



Universitat
Pompeu Fabra
Barcelona



REVERSE PAYMENTS AND PATENT STRENGTH UNDER ASYMMETRIC INFORMATION*

Anton-Giulio Manganelli

March 2016

CRES-UPF Working Paper #201603-89

Reverse Payments and Patent Strength under Asymmetric Information*

Anton-Giulio Manganelli[†]

CRES - UPF

March 28, 2016

Abstract

In the pharmaceutical industry, a reverse payment is a payment from an originator to a generic producer in exchange of a delay in her entry. In some recent cases, the US and the EU Antitrust Authorities have banned these agreements *per se*. This paper analyzes their dynamic effects and shows that this should not be the case when the parties' investment decisions are taken into account and the information over the patent strength is asymmetric. Reverse payments make the monopoly period longer and increase industry profits, which increases the entrant's investment. This makes generic entry possible before patent expiry and, because of the asymmetric information over the patent strength, increases the litigation rate. A fraction of it ends up in actual generic entry. Both effects increase consumer surplus. Reverse payments delay entry to increase entry. We derive the optimal policy and provide simulations. Reverse payments also create a tension in the originator's incentives to invest, absent from the traditional patent literature. Results suggest that a rule of reason is more suited than a ban *per se*.

JEL Classification: K21, L12, L41.

Keywords: reverse payments, pay-for-delay, cartel, pharmaceutical industry, litigation, settlement, asymmetric information.

*I thank Klenio Barbosa, Roberta Dessì, Bruno Jullien, Gerard Llobet, Xavier Martinez-Giralt, Patrick Rey, the CRESSE (2013) and the EPIP (2013) participants for their useful comments and discussions. I am deeply indebted, in particular, to Yassine Lefouili for his invaluable help. The views expressed in this paper are those of the author and do not necessarily reflect the position of CRES.

[†]E-mail: ag.manganelli@gmail.com.

1 Introduction

In the pharmaceutical industry, patents are the main assets of firms. Originators have the right to enforce them by litigating against potential infringers such as generic entrants that may infringe their patents. Litigations, however, are costly and their outcome is uncertain. As an alternative to litigation, the parties can settle. In some cases, such settlements may involve a value transfer (e.g. a payment) made by the originator to the generic. Such value transfers are denoted “reverse payments”. They are a topical and important issue in antitrust analysis as they could cover cartel-like agreements: a firm with a weak patent may agree to share its monopoly profits with a rival through a payment in exchange of a delay in her entry.

In *FTC v Actavis*, 17 June 2013, the Supreme Court in the US found that “reverse payment” agreements are not immune from antitrust scrutiny, ending a period of inconsistent treatments by different US Circuit Courts.¹ The Supreme Court ruled that reverse payment settlements within the scope of a patent could be assessed by competition law; it also found that such settlements are not *per se* legal, nor are they *per se* illegal, but must instead be assessed on a “rule of reason” basis.² In contrast, the Canadian Competition Bureau has stated that it will adopt a harsher line against reverse payment settlements than in the US.³

In Europe, reverse payments are under the spotlight as well. In 2008, the European Commission launched an inquiry into the pharmaceutical sector with dawn raids at the premises of several originator and generic companies, with particular attention to settlements involving reverse payments. In *Lundbeck*, the (at the time) Commission Vice-President Joaquín Almunia, in charge of competition policy, said: “*It is unacceptable that a company pays off its competitors to stay out of its market and delay the entry of cheaper medicines. Agreements of this type directly harm patients and national health systems, which are already under tight budgetary constraints. The Commission will not tolerate such anticompetitive practices*”.⁴ This suggests a presumption that reverse payments are harmful. More recently, while the Commission has indicated that settlement agreements with reverse payments are likely to at-

¹http://www.supremecourt.gov/opinions/12pdf/12-416_m5n0.pdf.

In 2003, in the Bristol-Myers case, the Cardizem case and the Valley Drug-Geneva Pharmaceuticals case, and in 2006 in the Tamoxifen case, the incumbent paid a potential generic competitor to avoid litigating over the patent and to stay out of the market until patent expiry. The FTC found these agreements anticompetitive.

At the appeal level, however, some of these decisions have been overturned. In the appeal of the Tamoxifen and Valley Drug-Geneva Pharmaceuticals cases, respectively the Sixth and the Eleventh Circuit reversed the initial judgement of the FTC and found the agreements not illegal, as they did not extend beyond the original patent terms.

²The Supreme Court defined a “reverse payment” as a case where a party with no claim for damages walks away with money simply so it will stay away from the patentee’s market. The Supreme Court contrasted that with settlements where a party (e.g. an enjoined party) receives a payment reflecting the potential liability for the originator to pay damages for loss sales during the enjoined period.

³<http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/03817.html>.

⁴European Commission Press Release 19 June 2013, Antitrust: Commission fines Lundbeck and other pharma companies for delaying market entry of generic medicines.

tract “the highest degree of antitrust scrutiny”, it has also indicated that agreements falling into this category would not always be incompatible with EU competition law – an assessment on the basis of the circumstances of each individual case being required.⁵

Proponents of the view that settlements with (large) reverse payments should be presumed anti-competitive typically rely on the argument that, by expanding contractual space, the reverse payment allows the originator scope to compensate the generic would-be entrant with a payment for delaying entry; such a payment is profitable for both parties because it allows the originator to earn monopoly profits during a longer period and because monopoly profits exceed the sum of duopoly profits – so there is scope to compensate the generic entrant. They go on to argue that absent a reverse payment, earlier entry would occur either as a result of settlement or, in the event of litigation, because the expected outcome of litigation would be earlier. These papers reflect the view that patents should be considered as probabilistic property rights (Shapiro 2003, Lemley and Shapiro 2005), as a patent may later be found invalid – or the entrant’s product may be found not to infringe a valid patent.⁶

Others, arguing that reverse payments should be assessed on a rule of reason basis put forward a number of arguments. For example, Willig and Bigelow (2004) put forward reasons why a settlement with reverse payments can be beneficial for consumers. These include differences in: (i) the information about the future states of the market; (ii) the expectation of success in the litigation; and (iii) the impact that entry of another firm has on the incumbent and the entrant. The key point is that a reverse payment can allow settlement to occur that would not otherwise be profitable when the parties have divergent views and, under certain conditions, may lead to earlier entry. Gratz (2012) compares regimes of *per se* legality, illegality and rule of reason. She finds that *per se* legality induces maximal collusion, *per se* illegality entirely prevents it and the rule of reason induces limited collusion when antitrust enforcement is subject to error. This limited collusion can be welfare enhancing, as it increases the expected settlement profits, thus fostering generic entry. This result, however, crucially depends on Antitrust Authorities making errors. Dickey & Rubinfeld (2012) also present an informal discussion of how permitting reverse payments may increase generic entry.

The aim of this paper is to contribute to the literature on reverse payments by analyzing their dynamic effects. We use the analytical framework of the literature of litigations and settlements.⁷ We deliberately adopt a framework consistent with a skeptical view over the

⁵See paragraph 17 of the European Commission’s 5th Report on the Monitoring of Patent Settlements (period: January-December 2013), 5 December 2014, available at http://www.mlex.com/EU/Attachments/2014-12-05_587L6F26816B614Y/patent_settlements_report5_en.pdf The EU Commission defines reverse payment settlements as those which foresee a value transfer from the originator to the generic company.

⁶Leonard and Mortimer (2005) argue for a rule of reason approach but consider that pro-competitive reverse payments are likely to be small in most cases. See also Elhauge & Krueger (2012), who argue that reverse payments in excess of the originator’s anticipated litigation costs are anti-competitive.

⁷Litigations and settlements have been studied by, among others, Salant and Rest (1982), P’Ng (1983), Bebchuk (1984), Salant (1984), Reinganum and Wilde (1986), Schweitzer (1989) and Daughety and Reinganum (1994). Almost all of these models assume that the bargaining process occurs sequentially, where

pro-competitive effects of reverse payments: by this we mean that if both the originator and the generic entrant would invest without reverse payments, then allowing for reverse payments necessarily delays entry (where the entry date in the latter regime is either that which would arise in the event of settlement or the expected entry date in the event of litigation).⁸ We include two ingredients in the setting: first, we consider the *ex ante* incentives to invest for both the originator and the entrant and, second, we assume there is asymmetric information over the patent strength – the likelihood that the originator wins a trial – between the originator and the entrant. In order to avoid the multiplicity of equilibria of signalling games, we assume that the asymmetry of information is one-sided and the non-informed party, the entrant, makes a take-it-or-leave-it offer. Results, as will become clear, hold for more general bargaining rules. We assume that the patent can be "strong" or "weak", and that only the originator knows the true strength of his patent. The entrant only knows its probability distribution. Results hold for continuous distribution of the patent strength.

We find that, while reverse payments delay the generic manufacturer's entry date if both parties would in any event have invested, they also increase the profits available for bargaining and therefore the entrant's profits.⁹ In turn, that increases her incentives to invest in the first place. This may increase consumer surplus in two ways: first, the asymmetric information over the patent strength makes it not always possible for the parties to reach a settlement, making a part of the higher generic investment end up in litigation, which with some probability ends up with the generic producer winning the trial and entering the market; second, when the settlement establishes an entry date for the generic firm prior to patent expiry, the higher generic investment makes the associated higher generic entry occur before patent expiry with certainty. Reverse payments, therefore, create a trade-off between *delaying generic entry* and *increasing generic entry*.

one part makes a take-it-or-leave-it offer and the other one accepts or rejects it. If the responder accepts, the terms of the offer are enforced, while, if he rejects, parties litigate. Except for Schweitzer (1989) and Daughety and Reinganum (1994), incomplete information is one-sided. Some models assume that the party making the offer is the informed one (P'Ng, Salant and Rest, Salant), in which case, due to the transmission of private information through the offer, equilibria are typically very numerous (a well known feature of signalling games), while Bebchuk assumes the opposite - which makes the equilibrium unique. Other models of bargaining assume that the identity of the proposer is determined by a coin flip, like Rubinstein and Wolinsky (1985, 1990), Gale (1986a, 1986b, 1987) and Binmore and Herrero (1988a, 1988b), or that both parties make simultaneous announcements (Wolinsky 1990).

⁸We assume that there is only one potential entrant. Appendix 5.6 discusses the impact of allowing for more entrants.

⁹We adopt a one-sided incomplete information game, where the less informed party (the entrant) makes a take-it-or-leave-it settlement offer. A similar approach was taken by Bebchuk (1984). This is the simplest way to ensure that the entrant obtains a non-negative share of the bargaining pie (i.e. the incremental profit available from widening contractual space to allow for reverse payment settlements relative to the fallback option of litigation). The originator may also secure a share of the pie due to its information advantage. Our results hold for more general bargaining rules (e.g. splitting the pie according to some pre-determined share) provided that the originator does not secure the entirety of the pie.

The originator's perspective is more complex. On the one hand, reverse payments allow him to benefit from greater profits in the event of generic entry (promoting his investment incentives); on the other hand, he may suffer because generic entry may occur when otherwise it would not have done (weakening the incentive to invest). Where the size of the reverse payment is regulated by an antitrust authority, we find that permitting a larger (positive) reverse payment increases the bargaining pie and makes the entrant more likely to prefer to settle than to engage in litigation. The entrant's less aggressive stance towards litigation permits the originator to earn a higher *ex ante* expected information rent, which may induce the originator to invest when otherwise he would not have done, thereby allowing a market to be created and increasing consumer surplus.

The impact of reverse payments on consumer surplus is, therefore, not trivial and deserves a careful analysis. There exist parameter sets where the positive effect of inducing entry offsets the negative effect of delaying it. This suggests that a rule of reason is more suited than a ban *per se*.

The optimal policy dictates that when investments by the originator and the entrant would in any case occur, or *preventing* the entrant from investing is essential to ensuring that the originator invests, a ban on reverse payments is optimal. However, in the other cases, it is optimal to permit reverse payments and the cap must be set at the minimum consistent with inducing the entrant to invest (when otherwise she would not and where such an investment does not deter the originator from investing) and/or (ii) inducing the entrant to make a less aggressive settlement demand to the originator (i.e. the entrant accepts a smaller share of the available bargaining pie, allowing more for the originator and thereby ensuring his investment is not deterred). In these cases reverse payments increase consumer surplus (and of course total welfare).

