 in the baseline model. It is thus renegotiation-proof.

Both entrants are also excluded if $k(0) = 2$, but with the difference that each entrant is willing to accept the low reverse payment $-t_i = (1-\theta)\,\pi(3) - c$ to stay out. Indeed, each entrant anticipates that, if it rejects the deal and successfully litigates, the incumbent renegotiates to accommodate the rival and offers only the triopoly profit to purchase the challenger's exit. Due to renegotiation, the incumbent is able to monopolize the market at the lowest possible cost, and cannot do better than this. Exactly the same applies if $k(0) = 3$ and (3) does not hold.[37] The outcome maps to Proposition 2, so conditional deals are in fact renegotiation-proof for these modes of competition.

Interestingly, when $k(0) = 3$ and (3) holds, both entrants are excluded from the market only if the patent is sufficiently strong. Moreover, the key insight from Proposition 4 that patents of intermediate strength are litigated, carries over:

Proposition 12. *Suppose that renegotiation is feasible. Then, both entrants are excluded from the market, unless*

$$\Pi(3) > \Pi(1) - 2\pi(2) > \Pi(2) - \pi(3)\,. \tag{4}$$

In this case, there is litigation for patents of intermediate strength, i.e. whenever

$$\tilde{G}(\theta) \triangleq \theta\Pi(1) + (1-\theta)\max\{\Pi(3), \Pi(1) - \pi(2) - \pi(3)\} - f(k(\theta);\theta) > \lambda.$$

The non-challenger waits if $\Pi(3) > \Pi(1) - \pi(2) - \pi(3)$ and is excluded otherwise. If instead $\tilde{G}(\theta) < \lambda$, then there is no litigation: $3 - k(\theta)$ entrants are excluded and $k(\theta) - 1$ buy a license.

Proof. It remains to consider the case when (4) holds. Then, there is renegotiation only if one entrant is active and the other one is out, in which case the incumbent pays the active entrant the duopoly profit to exclude it from the market and does not renegotiate with the originally excluded entrant. Thus, in equilibrium without litigation, by rejecting the incumbent's contract offer an entrant expects

[37] Notice that our numerical example satisfies these conditions.

the payoff $(1-\theta)\,\pi(2)-c$ if the rival entrant signs a pay-for-delay agreement and $(1-\theta)\,\pi(3)-c$ if the rival buys a license. The prospect of renegotiation thus does not affect the incumbent's payoff from no litigation, which is given by $f(k(\theta);\theta)+2c$. By (4) we can write:

$$f(k(\theta);\theta)=\begin{cases}\Pi(3)+\theta 2\pi(3) & \text{if } \theta\leq\tilde{\theta},\\ \Pi(1)-(1-\theta)\,2\pi(2) & \text{if } \theta\geq\tilde{\theta},\end{cases}$$

where

$$\tilde{\theta}=\frac{\Pi(3)-\Pi(1)+2\pi(2)}{2\pi(2)-2\pi(3)}.$$

However, if one of the entrants litigates and the non-challenger signs a pay-for-delay agreement, the incumbent will renegotiate to exclude the challenger if the court declares the patent invalid. Anticipating this, the incumbent's payoff is

$$\Pi(1)-(1-\theta)\,[\pi(2)+\pi(3)]-C,$$

where $(1-\theta)\,\pi(3)$ is the reverse payment made to the non-challenger (by waiting, it would obtain the triopoly profit if the patent is declared invalid, and with both entrants being active, there would be no renegotiation) and $(1-\theta)\,\pi(2)$ is the *expected* reverse payment made to the challenger. If the non-challenger waits instead, the incumbent obtains

$$\theta\Pi(1)+(1-\theta)\,\Pi(3)-C,$$

which is the same payoff as in the baseline model, because there is no renegotiation if both entrants are active or both stay out. Notice that this payoff is strictly higher than what the incumbent would obtain by licensing the patent to the non-challenger – the only difference is that the incumbent will renegotiate and exclude the licensee by paying it the duopoly profit if the court upholds the patent. The incumbent's gain from litigation is therefore

$$\tilde{G}(\theta)=\theta\Pi(1)+(1-\theta)\max\{\Pi(3),\Pi(1)-\pi(2)-\pi(3)\}-f(k(\theta);\theta),$$

where $\Pi(3)>\Pi(1)-\pi(2)-\pi(3)$ implies that the non-challenger waits and if the reverse inequality holds, it signs a pay-for-delay agreement. Finally, one can check that $\tilde{G}(0)\geq 0=\tilde{G}(1)$, and that the slope of $\tilde{G}$ is piecewise linear, strictly positive for $\theta < \tilde{\theta}$ and strictly negative for $\theta > \tilde{\theta}$. Hence, $\tilde{G}(\theta)$ is a concave function and for any $\lambda < \max \tilde{G}$, there exists a nonempty interval of patent strength, where litigation takes place in equilibrium. □

The general insight from the analysis is that, when the renegotiation is feasible, entry exclusion is likely to be a profitable strategy for the incumbent. On the one hand, the prospect of renegotiation tends to reduce the cost of excluding entry at the initial stage of contracting. The greater is the incumbent's incentive to accommodate entry after patent invalidation, the cheaper it is to buy exit in the first place. Anticipating new entry through renegotiation, the entrants have low litigation payoffs and are willing to sign pay-for-delay agreements with small reverse payments.

On the other hand, the greater is the incumbent's incentive to keep excluding entry after patent invalidation, by definition the higher is the price it is willing to pay in order to monopolize the market. Thus, even if the entrants anticipate high payoffs from litigation, the incumbent may well find it profitable to offer high reverse payments to purchase their exit.