Suppose, for sake of argument, that a "strong" patent is one that means there is a 67% chance of success if litigation occurs. In a static game where entry by the generic would in any event occur, a settlement with a reverse payment that gives rise to delay in excess of two thirds of the period that remains until patent expiry might be viewed as anti-competitive (since the agreed delay exceeds the expected delay that would arise with litigation). However, in a dynamic game, where generic investment occurs only if the entrant covers its up front entry cost, that greater delay due to reverse payments may nonetheless be pro-competitive through its impact on generic investment. The optimal permissible reverse payment must induce entry by the generic entrant with the minimum expected delay. Moreover, the asymmetric information over the patent strength may force parties to litigate, which further increases consumer surplus. Consumer surplus is maximized by choosing a cap on reverse payments inducing the earliest expected entry date consistent with litigation being possible – the entrant targets the weak realization of the patent strength – and just covering her entry cost. The reverse payment corresponding to that entry date may be zero (in which case a ban on reverse

payments is optimal) or positive (in which case a ban on reverse payments harms consumer surplus).

In order to show the intuitions set out above, we present simulations of hypothetical scenarios (Section 3). Specifically, these simulations allow a number of key parameters to vary such as the entrant's marginal cost, the degree of uncertainty over the patent strength, and the intensity of competition between the originator and the entrant (modelled by a conjectural variation parameter¹⁰).

We find that, where the generic entrant's investment cost is relatively low, consumer surplus falls as the permitted value of the reverse payment increases – such that a ban on reverse payments would maximize consumer welfare. However, as the entry cost increases, there comes a point when entry by the generic would not occur at all. At this stage, increasing the cap on reverse payments increases consumer surplus; however, this relation lasts only up to a point, after which further increases in the cap reduce consumer surplus. When the entrant has a cost disadvantage, it is more important to incentivize the generic to invest, so it is better to allow reverse payments. The same occurs when the competition is fierce on the market. Finally, the more uncertain the patent strength is, again, the better it is to allow reverse payment. The reason is that the more uncertain the patent strength is, the more costly it is for the entrant to make sure that a settlement will be reached. This incentivizes her to make an aggressive proposal, which increases consumer surplus through the possibility of litigation.

In general, industry features that reduce the entrant's ability to recover its entry cost tend to work in favour of a more permissive regime as regards reverse payments. For example, when the originator's product is perceived to be of higher quality, then a ban on reverse payments is less likely to be beneficial. Likewise, where the generic has a marginal cost disadvantage vis-à-vis the incumbent, it is less likely that the generic can recover the investment without a reverse payment. Indeed, in some cases a reverse payment can give rise to generic investment and entry (and higher consumer surplus) where the entry cost exceeds the entrant's profits – clearly entry would never occur in that case without a reverse payment.

The paper is organized as follows. Section 2 presents the model, Section 3 shows the numerical examples, Section 4 concludes, and Section 5 presents the Appendix.

¹⁰The conjectural variation is the belief a firm has over the reaction of the other firms in response to a change in its output or price. It is a way to model the intensity of competition in a market. It takes a value between -1 and 1. For example, in a Cournot game where firms choose quantities, Nash equilibrium means a conjectural variation of 0. A conjecture of -1 makes this game equal to a Bertrand game: each firm thinks that raising its own quantity makes the other firms reduce their quantity in such a way that total quantity remains the same. A conjecture of +1 is equal to the monopoly problem, as each firm believes that its choices will be imitated by the others.

2 The model

There are three players: an Antitrust Authority (AA), an originator and a generic manufacturer (the entrant). Normalize patent length to 1 and current date to 0.¹¹ In the first stage, the AA decides the maximal allowed amount of reverse payment \hat{R} . A maximal reverse payment of 0 means that they are banned, and one of ∞ means that no cap is set. In the second stage, the originator can invest a sum I_O to enter the market. In the third stage, if the originator has invested, the generic manufacturer can enter the market if he invests a sum I_E . In the fourth stage, if both the originator and the generic manufacturer have invested, the entrant makes a take-it-or-leave-it settlement offer.¹² The offer consists of an entry date $0 \leq D \leq 1$ and a payment $R \leq \hat{R}$ from the originator to the entrant. In the fifth stage, the originator learns the patent strength – the true probability $\theta \in \{\underline{\theta}, \bar{\theta}\}$ of winning the litigation, where $0 \leq \underline{\theta} < \bar{\theta} < 1$.¹³ The probability of drawing $\underline{\theta}$ is λ . This signal represents the information that arrives from the national patent office or from experts asked to evaluate the patent strength prior to the potential litigation. We assume that the originator has better information about the patent strength because he is the party that filed the patent application and, therefore, has better knowledge of its possible problems.¹⁴ In the sixth stage, if the originator accepts the offer D and R are enforced, otherwise the parties litigate. D represents the fraction of the patent period in which the entrant commits not to enter.

The timing is then the following:

1. **Policy choice.** The Antitrust Authority chooses a cap \hat{R} .
2. **Originator's investment.** The originator invests I_O to enter the market or stays out.
3. **Entrant's investment.** The entrant invests I_E to enter the market or stays out.
4. **Entrant's offer.** If the entrant has invested, the entrant makes a settlement offer.

¹¹Date 0 is the date when the entrant is ready to enter, which is the same as the one when the parties, having invested, decide whether to litigate or to settle. This will be shown in Appendix 5.2.

¹²The fact that the entrant makes a take-it-or-leave-it offer (or, better, a take-it-or-leave-it request) is not necessary for the results. Any form of bargaining that leaves the entrant with *some* additional surplus from the settlement compared to his threat point (the litigation payoff) yields our qualitative results. In other words, the only bargaining solution that is not compatible with the results is the originator making the take-it-or-leave-it offer.

¹³The patent strength represents both the probability that the patent is held valid and infringed - see the next footnote.

¹⁴The set-up with the originator receiving the private signal suits to the case where the entrant challenges the validity of the patent - so it is reasonable to assume that the originator has better information. However, results are qualitatively the same when the entrant has better information - which suits to the case where the entrant seeks to invent around the patent, provided that the identity of the party making the take-it-or-leave-it offer is inverted.

5. **Originator's signal.** The originator receives the private signal $\theta \in \{\underline{\theta}, \bar{\theta}\}$.¹⁵
6. **Originator's response.** The originator accepts or rejects it. Rejection implies litigation.

In case of litigation, the originator and the entrant bear, respectively, litigation costs C_O and C_E .¹⁶ Define H the originator's profits if he is the monopolist for the entire patent period, L the originator's profits if entry occurs immediately and E the entrant's profits if it enters immediately. Hence, $L + E$ are the joint profits of the originator and the entrant if entry occurs immediately. Assume that $H > L + E$: monopoly profits are larger than the industry duopoly profits.

When reverse payments are allowed, it can be easily shown that industry profits are higher (intuitively, because they are used only if they delay entry, so that the pledgeable profits increase - see Lemma 1). Therefore (i) the entrant always has more incentives to invest, (ii) the originator may have more or less incentives to invest (more if the entrant would have entered anyway and the higher profits make him offer more favorable settlement terms; less if they induce a generic manufacturer that would have otherwise stayed out to enter) and (iii) CS is lower for a *given* investment level. We will show that there exist several parameter sets where CS increases under $\hat{R} > 0$, even when no cap is set ($\hat{R} = \infty$), thanks to the pro-investment effects.

The following subsection analyzes the last stage of the game and computes the litigation and settlement profits.

2.1 Litigation-Settlement stage

The model is solved backwards. Denote $\check{\theta}$ the (hypothetical) realization for which the originator would be indifferent between litigating and accepting the entrant's offer, and $E_{\theta}[\theta | \theta > \check{\theta}]$ the entrant's Bayesian updating of θ given that the originator refuses a proposal based on $\check{\theta}$.

If the parties *litigate*, they expect to obtain:

Originator: $\theta H + (1 - \theta)L - C_O$

Entrant: $(1 - E_{\theta}[\theta])E - C_E$ before the settlement offer;

$(1 - E_{\theta}[\theta | \theta > \check{\theta}])E - C_E$ after the settlement offer if refused.

By litigating, the originator knows he has a probability θ of winning the case, in which case he gets H ; with probability $(1 - \theta)$ he loses and gets only L . Whether he wins or

¹⁵The originator is assumed to learn the true probability of winning the trial. Results are robust to variations to this assumption (e.g. the originator only receiving a noisy signal over the true probability or the entrant getting a signal over the patent strength). The only necessary feature is some asymmetric information between the originator and the entrant over the patent strength.

¹⁶They can be seen as the incremental legal costs of litigation – those in excess of any legal costs associated with settlement.

loses, litigation costs are C_O . The entrant, instead, knows *ex ante* that she has a probability $(1 - E_\theta[\theta])$ of winning, in which case she gets E , otherwise she earns nothing. If the originator refuses the settlement offer, this probability becomes $(1 - E_\theta[\theta|\theta > \check{\theta}])$. Her litigation costs are C_E .¹⁷

If the parties *settle*, they obtain:

Originator: $DH + (1 - D)L - R$

Entrant: $(1 - D)E + R$.

By settling, the originator earns DH in the period before the agreed entry date and $(1 - D)L$ in the period until patent expiry - in which he competes with the entrant. He also pays R to him. The entrant earns $(1 - D)E$ if she enters at date D and receives the payment R .

Solving the model backwards, in stage 6 the originator settles if and only if this is at least as profitable as litigating: $DH + (1 - D)L - R \geq \theta H + (1 - \theta)L - C_O$, which yields:

$$D \geq D^* = \theta + \frac{R - C_O}{H - L}. \quad (1)$$

The minimal entry date that the originator is willing to accept is increasing in R and θ and decreasing in C_O and $(H - L)$ (as long as R is larger than C_O). A higher patent strength makes the originator more confident of winning the litigation and, therefore, less willing to accept an early entry. A higher R means that accepting the settlement is more costly, which also makes the originator less keen on settling. However, a settlement makes the originator save the litigation costs C_O . The net cost of the settlement $R - C_O$ is weighted over the gain of making monopoly profits longer, $(H - L)$.

When reverse payments are allowed, the following Lemma shows that the potential bidimensionality of the settlement (over R and D) reduce to monodimensional.

Lemma 1 *If the entrant has invested, she asks for the maximal possible payment \hat{R} and enters at D^* if D^* is not greater than 1, and asks for $R = (1 - \theta)(H - L) + C_O$ and enters at the patent expiry otherwise.*

Proof. See Appendix 5.1 for the first part of the Lemma. The intuition is that a larger reverse payment more than compensates the profit loss due to the later entry needed to keep the originator willing to settle. A higher reverse payment implies a later entry date – see (1): a marginally higher R makes the entrant gain dR through R and lose $\frac{E}{H-L}dR$ through D . Being the loss in the originator's profits $(H - L)$ higher than the entrant's profits E , the optimal R is the maximal possible one: $R = \hat{R}$. The second part comes from the fact that that the parties cannot agree on an entry date after patent expiry $D > 1$. Therefore,

¹⁷The fact that each party bears his own litigation costs is the so called American rule. Results are robust to changes in the allocation of litigation costs.

$D = 1$ is an implicit constraint on the maximal reverse payment: the entrant cannot ask for a reverse payment so high that the originator would need a monopoly period even longer than the patent duration to accept the settlement. Any policy $\hat{R} > (1 - \theta)(H - L) + C_O$, therefore, is equivalent to $\hat{R} = (1 - \theta)(H - L) + C_O$. For the sake of exposition, we assume in the following that $\hat{R} \leq (1 - \theta)(H - L) + C_O$.

Note, moreover, the duality between the imposition of a latest entry date \hat{D} and the maximal reverse payment \hat{R} .

Lemma 2 *The choice of the maximal reverse payment \hat{R} is a perfect substitute for the choice of the latest entry date \hat{D} .*

Proof. See (1). There is a biunivocal correspondence between R and D^* , so setting a cap on R or D is equivalent.

We can therefore restrict our attention to a cap on R . This Lemma also implies that the model can be reinterpreted in terms of latest entry date instead of maximal reverse payment.

2.2 Optimal offer of the entrant

In principle, the entrant could behave as the principal in a principal-agent relationship with the originator, i.e. he could write a menu of contracts and make the originator truthfully reveal his type. However, this is impossible, as the originator's settlement profits do not depend on his type (see Appendix 5.5). Therefore, the contract for the high type $\bar{\theta}$ must yield the same profits for the originator as the one for the low type $\underline{\theta}$, in order to make the originator not lie about his own type. The entrant can decide to make the participation constraint of the high type binding, in which case both types of originator accept the settlement, or the one of the low type, in which case the high type will litigate. In other words, given that it is not possible to elicit the originator's type and that θ has 2 possible realizations, the entrant (provided that he has invested) has two potential optimal strategies: one that leaves the weak originator (i.e. when he draws $\theta = \underline{\theta}$) indifferent between litigating and settling and one that leaves the strong one (i.e. when he draws $\theta = \bar{\theta}$) indifferent. Call the realization $\hat{\theta} \in \{\underline{\theta}, \bar{\theta}\}$ leaving the originator indifferent between litigating and settling the realization "targeted" by the entrant. Denote $\hat{\theta}$ the equilibrium targeted realization under the policy \hat{R} . From Lemma 1, we know that we can substitute R with \hat{R} . Upon entry, the entrant will optimally propose a reverse payment \hat{R} and an entry date $D(\hat{\theta})$ such that the originator is indifferent between accepting the settlement and litigating when the realized θ is $\hat{\theta}$.

Denote $\pi_i^{\hat{R}}(\hat{\theta})$ the expected profits of party i , where $i = E$ is the entrant and $i = O$ is the originator, under policy \hat{R} when the entrant targets the realization $\hat{\theta}$. In other words, $\pi_i^{\hat{R}}(\hat{\theta})$ represents the expected profits after stage 4 (the entrant makes the offer) and before stage 5 (the originator does not know his private signal yet). Under policy \hat{R} , in equilibrium the

originator will accept the request $\{R = \hat{R}, D = D(\hat{\theta}) = \hat{\theta} + \frac{\hat{R}-C_O}{H-L}\}$ if and only if the realized patent strength is not larger than the targeted one: $\theta \leq \hat{\theta}$ – see (1).

When the entrant targets $\underline{\theta}$, we have

$$\pi_E^{\hat{R}}(\underline{\theta}) = \lambda[(1 - \underline{\theta} - \frac{\hat{R} - C_O}{H - L})E + \hat{R}] + (1 - \lambda)[(1 - \bar{\theta})E - C_E]. \quad (2)$$

When the entrant targets $\underline{\theta}$, the parties settle if and only if the realized θ is $\underline{\theta}$. This occurs with probability λ . With probability $(1 - \lambda)$, the realized θ is $\bar{\theta}$, which makes the originator litigate. Therefore, in case of litigation also the entrant knows that the patent strength is $\theta = \bar{\theta}$, so both parties compute their litigation payoffs accordingly.

When the entrant targets $\bar{\theta}$, we have

$$\pi_E^{\hat{R}}(\bar{\theta}) = (1 - \bar{\theta} - \frac{\hat{R} - C_O}{H - L})E + \hat{R}. \quad (3)$$

In this case, litigation never occurs in equilibrium. When the entrant targets the high realization, he is paying for an insurance: he is leaving some information rent to the originator in exchange for the certainty of avoiding litigation.

Given (2) and (3), the following Lemma describes the optimal offer of the entrant.

Lemma 3 *The entrant targets $\hat{\theta} = \underline{\theta}$ under policy \hat{R} if and only if*

$$\lambda > \lambda^{\hat{R}} = \frac{\hat{R} + C_E - \frac{\hat{R}-C_O}{H-L}E}{\hat{R} + C_E - \frac{\hat{R}-C_O}{H-L}E + (\bar{\theta} - \underline{\theta})E}.$$

Corollary 1 $\frac{d\lambda^{\hat{R}}}{d\hat{R}} = \frac{(1 - \frac{E}{H-L})(\bar{\theta} - \underline{\theta})E}{[\hat{R} + C_E - \frac{\hat{R}-C_O}{H-L}E + (\bar{\theta} - \underline{\theta})E]^2} > 0.$

This important corollary implies that the AA can affect the realization targeted by the entrant, as $\lambda^{\hat{R}}$ is strictly increasing in \hat{R} . This means that an entrant that targets $\underline{\theta}$ switches target to $\bar{\theta}$ when the allowed reverse payment is sufficiently larger. In other words, the entrant offers a better offer to the originator, in order to avoid litigation. This result will be useful when the entrant must be incentivized to offer a more favorable settlement to the originator in order to make him invest. Lemma 3 yields also the following corollary.

Corollary 2 $\frac{d\lambda^{\hat{R}}}{d(\bar{\theta} - \underline{\theta})} = -\frac{(\hat{R} + C_E - \frac{\hat{R}-C_O}{H-L}E)E}{[\hat{R} + C_E - \frac{\hat{R}-C_O}{H-L}E + (\bar{\theta} - \underline{\theta})E]^2} < 0.$

The term $(\bar{\theta} - \underline{\theta})$ can be interpreted as the uncertainty over the patent strength. For the same expected patent strength, the more uncertain the patent strength is, the more it becomes costly to make sure that litigation will not occur. Litigation indeed depends on the high realization – now higher. Therefore it becomes more appealing for the entrant to be aggressive and try to get a profitable settlement by targetting the low realization.

Consider now consumer surplus (CS). If a settlement takes place, CS is simply the monopoly CS until the generic manufacturer's entry date D and the duopoly CS from that moment until patent expiry $(1-D)$. Denoting \underline{S} the monopoly CS and $\bar{S} > \underline{S}$ the duopoly CS, we have $CS_S = D\underline{S} + (1-D)\bar{S}$. If litigation occurs, we follow Shapiro (2003) by assuming that CS is equal to the probability that the originator wins the case times the monopoly CS, plus the probability that the entrant wins times the duopoly CS. Therefore $CS_L = \theta\underline{S} + (1-\theta)\bar{S}$.

Disregarding the investment decisions, it is easy to see that CS is higher when reverse payments are banned ($\hat{R} = 0$). A "laissez faire" policy ($\hat{R} = \infty$) makes firms choose $D = 1$, while banning them makes the entrant propose $D(\hat{\theta}) = \hat{\theta} - \frac{C_O}{(H-L)}$, which is smaller than 1. Being $\bar{S} > \underline{S}$, it is clear that CS is higher when reverse payments are banned, *provided that both firms invest*. When they are banned, the originator must allow the entrant to enter prior to patent expiry to make her willing to accept the settlement. This supports the opinion shown by the FTC and the EC in some cases that reverse payments should be banned *per se*. However, when considering both parties' incentives to invest, banning reverse payments can reduce CS. A ban on reverse payments, indeed, always reduces the entrant's incentives to invest and creates a tension in the originator's ones. The originator's incentives may be reduced because the smaller industry profits can make the entrant more aggressive in his settlement proposal, while the entrant's incentives are always reduced because of the smaller industry profits. When the originator does not invest the CS falls to zero, while when the entrant does not invest there is no entry before the patent expiry nor litigation, which keeps CS at the monopoly level. For several parameter sets, as will be shown, the entry-enhancing effects of reverse payments dominate the entry-delaying ones. The following subsection describes all the subgame perfect equilibria of the game.

2.3 Subgame perfect equilibria

The possible outcomes of the game depend on (i) the originator's investment, (ii) the entrant's investment and (iii) the entrant's settlement offer.

1) If $I_O > H$, the originator does not invest and the game ends. Consumer surplus is $CS(O_{out}) = 0$. This case is trivial: if monopoly profits are smaller than the investment cost, the originator stays out of the market.

2) If $\pi_O(\bar{\theta}) \leq I_O \leq H$, then the originator invests if and only if the entrant does not enter, i.e. when $I_E > \pi_E^{\hat{R}}(\hat{\theta})$. In this case consumer surplus is $CS(E_{out}) = \underline{S}$, the monopoly outcome. If $I_E \leq \pi_E^{\hat{R}}(\hat{\theta})$, the entrant would invest, which deters the originator from investing in the first place. In this case consumers surplus is $CS(O_{out}) = 0$.

3) If $\pi_O(\underline{\theta}) < I_O \leq \pi_O(\bar{\theta})$, then the originator invests if the entrant does not enter or if he enters and targets $\bar{\theta}$. The entrant enters if and only if $I_E \leq \pi_E^{\hat{R}}(\hat{\theta})$, where $\hat{\theta}$ is the targeted patent strength when the policy is \hat{R} . If $I_E > \pi_E^{\hat{R}}(\hat{\theta})$, the entrant stays out and consumer

surplus is $CS(E_{out}) = \underline{S}$. If the entrant enters, he targets $\underline{\theta}$ if $\lambda > \lambda^{\hat{R}}$ and $\bar{\theta}$ if $0 < \lambda \leq \lambda^{\hat{R}}$.¹⁸ Recall that the targeted realization depends on \hat{R} , as \hat{R} has a positive impact on $\lambda^{\hat{R}}$. For any policy, if the entrant would enter and target $\underline{\theta}$, consumer surplus is $CS(O_{out}) = 0$, as the originator would not invest. On the other hand, if the entrant enters and targets $\bar{\theta}$, then

$$CS^{\hat{R}}(\bar{\theta}) = (\bar{\theta} + \frac{\hat{R} - C_O}{H - L})\underline{S} + (1 - \bar{\theta} - \frac{\hat{R} - C_O}{H - L})\bar{S}.$$

Note that the policy choice has an impact not only on CS, but also on the entrant's profits and therefore on the originator's incentives to invest.

4) If $0 \leq I_O \leq \pi_O(\underline{\theta})$, then the originator invests for any choice of the entrant and, therefore, for any policy. The entrant invests, as usual, if and only if $I_E \leq \pi_E^{\hat{R}}(\hat{\theta})$. If $I_E > \pi_E^{\hat{R}}(\hat{\theta})$, he does not enter and consumer surplus is $CS = \underline{S}$. If he enters, he targets $\underline{\theta}$ if $\lambda > \lambda^{\hat{R}}$ and $\bar{\theta}$ if $0 < \lambda \leq \lambda^{\hat{R}}$.

If the entrant targets $\underline{\theta}$, consumer surplus is

$$CS^{\hat{R}}(\underline{\theta}) = \lambda[(\underline{\theta} + \frac{\hat{R} - C_O}{H - L})\underline{S} + (1 - \underline{\theta} - \frac{\hat{R} - C_O}{H - L})\bar{S}] + (1 - \lambda)[\bar{\theta}\underline{S} + (1 - \bar{\theta})\bar{S}]$$

The originator draws $\underline{\theta}$ with probability λ , in which case he settles, and $\bar{\theta}$ with probability $(1 - \lambda)$, in which case he litigates. In case of litigation, consumer surplus is computed with the true probability that the originator wins, $\bar{\theta}$.

From the analysis of these subgame perfect equilibria, we get to our main result.

Proposition 1 *There exist parameter sets where banning reverse payments reduces consumer surplus.*

Proof and Explanation. Banning reverse payments reduces CS when (i) it impedes generic entry that would otherwise take place, provided that the originator invests, or (ii) it deters the originator's investment, because the lower industry profits make the entrant more aggressive in the settlement offer. We show these two cases in detail. Recall that the originator's profits depend only on the realization θ targeted by the entrant: we have $\pi_O(\underline{\theta}) \equiv \pi_O^0(\underline{\theta}) = \pi_O^\infty(\underline{\theta}) < \pi_O(\bar{\theta}) \equiv \pi_O^0(\bar{\theta}) = \pi_O^\infty(\bar{\theta})$.¹⁹

Case 1: more generic entry. Consider now $\lambda^0 < \lambda^\infty < \lambda$. This means that the probability that the patent is weak is so high that the entrant targets the low realization even when no cap is set on reverse payments. Consider now the investment costs I_E and I_O . If $I_O > \pi_O^{\hat{R}}(\underline{\theta})$, the originator never invests and the game ends. If $I_O \leq \pi_O^{\hat{R}}(\underline{\theta})$, we have three cases:

If (I) $0 \leq I_E \leq \pi_E^0(\underline{\theta})$, then the entrant invests under both policies. This makes CS higher under $\hat{R} = 0$, because this makes entry occur as soon as possible - see (1).

¹⁸i.e. her settlement proposal is $R = \hat{R}$ and $D = \underline{\theta} + \frac{\hat{R} - C_O}{H - L}$ if $\lambda \geq \lambda^{\hat{R}}$ and $D = \bar{\theta} + \frac{\hat{R} - C_O}{H - L}$ otherwise.

¹⁹The equalities $\pi_O^0(\underline{\theta}) = \pi_O^\infty(\underline{\theta})$ and $\pi_O^0(\bar{\theta}) = \pi_O^\infty(\bar{\theta})$ are due to the fact that the entrant makes a take-it-or-leave-it offer. A more general bargaining rule without this feature does not change the qualitative results.

If (II) $\pi_E^0(\underline{\theta}) < I_E \leq \pi_E^\infty(\underline{\theta})$, the entrant invests if and only if the allowed reverse payment is sufficiently high. CS is therefore higher under such a cap, because it makes the entrant enter the market before patent expiry when the entry date associated with \hat{R} is smaller than 1 and because, when the originator draws $\theta = \bar{\theta}$, litigation occurs. Note that in a more general framework, where the possible realizations of θ are more than two, the only necessary condition for litigation being possible is that the entrant does not target the highest realization of θ (or its upper bound, in the case of the continuous distribution).

If (III) $\pi_E^\infty(\underline{\theta}) < I_E$, the entrant does not invest under any policy. CS is then the same under any policy.

The existence of case (II) completes the proof. The intuition is that if the originator's investment cost is sufficiently small and the entrant's cost is intermediate, the originator invests, but the entrant will if and only if reverse payments are allowed and sufficiently high. This increases the size of the pledgeable profits, making the entrant willing to invest. With probability λ this investment will end up in litigation and, when it ends up with a settlement, will make the entrant enter before patent expiry. Both effects increase CS.

Case 2: more originator's investment. Consider now $\lambda^0 < \lambda < \lambda^\infty$. The probability of the low realization of the patent strength is such that the entrant targets $\underline{\theta}$ under $\hat{R} = 0$ and targets $\bar{\theta}$ when \hat{R} is sufficiently high. Therefore, the originator obtains higher profits when reverse payments are allowed and \hat{R} is sufficiently high. When $\pi_O(\underline{\theta}) < I_O < \pi_O(\bar{\theta})$, allowing reverse payments makes the originator invest, which increases CS. The intuition is that reverse payments increases industry profits, which makes the entrant keener on settling on more favorable terms to the originator, in order to reduce the risk of litigation. This increases the originator's profits and triggers the originator's investment - which would not have taken place otherwise. This raises CS, as it creates a market that would not have existed.

These two cases show that a ban *per se* is suboptimal. The next subsection derives the impact of the policy of the AA on CS and derives the optimal policy.

2.4 Ranking of CS and optimal policies

Using the results of the previous subsection, we can rank the outcomes depending on their CS.

1) The best outcome for CS is that both parties invest, the entrant targets $\underline{\theta}$ and the policy is $\hat{R} = 0$. In this case entry occurs as soon as possible, if parties settle, and litigation is possible. In this case, consumer surplus is

$$CS^0(\underline{\theta}) = \lambda[(\underline{\theta} - \frac{C_O}{H-L})\underline{S} + (1 - \underline{\theta} + \frac{C_O}{H-L})\bar{S}] + (1 - \lambda)[\bar{\theta}\underline{S} + (1 - \bar{\theta})\bar{S}].$$

2) The second-best outcome is that, under $\hat{R} = 0$, the entrant enters and targets $\bar{\theta}$. In

this case consumer surplus is

$$CS^0(\bar{\theta}) = (\bar{\theta} - \frac{C_O}{H-L})\underline{S} + (1 - \bar{\theta} + \frac{C_O}{H-L})\bar{S}.^{20}$$

3) The third-best outcome is that, under $\hat{R} = 0$, the entrant would not enter because his profits are smaller than his investment cost, but the AA can equalize them by choosing the appropriate maximal reverse payment. In this case consumer surplus is

$$CS^{\hat{R}}(\underline{\theta}) = \lambda[(\underline{\theta} + \frac{\hat{R} - C_O}{H-L})\underline{S} + (1 - \underline{\theta} - \frac{\hat{R} - C_O}{H-L})\bar{S}] + (1 - \lambda)[\bar{\theta}\underline{S} + (1 - \bar{\theta})\bar{S}].^{21}$$

when the entrant still targets $\underline{\theta}$, and $CS^{\hat{R}}(\bar{\theta}) = (\bar{\theta} + \frac{\hat{R} - C_O}{H-L})\underline{S} + (1 - \bar{\theta} - \frac{\hat{R} - C_O}{H-L})\bar{S}$ if the cap \hat{R} is such that the entrant now targets $\bar{\theta}$.

4) If, under $\hat{R} = 0$, the entrant would enter and target $\underline{\theta}$ but the originator's investment cost I_O is between the profits he would get when the entrant targets the low realization and the ones when the entrant targets the high realization, the originator does not invest. Consumer surplus, therefore, would be $CS^0(O_{out}) = 0$. The fourth-best outcome consists, therefore, of the AA allowing reverse payments with a cap such that the entrant targets $\bar{\theta}$. The entrant targets $\bar{\theta}$ if and only if the probability λ that the patent is weak is smaller than $\lambda^{\hat{R}}$, which occurs when the entry date is sufficiently high. Under such a policy, consumer surplus is $CS^{\hat{R}}(\bar{\theta}) = (\bar{\theta} + \frac{\hat{R} - C_O}{H-L})\underline{S} + (1 - \bar{\theta} - \frac{\hat{R} - C_O}{H-L})\bar{S}$.

5) If the originator's investment cost I_O is higher than the profits $\pi_O(\bar{\theta})$ he makes when the entrant targets $\bar{\theta}$, the originator never invests if the entrant would eventually enter the market. If he would eventually enter, therefore, consumer surplus would be $CS^0(O_{out}) = 0$. The fifth-best outcome, therefore, consists of the AA implementing a policy such that the entrant's entry cost I_E is larger than the profits he makes by entering. This makes the generic producer stay out of the market and, therefore, it makes the originator invest. Therefore $CS^0(E_{out}) = \underline{S}$. In this case, therefore, the AA has an incentive to reduce the entrant's profits, to make the originator invest (a monopoly is better than nothing).

6) The worst outcome for CS occurs when the originator's investment cost I_O is higher than the monopoly profits H . In this case, nothing can be done to push (at least) the originator to invest, therefore $CS^0(O_{out}) = 0$.

The objective of the AA is to maximize CS. The full characterization of the optimal policy is complex (see Appendix 5.3). The following proposition focuses on the cases where allowing reverse payments is optimal.

Proposition 2 *Allowing reverse payments is optimal when:*

²⁰ $CS^0(\bar{\theta})$ is smaller than $CS^0(\underline{\theta})$ if and only if $\lambda \geq \frac{\frac{C_O}{H-L}}{(\bar{\theta} - \underline{\theta}) + \frac{C_O}{H-L}}$. We assume that this is the case.

²¹ The ranking $CS^0(\bar{\theta}) > CS^{\hat{R}}(\underline{\theta})$ is true as long as $\hat{R} > (1 + \bar{\theta} - \underline{\theta})(H - L) - \frac{(1-\lambda)}{\lambda}C_O$, otherwise it is reversed.

(1) $\pi_O(\underline{\theta}) < I_O \leq \pi_O(\bar{\theta})$, (1.a) $\lambda \leq \lambda^0$ and (1.a.i) $I_E > \pi_E^0(\bar{\theta})$: the optimal reverse payment \hat{R} is the one such that $I_E = \pi_E^{\hat{R}}(\bar{\theta})$; (1.b) $\lambda^0 < \lambda \leq \lambda^\infty$: the optimal reverse payment is $\hat{R} = \min\{R \text{ s.t. } I_E \leq \pi_E^{\hat{R}}(\bar{\theta}) \ \& \ \lambda \leq \lambda^{\hat{R}}\}$.

(2) $0 \leq I_O \leq \pi_O(\underline{\theta})$ and (2.a) $\lambda \leq \lambda^0$ and $I_E > \pi_E^0(\bar{\theta})$: the optimal reverse payment \hat{R} is the one such that $I_E = \pi_E^{\hat{R}}(\bar{\theta})$; (2.b) $\lambda^0 < \lambda \leq \lambda^\infty$ and $I_E > \pi_E^0(\underline{\theta})$, then the optimal reverse payment is $\hat{R} = \min\{R \text{ s.t. } I_E \leq \pi_E^{\hat{R}}(\underline{\theta}) \ \& \ \lambda \geq \lambda^{\hat{R}}, R \text{ s.t. } I_E \leq \pi_E^{\hat{R}}(\bar{\theta}) \ \& \ \lambda \leq \lambda^{\hat{R}}\}$; if (2.c) $\lambda > \lambda^\infty$ and $I_E > \pi_E^0(\underline{\theta})$, then the optimal reverse payment \hat{R} is the one such that $I_E = \pi_E^{\hat{R}}(\underline{\theta})$.

Proof and Explanation. In case (1), the originator's investment cost is such that he invests if and only if the entrant has not invested or, having invested, targets the high realization $\bar{\theta}$. Therefore, the objective of the AA is to make the entrant enter and target $\bar{\theta}$. In subcase (1.a), the entrant would target the high realization under any policy, but if reverse payments are banned he does not enter. Therefore, reverse payments must be allowed and their size should just make the entrant's profits match his investment cost. This makes the entrant enter as soon as possible. In case (1.b), the probability of the low realization is such that the entrant targets the low realization if reverse payments are banned, but he targets the high one if the allowed reverse payment is sufficiently large. Therefore, it is optimal to allow reverse payments with a size such that the entrant can recoup his investment cost and he targets the high realization. This policy makes the originator invest, the entrant enter and the monopoly period is minimized.

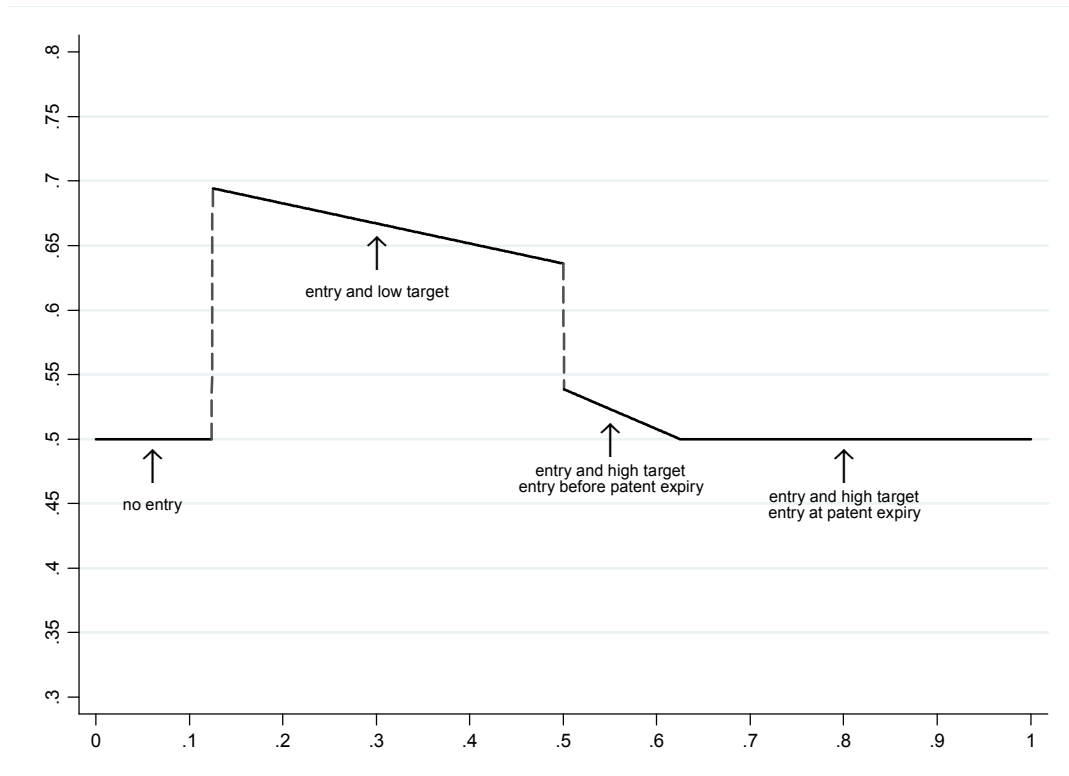
In case (2), the originator's investment cost is so small that he invests for any realization the entrant eventually targets. Therefore, the only problem is to make the entrant enter. Subcase (2.a) is the same as (1.a) above. For a higher probability of the low realization and an investment cost that does not make the entrant enter when reverse payments are banned (2.b), the optimal policy is $\hat{R} = \min\{R \text{ s.t. } I_E \leq \pi_E^{\hat{R}}(\underline{\theta}) \ \& \ \lambda \geq \lambda^{\hat{R}}, R \text{ s.t. } I_E \leq \pi_E^{\hat{R}}(\bar{\theta}) \ \& \ \lambda \leq \lambda^{\hat{R}}\}$. This policy makes the entrant recoup his investment cost with the earliest entry date. For the case (2.c), where the probability of the low realization is so high that the entrant always targets it, the optimal policy is $\hat{R} = R \text{ s.t. } I_E = \pi_E^{\hat{R}}(\underline{\theta})$; so that the entrant can recoup it and enter as soon as possible.

A natural question is what is the role of patent strength. Patent strength has two opposite effects on the optimal policy. On one hand, reverse payments with a strong patent make the additional generic entry ending up in litigation have a lower probability of increasing consumer surplus; on the other one they also induce just a small delay, because the expected entry under litigation is already late. In other words, a strong patent makes both the benefits and the costs of reverse payments smaller. These two opposite forces make the impact of patent strength on the optimal treatment of reverse payments ambiguous. Appendix 5.4 discusses this. The next section provides simulations of different scenarios.

3 Simulations

In order to test the intuitions set out above, we run a number of hypothetical scenarios. Specifically, we assume demand is linear of the form $P = 2 - Q$ and that values that the low and high realizations may take are respectively 0.4 and 0.5. We then allow a number of key parameters to vary such as: the entrant's investment cost; the intensity of competition between the originator and the entrant (modelled by a conjectural variation parameter); the uncertainty of the patent strength; and the degree of symmetry between the originator and the entrant (modelled by variations in the entrant's marginal cost). Finally we assume that the probability of the low realization is 0.5. While this set up is simple, it is nonetheless a helpful way to test our earlier intuitions.

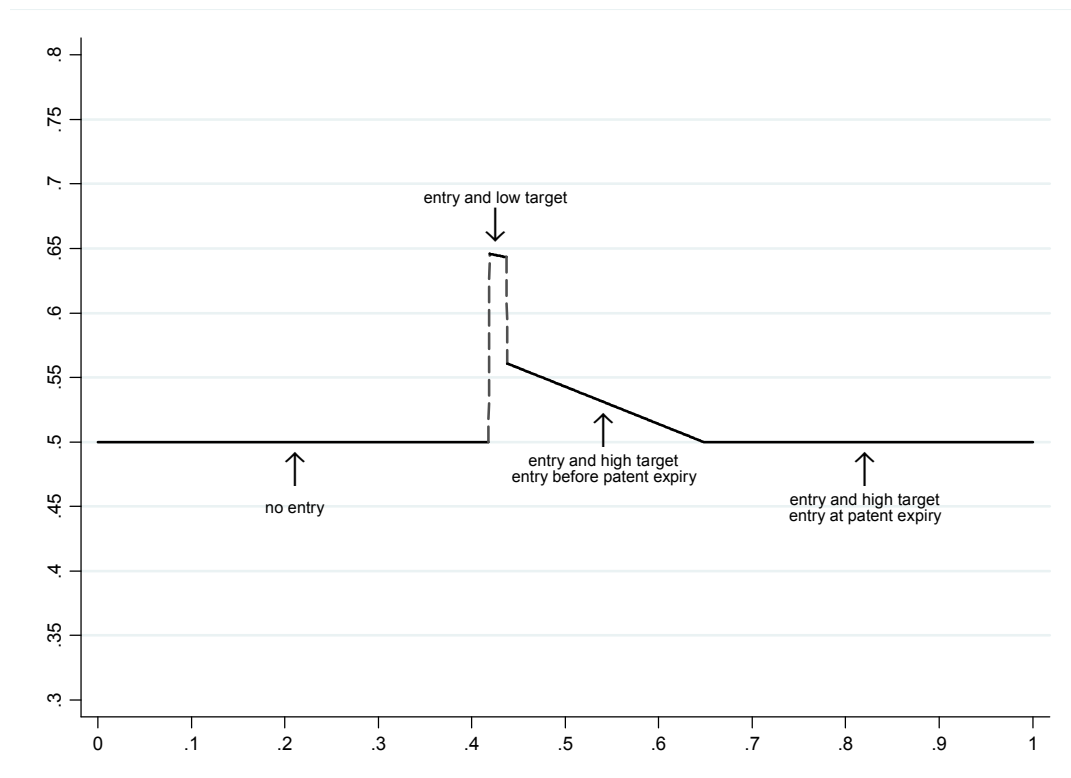
For each case we consider, we identify expected consumer surplus (which we refer to as “consumer surplus” for convenience). The baseline scenario has a conjectural variation parameter of 0 (Cournot competition), the originator and the entrant's marginal cost are 0, the entrant's investment cost I_E is 25% of the monopoly profits H (which are equal to 1, given the demand function and the originator's marginal cost) and litigation costs are assumed to be 0. These parameters make consumer surplus under monopoly equal to 0.5. We evaluate consumer surplus (y-axis) as a function of the allowed reverse payment (x-axis), considered as a fraction of the entrant's profits E . This baseline simulation yields the following graph.



This scenario shows that when reverse payments are banned, or too low (below 12% of the entrant's profits), the entrant cannot recoup her investment cost. This makes the

originator be the monopolist. For a sufficiently high reverse payment – greater than 12% of the entrant’s profits but below 51% – the entrant can recoup her investment cost, she invests and makes an aggressive settlement proposal. The aggressiveness of the proposal – targetting the low realization of the patent strength – is due to the reverse payment being still relatively small. This makes the settlement not too attractive, so the entrant tries to get a large share of the pie by asking an early entry on the market. This is the situation where consumer surplus is maximal: the possibility of using a reverse payment makes the entrant invest, asking for an early entry and, moreover, making litigation possible. Consumer surplus declines slowly as the cap on reverse payments increases. The reason of the decline is that a higher reverse payment must be coupled with a later entry date (to make the originator accept the settlement). When the cap is sufficiently high (51% of the entrant’s profits), the entrant prefers to make a less aggressive settlement proposal, in order to be sure she can reap the reverse payment. Now the entrant targets the high realization of the patent strength and litigation is not possible. However, when the cap is not too high (between 51% and 63% of the entrant’s profits), consumer surplus is still higher than in the monopoly case, because the entrant enters before patent expiry. For a cap higher than that, the entrant enters at patent expiry and consumer surplus is at the monopoly level.

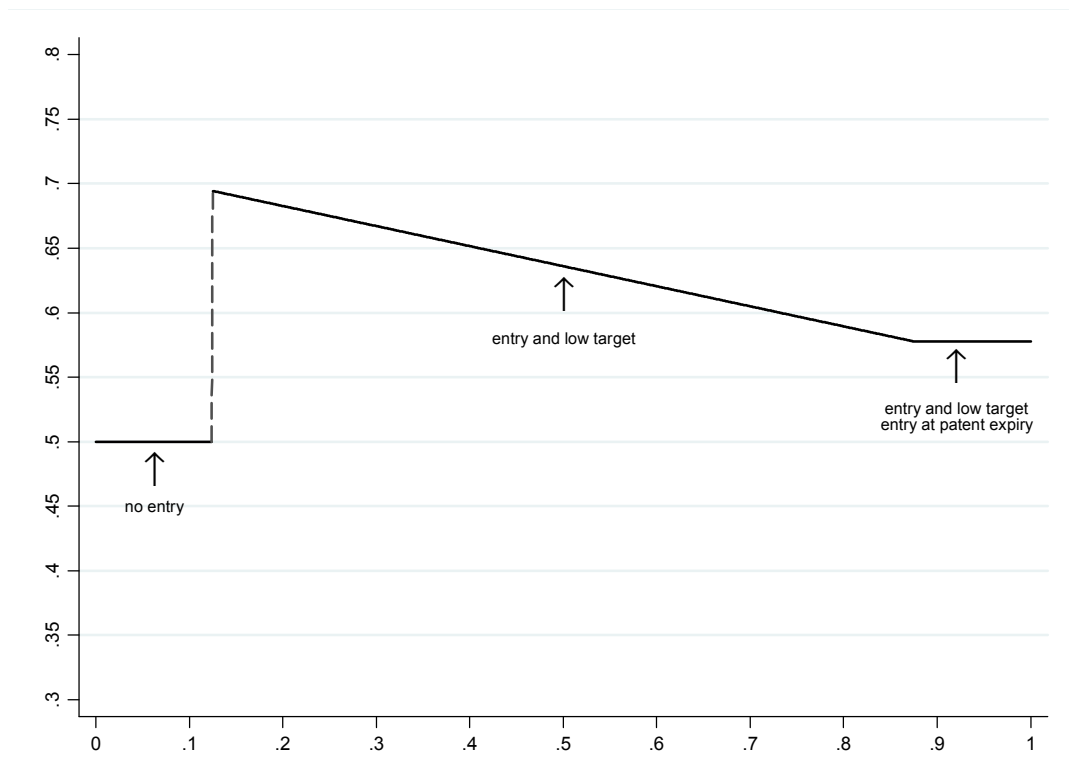
We have also run other simulations. The second one introduces a cost disadvantage for the entrant: all the parameters are the same as above, except the entrant’s marginal cost, which becomes 0.03.



The qualitative conclusions remain the same, but the area of no entry widens, the area

of entry with low target shrinks and the area of entry with high target grows. The intuition is that, given that now the entrant has a cost disadvantage, competing on the market with the originator is less attractive for her. This makes her less willing to ask for an aggressive settlement, which undermines her incentives to invest in the first place (larger no entry area). When the reverse payment allows the recoupment of the investment cost, she prefers to target the low realization in a much tinier parameter set, in order to reduce the risk that she actually has to compete with the originator. On the other hand, entering and making a less aggressive proposal becomes even more appealing.

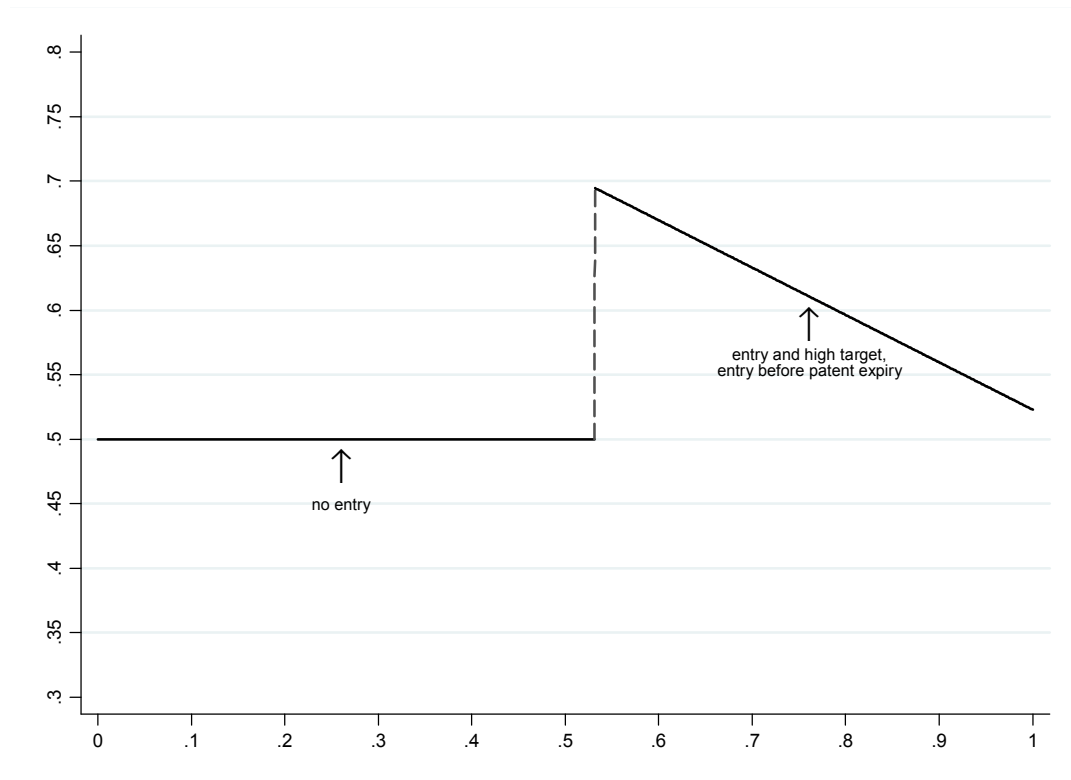
The third simulation uses the parameters of the baseline simulation but increases the uncertainty over the patent strength (keeping the expected patent strength the same). Here, instead of $\underline{\theta} = 0.4$ and $\bar{\theta} = 0.5$, we have $\underline{\theta} = 0.3$ and $\bar{\theta} = 0.6$.



When the uncertainty over the patent strength grows, it becomes more costly to make sure that litigation will not occur, as litigation depends on the high realization – now higher. Therefore it becomes more appealing for the entrant to be aggressive and try to get a profitable settlement by targetting the low realization. In this case, it is convenient for the entrant to be aggressive even with a reverse payment so high that entry would happen at patent expiry (R at least equal to 88% of the entrant's profits). Even for such a reverse payment, consumer surplus is higher than under monopoly, as litigation remains possible.

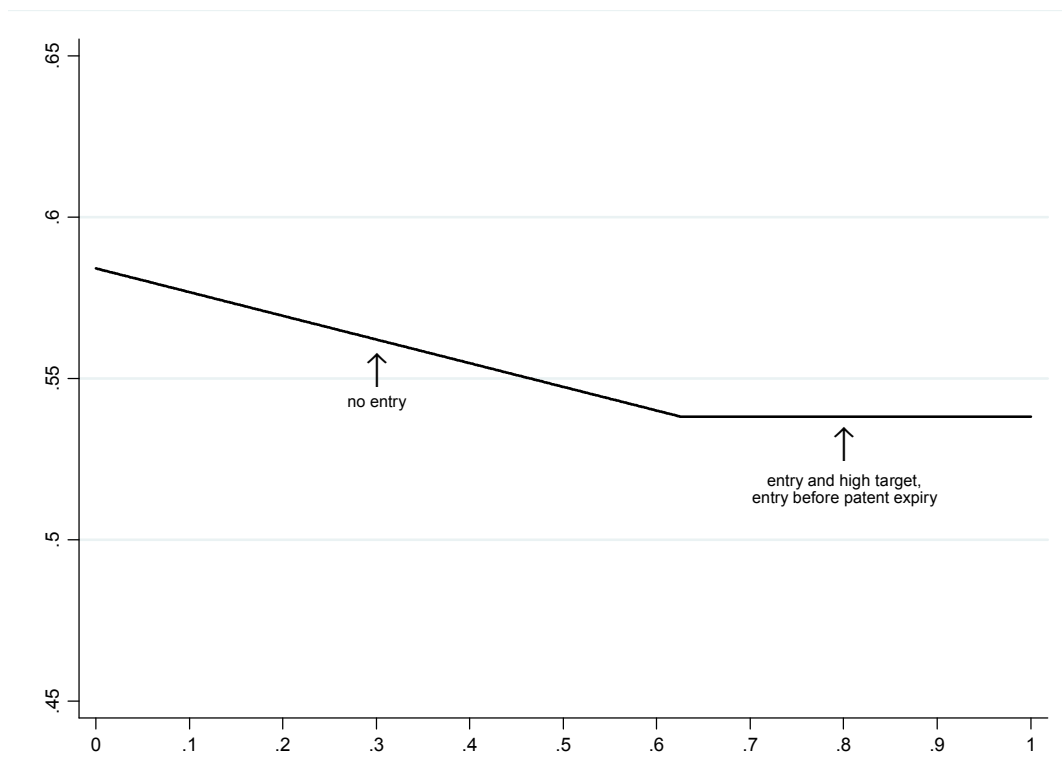
The fourth simulation changes the conjectural variation parameter of the baseline case

to mimick higher intensity of competition. Here we use a value of -0.5 (recall that -1 means Bertrand competition, 0 means Cournot competition and $+1$ means full collusion).



When competition on the market is fierce, the entrant has lower incentives to enter, as prices will be pushed towards marginal cost. Therefore the no entry area becomes wider and, when reverse payment is so high that the entrant invests, she makes a less aggressive settlement proposal, in order to avoid having to actually compete fiercely. However, also in this case, the entry of the generic firm increases consumer surplus, as entry will occur before patent expiry.

The fifth simulation, on the contrary, depicts a scenario of soft competition. We use a conjectural variation parameter of $+0.5$.



In this scenario, the possibility that the settlement offer is rejected by the originator and that parties may have to litigate and actually compete is just a weak deterrent for aggressive proposals. Therefore the entrant has higher incentives to invest – in this simulation she finds it optimal to invest even in absence of reverse payments – and to target the low realization. When the cap on reverse payments is low, the entrant enters before patent expiry, while if it is large the entrant enters at patent expiry. In this simulation a ban on reverse payments would increase consumer surplus, as the entrant would invest in any case and the entry delay is minimized.

In terms of the optimal policy, as these simulation show, we find that the relationship between consumer surplus and the cap on reverse payments is highly non-monotonic and highly dependent on each specific situation. A simple ban on reverse payments seems therefore a too simple rule, as it may reduce consumer surplus – especially when the entrant faces high investment costs or a marginal cost higher than the originator.

4 Discussion and conclusions

When the investment decisions of the originator and the generic manufacturer are taken into account and there is asymmetric information over the patent strength, banning reverse payments may reduce consumer surplus. The reason is two-fold. First, banning reverse payments reduces the industry profits, which reduces the entrant's incentives to invest. This

reduces consumer surplus both because the decreased generic investment implies less litigations (avoiding, therefore, the possibility of the generic entering the market) and because no generic entry occurs before patent expiry. This result is robust to other types of asymmetric information and to any bargaining rule between the originator and the entrant, except when the originator makes a take-it-or-leave-it offer. There is no need that it be the originator who receives a private signal – results hold also if it is the entrant who gets the private signal. Results just requires the entrant to get some, however small, additional surplus from the settlement compared to litigation. The only necessary features for the results are some asymmetric information between the parties and that the informed party does not make a take-it-or-leave-it offer.²² Second, reverse payments creates a tension in the originator's incentives to invest. They make industry profits larger, which can make the entrant willing to enter, which reduces the originator's incentives to invest, but also they make the entrant keener on settling on more favorable terms to the originator, which increases the originator's incentives to invest. This result is robust to any bargaining rule between the originator and the entrant.²³

The main result is that allowing reverse payments *delays* generic entry but *increases* it: when reverse payments are actually used, entry is delayed compared to when generic entry would have occurred without it, but the very possibility of using them increases it. This suggests that a rule of reason is more suited than a ban *per se*. Note, moreover, that even a *laissez-faire* policy that allows complete freedom over the use of reverse payments may be superior to a ban *per se*. Finally, note that patent strength has an ambiguous impact on the optimal policy. Two forces are present: on one hand, a reverse payment on a strong patent makes it unlikely that litigation ends up in generic entry; on the other one, a reverse payment on a strong patent also involves a small cost, given that the additional delay with respect of the expected one is small. For a more detailed discussion of patent strength and optimal policy, see Appendix 5.4. The possibility that several entrants exist reduce the utility of reverse payments for the originator, which makes him use them less often. Appendix 5.5 discusses this.

Though it is difficult to derive an easily feasible rule, we can derive some suggestions. The

²²To see that this result holds also when the entrant gets the private signal, consider the following example. Now the entrant gets the private signal and the originator makes a take-it-or-leave-it settlement offer. The entrant now earns strictly more than his expected litigation payoffs not because of his bargaining power, but because of his information rent. Reverse payments increase the industry profits and, therefore, the originator's willingness to avoid litigation. This makes the originator keener on settling on more favorable terms to the entrant. This increases the entrant's profits, thus increasing his incentive to enter the market. Litigation is still possible and, therefore, the main result that reverse payments delay entry to increase entry still holds.

²³Any other bargaining rule that gives some settlement surplus to the originator still yields the tension explained above. This rule would reduce the negative impact on the originator's incentives when the entrant enters (because the originator is now able to extract some additional surplus from the settlement), but it would also reduce the willingness of the entrant to settle on favorable terms to the originator (exactly because the entrant, now, enjoys less profits). Qualitatively the tension for the originator would still exist.

higher the cap on reverse payments, the higher the profits the generic can reap. This may motivate the generic to enter the market and, as long as the cap is relatively small, make her offer an aggressive settlement proposal to the originator. When this happens, consumer surplus is maximal. If the cap is slightly higher than the optimal one, entry is delayed in case of settlement, but still the entrant will make an aggressive proposal, making litigation possible. Consumer surplus remains relatively high, as entry occurs before patent expiry and litigation is possible. If the cap goes beyond a certain point, then the reverse payment is so big that it becomes convenient for the entrant to make a generous settlement proposal to the originator, to make sure they can split monopoly profits without the risk of litigating. Here consumer surplus decreases, but still remains higher than in the monopoly scenario – as entry occurs before patent expiry if the cap is not too high. For a cap even higher, entry will occur at patent expiry and parties will share monopoly profits. This suggests that a very high reverse payment should be regarded as probably anticompetitive, while a smaller one may be procompetitive. If the entrant has a high investment cost, is perceived of lower quality or has a cost disadvantage, the cap on reverse payments should be higher to make the generic invest. The policy on reverse payments should be more lenient.

In all of these cases, an excessively high reverse payment keeps consumer surplus at the monopoly level. However, often an intermediate reverse payment increases consumer surplus, suggesting that a rule of reason is better than a simple ban *per se*.

References

- [1] Bebchuk L.A (1984), "Litigation and Settlement under Imperfect Information", *RAND Journal of Economics* 15, 404-415.
- [2] Bessen J.E. and Meurer M.J., (2006). "Patent Litigation with Endogenous Disputes," *American Economic Review*, American Economic Association, vol. 96(2), pages 77-81, May.
- [3] Binmore K.G. and Herrero M.J. (1988a), "Matching and Bargaining in Dynamic Markets", *Review of Economic Studies* 55, 17-31.
- [4] Binmore K.G. and Herrero M.J. (1988b), "Security Equilibrium", *Review of Economic Studies* 55, 33-48.
- [5] Cho, I-K. & Kreps, D. M. (1987), "Signaling games and stable equilibria", *Quarterly Journal of Economics* 102:179-221."
- [6] Daughety A.F. and Reinganum J.F. (1994), "Settlement Negotiations with Two-Sided Asymmetric Information: Model Duality, Information Distribution and Efficiency", *Game Theory and Information*, 9403009, EconWPA.

- [7] Gale D. (1986a), "Bargaining and Competition Part I: Characterization", *Econometrica* 54, 785-806.
- [8] Gale D. (1986b), "Bargaining and Competition Part II: Existence", *Econometrica* 54, 807-818.
- [9] Gale D. (1987), "Limit Theorems for Markets with Sequential Bargaining", *Journal of Economic Theory*, 43, 20-54.
- [10] Gratz L. (2012), "Economic Analysis of Pay-for-delay Settlements and Their Legal Ruling", Munich Discussion Paper No. 2012-6.
- [11] Lemley M.A. and Shapiro C. (2005), "Probabilistic Patents", *Journal of Economic Perspectives*, Vol. 19, Number 2 — Spring 2005—Pages 75–98
- [12] Leonard G. and & Mortimer R., "Antitrust Implications of Pharmaceutical Patent Litigation Settlements", in *Economic Approaches to Intellectual Property Policy, Litigation and Management*, 251.
- [13] P'Ng I.P.L. (1983), "Strategic Behavior in Suit, Settlement and Trial", *Bell Journal of Economics*, 14, 539-550.
- [14] Reinganum J.F. and Wilde L.L. (1986), "Settlement, Litigation and the Allocation of Litigation Costs", *RAND Journal of Economics*, 17, 557-566.
- [15] Rubinstein A. and Wolinsky A. (1985), "Equilibrium in a Market with Sequential Bargaining", *Econometrica* 53, 1133-1150.
- [16] Rubinstein A. and Wolinsky A. (1990), "Decentralized Trading, Strategic Behavior and the Walrasian Outcome", *Review of Economic Studies* 57, 63-78.
- [17] Salant S.W. (1984), "Litigation of Settlement Demands Questioned by Bayesian Defendants", working paper 516, California Institute of Technology, Div. of the Humanities and Social Sciences.
- [18] Salant S.W. and Rest G. (1982), "Litigation of Questioned Settlement Claims: A Bayesian Nash-Equilibrium Approach", Rand Corporation Discussion Paper p-6809.
- [19] Schweizer Urs (1989), "Litigation and Settlement under Two-Sided Incomplete Information", *The Review of Economic Studies*, Vol. 56, No. 2, pp. 163-177.
- [20] Shapiro, C. (2003) " Antitrust Limits to Patent Settlements," *RAND Journal of Economics*, The RAND Corporation, Vol. 34(2), pages 391-411, Summer.
- [21] Willig R. and Bigelow J.P. (2004), "Antitrust Policy towards Agreements that settle Patent Litigation", *The Antitrust Bulletin*, Fall.

5 Appendix

5.1 Reverse payments and late entry

Proof of Lemma 1. Compare the profits the entrant obtains from offering the reverse payment $R = \hat{R}$ and the associated entry date $D(\hat{\theta}) = \hat{\theta} + \frac{\hat{R}-C_O}{H-L}$ with the ones from asking for a smaller reverse payment $\hat{R} < \hat{R}$ and the associated entry date $D(\hat{\theta}) = \hat{\theta} + \frac{\hat{R}-C_O}{H-L}$. Note that, from (1), $D(\hat{\theta})$ is the optimal entry date given the lower reverse payment, as it keeps the originator indifferent between accepting and refusing the offer for the (possibly new) targeted realization $\hat{\theta}$. Consider, first, the case where the entrant targets the same realization: $\hat{\theta} = \hat{\theta}$. In this case the probability of litigation is the same, so we can just compare the entrant's settlement profits from the offer $\{\hat{R}, D(\hat{\theta})\}$ with the ones from the alternative offer $\{\hat{R} < \hat{R}, D(\hat{\theta})\}$. Note that a lower reverse payment implies an earlier entry date: a marginally lower R makes the entrant lose dR through R and gain $\frac{E}{H-L}dR$. Being the loss in the originator's profits $(H-L)$ higher than the entrant's profits E , the optimal R is the maximal possible one: $R = \hat{R}$. Consider now the case where the entrant, as a consequence of lowering R from \hat{R} to \hat{R} , targets a different realization: $\hat{\theta} \neq \hat{\theta}$. This represents a further distortion with respect to the optimal offer: even under this different target, $\hat{R} < \hat{R}$ is not optimal, because the entrant could now raise the reverse payment \hat{R} to \hat{R} and extract the additional originator's surplus through it. Again, therefore, the optimal R is \hat{R} . This Lemma holds under any bargaining rule.²⁴

5.2 No Entry Delay

Here we show that the entrant enters as soon as he can (footnote 12). Change the notation in the following way: 0 is now the date when the entrant is ready to enter and T the entry date he actually chooses. D and 1 remain, respectively, the entry date and the patent expiry date. The entrant now can choose the moment T when he will make an offer to the originator. In order to simplify the notation, just assume that the patent strength is common knowledge.²⁵

Now, if the parties *litigate*, they expect to obtain:

$$\text{Originator: } TH + (1 - T)[\theta H + (1 - \theta)L] - C_O$$

$$\text{Entrant: } (1 - T)(1 - \theta)E - C_E$$

The only additional element, here, is T : the larger T , the higher the monopoly profits the originator earns before the settlement-litigation decision.

If the parties *settle*, they obtain:

²⁴The larger is \hat{R} , the longer the monopoly period - it is in both parties' interest to make it as long as possible.

²⁵Results, as will become clear, do not depend on this.

Originator: $TH + (1 - T)[DH + (1 - D)L] - R$

Entrant: $(1 - T)(1 - D)E + R$

The originator settles if and only if this is more profitable than litigating, i.e. if

$$TH + (1 - T)[DH + (1 - D)L] - R > TH + (1 - T)[\theta H + (1 - \theta)L] - C_O.$$

Substituting R with \hat{R} , we have:

$$D \geq D^*(T) = \theta + \frac{\hat{R} - C_O}{(1 - T)(H - L)}.$$

The minimal entry date that the originator is ready to accept is increasing in T , as long as $\hat{R} > C_O$. In this case, the larger the time elapsed between the moment when the entrant is ready to enter and the moment when he discusses the settlement with the originator, the later the entry date the originator is willing to accept. In other words, waiting is counterproductive for the entrant, as it only reduces the amount he can get. On the other hand, when $\hat{R} < C_O$, the minimal entry date that the originator is willing to accept is decreasing in T . The larger the time elapsed between the moment when the entrant is ready to enter and the moment when he discusses the settlement with the originator, the earlier he can actually enter (in case of settlement) *relatively to the (lower) patent validity*. This positive effect for the entrant must be weighted with the later date T when the settlement is discussed. Substituting $D^*(T)$ in the entrant's profits, we get $\pi_E^{\hat{R}} = (1 - T)(1 - \theta - \frac{\hat{R} - C_O}{(1 - T)(H - L)})E + \hat{R}$. Its derivative with respect to T is $\frac{d\pi_E^{\hat{R}}}{dT} = -(1 - \theta)E < 0$, therefore the entrant prefers to enter as soon as possible.

In conclusion, the entrant chooses $T = 0$ and discusses the settlement as soon as possible.

5.3 Optimal Policy

This proposition gives the full characterization of the optimal policy.

Proposition 3 *An optimal policy²⁶ is the following:*

if (1) $I_O > H$, any policy is equivalent;

if (2) $\pi_O(\bar{\theta}) < I_O \leq H$, then $\hat{R} = 0$;

if (3) $\pi_O(\underline{\theta}) < I_O \leq \pi_O(\bar{\theta})$, then if (3.a) $\lambda \leq \lambda^0$ and (3.a.i) $I_E \leq \pi_E^0(\bar{\theta})$ then $\hat{R} = 0$; if (3.a.ii) $\pi_E^0(\bar{\theta}) < I_E \leq \pi_E^\infty(\bar{\theta})$, then $\hat{R} = R$ s.t. $I_E = \pi_E^{\hat{R}}(\bar{\theta})$; if (3.a.iii) $I_E > \pi_E^\infty(\bar{\theta})$ then any policy is equivalent; if (3.b) $\lambda^0 < \lambda \leq \lambda^\infty$, then $\hat{R} = \min\{R \text{ s.t. } I_E \leq \pi_E^R(\bar{\theta}) \text{ \& } \lambda \leq \lambda^R\}$; if (3.c) $\lambda > \lambda^\infty$, then any policy is equivalent;

if (4) $0 \leq I_O \leq \pi_O(\underline{\theta})$, (4.a) $\lambda < \lambda^0$ and (4.a.i) $I_E \leq \pi_E^0(\bar{\theta})$, then $\hat{R} = 0$; if (4.a.ii) $\pi_E^0(\bar{\theta}) < I_E \leq \pi_E^\infty(\bar{\theta})$, then $\hat{R} = R$ s.t. $I_E = \pi_E^R(\bar{\theta})$; if (4.a.iii) $I_E > \pi_E^\infty(\bar{\theta})$, then any

²⁶It is "an" optimal policy because there can be other policies that yield the same CS. For example, in case (2), any policy other than $\hat{R} = 0$ yields the same CS if $I_E > \max\{\pi_E^{\hat{R}}(\bar{\theta}), \pi_E^{\hat{R}}(\underline{\theta})\}$, because the generic producer would not enter in any case.

policy is equivalent; if (4.b) $\lambda^0 < \lambda < \lambda^\infty$ and (4.b.i) $I_E \leq \pi_E^0(\underline{\theta})$ then $\hat{R} = 0$; if (4.b.ii) $\pi_E^0(\bar{\theta}) < I_E \leq \pi_E^\infty(\bar{\theta})$ then $\hat{R} = \min\{R \text{ s.t. } I_E \leq \pi_E^R(\underline{\theta}) \ \& \ \lambda > \lambda^\infty, R \text{ s.t. } I_E \leq \pi_E^R(\bar{\theta}) \ \& \ \lambda < \lambda^R\}$; if (4.b.iii) $I_E > \pi_E^\infty(\bar{\theta})$, any policy is equivalent; if (4.c) $\lambda > \lambda^\infty$ and (4.c.i) $I_E \leq \pi_E^0(\underline{\theta})$, then $\hat{R} = 0$; if (4.c.ii) $\pi_E^0(\bar{\theta}) < I_E \leq \pi_E^\infty(\bar{\theta})$ then $\hat{R} = R \text{ s.t. } I_E \leq \pi_E^{\hat{R}}(\underline{\theta})$; if (4.c.iii) $I_E > \pi_E^\infty(\bar{\theta})$, any policy is equivalent.

Proof and Explanation. In case (1) the originator's investment cost is higher than monopoly profits: trivially, nothing can be done to push him to invest. In case (2), where the originator's investment cost is smaller than monopoly profits but higher than the profits he makes when the entrant is present, the AA needs to deter the generic manufacturer's entry. The originator's investment cost is, indeed, higher than the profits he would get if the entrant entered - even if he targeted the high realization $\bar{\theta}$. Reducing the entrant's profits as much as possible, through $\hat{R} = 0$, is the best the AA can make in order to make the entrant not invest and, therefore, to make the originator invest. This case highlights the *negative* impact of reverse payments on investment, which is absent from the traditional patent literature. In case (3), the originator invests if and only if the entrant has not invested or, having invested, targets the high realization $\bar{\theta}$. Therefore, the objective of the AA is to make the entrant enter and target $\bar{\theta}$. In subcase (3.a), the entrant targets $\bar{\theta}$ upon entry, which occurs if his investment cost is sufficiently small (3.a.i): in this case the optimal policy is $\hat{R} = 0$. If the entrant's investment cost is intermediate (3.a.ii), then the AA must set a policy that allows the entrant to recoup it. In this case the optimal policy is $\hat{R} = R \text{ s.t. } I_E = \pi_E^R(\bar{\theta})$. This policy makes the entrant recoup his investment *and* target the high realization $\bar{\theta}$. If the entrant's investment cost is high (3.a.iii), then any policy is indifferent, because the generic manufacturer would never enter and the originator would invest anyway (getting $H > I_O$). In case (3.b), the entrant would target $\underline{\theta}$ under $\hat{R} = 0$ and $\bar{\theta}$ for a sufficiently high reverse payment: the optimal policy is therefore to allow reverse payments with a cap such that the entrant still targets the high realization and recoups his investment cost. In case (4), the originator's investment cost is so small that he would invest for any realization the entrant may target. Therefore, when the probability of the low realization is sufficiently small (4.a), an optimal policy is $\hat{R} = 0$. The entrant would target the high realization under both policies, so it is better to make him target it when reverse payments are banned. If the entrant's investment cost is higher (4.a.ii and 4.a.iii), then the optimal policy is to allow reverse payments in order to make the entrant recoup his investment cost, when possible. When the probability of the low realization is slightly higher (4.b), the optimal policy is similar to the previous case. The only difference is that, when the profits are intermediate (4.b.ii), the optimal reverse payment makes the entrant target the low realization, if this allows him to recoup the investment cost. In other words, this policy makes the entrant recoup his investment cost with the earliest entry date and keeps him willing to target the low realization, or just implies the earliest possible entry date if it is impossible to make the entrant recoup his investment and target the low realization. For the case (4.c), where the probability of the low realization is so high

that the entrant always targets the low one, then the optimal policy is very similar to the previous case, with the only difference that $\pi_E^R(\bar{\theta})$ may be disregarded and only $\pi_E^R(\underline{\theta})$ needs to be considered when the entrant's investment cost is intermediate.

5.4 Patent Strength and Optimal Policy

This paragraph discusses the ambiguous impact of the patent strength θ on the optimal policy. Recall that there are two opposite forces: a strong patent makes the generic entry ending up in litigation have a lower probability of increasing consumer surplus, but also it makes reverse payments induce just a small delay, because the expected entry under litigation is already late. We show here how each area of the optimal policy changes after an increase in patent strength. Assume, for simplicity, that an increase in patent strength θ consists of an increase of the same size of the two realizations $\underline{\theta}$ and $\bar{\theta}$.

In Area (1) of Proposition 4 nothing can be done to induce the originator to invest, so the patent strength has no impact.

In Area (2) it is optimal to ban reverse payments: the condition $\pi_O(\bar{\theta}) < I_O \leq H$ can be rewritten as $\bar{\theta}H + (1 - \bar{\theta})L - C_O < I_O \leq H$, so an increase in θ makes the area smaller.

In Area (3) the optimal policy depends on the size of λ and I_E : the condition of this area can be rewritten as $\underline{\theta}H + (1 - \underline{\theta})L - C_O < I_O \leq \bar{\theta}H + (1 - \bar{\theta})L - C_O$. Note, first, that an increase in θ has no impact on its size - both sides increase by the same amount. When (3.a) $0 < \lambda \leq \lambda^0$, which can be rewritten as $0 < \lambda \leq \frac{EC_O + (H-L)C_E}{(\bar{\theta} - \underline{\theta})E(H-L) + EC_O + (H-L)C_E}$, an increase in θ has no impact either. When (3.a.i) $0 \leq I_E \leq \pi_E^0(\bar{\theta})$, which can be rewritten as $0 \leq I_E \leq (1 - \bar{\theta} + \frac{C_O}{H-L})E$, an increase in θ reduces the area, so, like in area (2), an increase in θ reduces the parameter region where a ban on reverse payments is optimal. However, this is not the case when (3.a.ii) $\pi_E^0(\bar{\theta}) \leq I_E \leq \pi_E^\infty(\bar{\theta})$, which can be rewritten as $(1 - \bar{\theta} + \frac{C_O}{H-L})E \leq I_E \leq (1 - \bar{\theta} + \frac{C_O}{H-L})(H-L)$. An increase in θ makes this area smaller, like above, but here it is optimal to allow reverse payments. The area that grows is the one where (3.a.iii) $I_E > \pi_E^\infty(\bar{\theta})$, i.e. $I_E > (1 - \bar{\theta} + \frac{C_O}{H-L})(H-L)$: in this area any policy is ineffective, as it is impossible to induce the generic manufacturer to enter. When (3.b) $\lambda^0 < \lambda \leq \lambda^\infty$, which can be rewritten as

$$\frac{EC_O + (H-L)C_E}{(\bar{\theta} - \underline{\theta})E(H-L) + EC_O + (H-L)C_E} < \lambda \leq \frac{(1 - \bar{\theta})(H-L) - (1 - \bar{\theta})E + C_O + C_E}{(1 - \underline{\theta})(H-L) - (1 - \bar{\theta})E + C_O + C_E},$$

the optimal policy is to allow reverse payments and an increase in θ reduces the area. When (3.c) $\lambda > \lambda^\infty$, i.e. $\lambda > \frac{(1 - \bar{\theta})(H-L) - (1 - \bar{\theta})E + C_O + C_E}{(1 - \underline{\theta})(H-L) - (1 - \bar{\theta})E + C_O + C_E}$, no policy can make the entrant target the high type $\bar{\theta}$, which means that the originator will never enter at the first place. An increase in θ makes this area larger.

In Area (4), when (4.a) $\lambda \leq \lambda^0$ the optimal policy is to ban reverse payments and an increase in θ has no impact. In areas (4.a.i), where it is optimal to ban reverse payments, and in (4.a.ii), where it is optimal to allow them, the impact is ambiguous, because a higher θ makes the condition for I_O larger but the one for I_E smaller. In area (4.a.iii), where the

policy has no effects, a larger θ makes the area larger. When (4.b) $\lambda^0 < \lambda \leq \lambda^\infty$, it is optimal to allow reverse payments and an increase in θ makes the area smaller. In the subcase of (4.b) the impact is always ambiguous because θ makes the area of I_O larger but the one of λ smaller.

When (4.c) $\lambda > \lambda^\infty$, an increase in θ has an ambiguous effect too: when (4.c.i) $I_E \leq \pi_E^0(\underline{\theta})$, the optimal policy is to ban reverse payments and an increase in θ makes the area of the originator's investment cost larger, but the areas of λ and the entrant's investment cost smaller; and when (4.c.ii) $I_E > \pi_E^0(\underline{\theta})$, the optimal policy is to allow reverse payments and an increase in θ makes the areas of the originator and the entrant's investment costs larger, but the area of λ smaller. In (4.c.iii) θ makes the parameter area larger and the policy is ineffective.

The overall interaction between the optimal policy and the patent strength is, therefore, ambiguous and no robust policy implications on patent strength can be derived.

The following table resumes the results. The columns for I_O , and I_E represent the range of these parameters in each area, while column \hat{R}^* represents the optimal policy. *Ind* means that any policy is indifferent on CS, *yes* means that the optimal policy sets a positive cap on R and *no* means that the optimal policy is a ban on reverse payments. The arrows (and the equality signs) in the last column represent the impact of the patent strength on the areas of, respectively, the originator's investment cost I_O , the probability of the low patent strength realization λ and the entrant's investment cost I_E . The question marks mean that an increase in patent strength has an ambiguous impact on the considered area.

Area	I_O	λ	I_E	\hat{R}^*	$\frac{d(Area)}{d\theta}$
1	$[H, +\infty)$	any	any	ind.	=
2	$[\pi_O(\bar{\theta}), H]$	any	any	no	\downarrow
3.a.i	$[\pi_O(\underline{\theta}), \pi_O(\bar{\theta})]$	$[0, \lambda^0]$	$[0, \pi_E^0(\bar{\theta})]$	no	$\downarrow (=, =, \downarrow)$
3.a.ii	$[\pi_O(\underline{\theta}), \pi_O(\bar{\theta})]$	$[0, \lambda^0]$	$[\pi_E^0(\bar{\theta}), \pi_E^\infty(\bar{\theta})]$	yes	$\downarrow (=, =, \downarrow)$
3.a.iii	$[\pi_O(\underline{\theta}), \pi_O(\bar{\theta})]$	$[0, \lambda^0]$	$[\pi_E^\infty(\bar{\theta}), +\infty)$	ind.	$\uparrow (=, =, \uparrow)$
3.b	$[\pi_O(\underline{\theta}), \pi_O(\bar{\theta})]$	$[\lambda^0, \lambda^\infty]$	any	yes	$\downarrow (=, \downarrow)$
3.c	$[\pi_O(\underline{\theta}), \pi_O(\bar{\theta})]$	$[\lambda^\infty, +\infty]$	any	ind.	$\uparrow (=, \uparrow)$
4.a.i	$[0, \pi_O(\underline{\theta})]$	$[0, \lambda^0]$	$[0, \pi_E^0(\bar{\theta})]$	no	? ($\uparrow, =, \downarrow$)
4.a.ii	$[0, \pi_O(\underline{\theta})]$	$[0, \lambda^0]$	$[\pi_E^0(\bar{\theta}), \pi_E^\infty(\bar{\theta})]$	yes	? ($\uparrow, =, \downarrow$)
4.a.iii	$[0, \pi_O(\underline{\theta})]$	$[0, \lambda^0]$	$[\pi_E^\infty(\bar{\theta}), +\infty)$	ind.	? ($\uparrow, =, \uparrow$)
4.b.i	$[0, \pi_O(\underline{\theta})]$	$[\lambda^0, \lambda^\infty]$	$[0, \pi_E^0(\underline{\theta})]$	no	? ($\uparrow, \downarrow, \downarrow$)
4.b.ii	$[0, \pi_O(\underline{\theta})]$	$[\lambda^0, \lambda^\infty]$	$[\pi_E^0(\underline{\theta}), \pi_E^\infty(\bar{\theta})]$	yes	? ($\uparrow, \downarrow, \downarrow$)
4.b.iii	$[0, \pi_O(\underline{\theta})]$	$[\lambda^0, \lambda^\infty]$	$[\pi_E^\infty(\underline{\theta}), +\infty)$	ind.	? ($\uparrow, \downarrow, \uparrow$)
4.c.i	$[0, \pi_O(\underline{\theta})]$	$[\lambda^\infty, +\infty]$	$[0, \pi_E^0(\underline{\theta})]$	no	? ($\uparrow, \uparrow, \downarrow$)
4.c.ii	$[0, \pi_O(\underline{\theta})]$	$[\lambda^\infty, +\infty]$	$[\pi_E^0(\underline{\theta}), \pi_E^\infty(\bar{\theta})]$	yes	? ($\uparrow, \uparrow, \downarrow$)
4.c.iii	$[0, \pi_O(\underline{\theta})]$	$[\lambda^\infty, +\infty]$	$[\pi_E^\infty(\underline{\theta}), +\infty)$	ind.	$\uparrow (\uparrow, \uparrow, \uparrow)$

5.5 No Menu of Contracts

This subsection shows that the entrant cannot write a menu of contracts to make the originator truthfully reveal his type. Consider a candidate menu of contracts $\{(\underline{D}, \underline{R}), (\bar{D}, \bar{R})\}$, where $(\underline{D}, \underline{R})$ is designed for the low type and (\bar{D}, \bar{R}) for the high type. The constraints to fulfill are:

$$\bar{D}H + (1 - \bar{D})L - \bar{R} \geq \underline{D}H + (1 - \underline{D})L - \underline{R}, \quad (IC_{\bar{\theta}})$$

$$\underline{D}H + (1 - \underline{D})L - \underline{R} \geq \bar{D}H + (1 - \bar{D})L - \bar{R}, \quad (IC_{\underline{\theta}})$$

$$\bar{D}H + (1 - \bar{D})L - \bar{R} \geq \bar{\theta}H + (1 - \bar{\theta})L - C_O, \quad (PC_{\bar{\theta}})$$

$$\underline{D}H + (1 - \underline{D})L - \underline{R} \geq \underline{\theta}H + (1 - \underline{\theta})L - C_O. \quad (PC_{\underline{\theta}})$$

The first two inequalities are the incentive compatibility constraints to make each originator's type prefer not to pretend to be the other type. Note that the originator's true type does *not* enter these equations - it only enters the originator's litigation payoff. Therefore the only way to fulfill these incentive compatibility constraints is to make them have the same value. We have, therefore, $\bar{D}H + (1 - \bar{D})L - \bar{R} = \underline{D}H + (1 - \underline{D})L - \underline{R}$, that yields

$$(\bar{D} - \underline{D})(H - L) = \bar{R} - \underline{R}. \quad (IC_{\bar{\theta}} = IC_{\underline{\theta}})$$

The third and the fourth inequalities are the participation constraints that make each type prefer not to litigate. Given that the left hand sides of the four inequalities above must be the same ($IC_{\bar{\theta}} = IC_{\underline{\theta}}$), only the inequality with the larger right hand side can bind. This inequality is $PC_{\bar{\theta}}$, as $\bar{\theta}H + (1 - \bar{\theta})L - C_O$ is larger than $\underline{\theta}H + (1 - \underline{\theta})L - C_O$ because $(\bar{\theta} - \underline{\theta})(H - L) > 0$. Therefore, type $\bar{\theta}$ is left with no rent and type $\underline{\theta}$ enjoys an information rent. Consider now the entrant's profits. Recall that λ is the probability that the originator's type is $\underline{\theta}$. The entrant's problem is:

$$\max_{(\underline{R}, \bar{R})} \pi_E = \lambda[(1 - \underline{\theta} - \frac{\underline{R} - C_O}{H - L})E + \underline{R}] + (1 - \lambda)[(1 - \bar{\theta} - \frac{\bar{R} - C_O}{H - L})E + \bar{R}]$$

$$s.t. (\bar{D} - \underline{D})(H - L) = \bar{R} - \underline{R}.$$

The derivative of π_E with respect to \underline{R} is $\lambda \frac{H-L-E}{H-L} > 0$ and the one with respect to \bar{R} is $(1 - \lambda) \frac{H-L-E}{H-L} > 0$, therefore the entrant asks for the maximal allowed reverse payment. This makes $\underline{R} = \bar{R} = \hat{R}$. This implies (for $IC_{\bar{\theta}} = IC_{\underline{\theta}}$) that we also have $\underline{D} = \bar{D}$. The entry dates \underline{D} and \bar{D} associated with these reverse payments are the ones that make the $PC_{\bar{\theta}}$ binding: $\underline{D} = \bar{D} = \bar{\theta} + \frac{\hat{R} - C_O}{H - L}$. Given that $\underline{D} = \bar{D}$ and $\underline{R} = \bar{R}$, the candidate menu of contracts reduces to a single contract that leaves an information rent to the $\underline{\theta}$ -type. This contract is the

settlement offer that targets $\bar{\theta}$ in the main text. When the probability that the originator's type is $\bar{\theta}$ is too small, it is optimal to "shutdown" this type, offer a contract that extracts all the rent of the $\underline{\theta}$ -type and causes litigation when the originator draws the high realization $\bar{\theta}$. This is exactly the settlement offer that targets $\underline{\theta}$ in the main text.

5.6 Several entrants

This section discusses the case of having n generic firms, instead of one, able to invest and enter the market. If they were to enter simultaneously, they would face a coordination problem – entry when other generics enter is less profitable than if they stay out. This reduces the expected profits of each entrant and, in turn, increases the level of reverse payments needed to make any of the generic entrants willing to invest. At the same time, the existence of many entrants reduce the willingness to pay a reverse payment of the originator. Compared to the single entrant case, the existence of many entrants would require a higher cap to be effective and, in any case, the parameter sets in which the originator invests and uses a reverse payment would shrink. If they enter sequentially, on one hand, each entrant creates a negative externality on the following ones (like in the simultaneous entry case) and, on the other one, it creates an asymmetry between the originator and the generics: every time a generic enters the market, all the players already on the market suffer, while only the originator pays the cost of the reverse payment. The existence of n entrants investing sequentially therefore shrinks the parameter set in which reverse payments are used (compared to the case of single entrant), as it may be in the interest of the originator to allow for the entry of some generics in order to reduce the incentives of the following ones to invest.



Universitat
Pompeu Fabra
Barcelona

