“Pay for Delay”: What do we disagree on?  

Pierre Régibeau  
CRA and Imperial College  

Abstract  
Antitrust concerns about “Pay For Delay” patent settlements are based on two theory of harms, one that stresses the need for Courts to review the validity of patents and one that emphasises the “probabilistic” nature of patent rights. The main weakness of the first theory of harm is that it fails to explain why some forms of patent settlements would be less desirable than others. The “probabilistic” theory of harm raises fundamental questions about the legal obligations of a patent-holdr, the type of uncertainty that should be reflected in the probabilistic nature of the patents and whether the theory can be applied to anything but the simplest PFD settlements. The paper also discusses the likely effect of a PDF ban on innovation and reviews both the European approach to recent and on-going PDF cases and the recent Actavis decision of the US Supreme Court.  

Keywords: Patents, Litigation, Settlements  
JEL Classification: K21, K41, L24, L41.  

1. THEORIES OF HARM  

“Pay for delay” or “value transfer” licensing settlements have been investigated by antitrust authorities in a number of jurisdictions, including the US, the EU and the UK. “Pay for Delay” (PFD) refer to agreements reached between a pharmaceutical firm that produces a drug that is still protected by some patents and a (potential) generic entrant in settlement of litigation about the infringement and/or validity of the relevant patents. The key feature of a PDF agreement is that the generic agrees to enter only after a specified period and receives a positive transfer from the patent holder. The antitrust authorities’ objection to this type of settlement relies generally on two theories of harm.  

According to the first theory, the fact that the patent-holder actually pays the generic challenger decreases the probability that patents will be effectively reviewed in Court. As such review is an integral part of the patent system, this amounts to depriving society of the opportunity to “weed out” weak patents, thereby preserving unwarranted exclusion rights to the detriment of consumers. The second theory of  

1 Disclaimer: This paper should not be taken as reflecting the opinion of CRA or of any of CRA’s clients.  
2 Disclosure: The author has been working on behalf of patent holders in several European “Pay for Delay” cases.
harm – commonly associated with the work of Carl Shapiro\(^3\) – claims that any settlement in which the generic and the patent holder agree to allow independent generic entry prior to patent expiration but which involves a transfer from patent-holder to generic that exceeds the litigation costs expected by the patent-holder must involve a date of generic entry that is later than the “average” date that the patent-holder would – rightly or wrongly – expect to result from the continuation of litigation. This implies that the patent-holder essentially shares some of its expected monopoly rents to “delay” generic entry and that expected consumer surplus is lower than if litigation had been pursued to a final judgement. While this article tries to focus on issues that are common to both the US and the European “versions” of “pay for delay” cases, relatively more emphasis is given to the European approach. It is therefore important to remember that there is no equivalent to the Hatch-Waxman Act in Europe, so that keeping one generic out has no direct effect on keeping other generics out of the market.

2. THE MAIN POINTS OF DISAGREEMENT

I’ll start with what I am not going to claim. I am not going to claim that any form of patent settlement should be acceptable under Antitrust Law. Clearly settlement terms that extend beyond the scope or lifetime of the patents deserve to be closely scrutinised. I would even agree that “value transfer” agreements should also be subject to some oversight as they could otherwise be used to protect extremely weak (or even sham) patents in a manner that would be hard to distinguish from a blatant market-sharing agreement. So, in my view at least, the main point of disagreement is not whether or not it is legitimate for competition authorities to be concerned about PFD agreements. Rather, the continuing disagreement between PDF “hawks” and “doves” stems from different views of the two theories of harm described above and of how these theories of harm can be applied to concrete cases. As I explain in more details below, the main question raised by the first theory of harm is why it should apply with special urgency to PFD deals. After all, any patent settlement effectively deprives society from the opportunity to invalidate patents that were granted in error. There are two main sources of disagreement about the second theory of harm. Firstly, as we will soon discuss, that theory is squarely based on a “probabilistic” view of patent protection. While undoubtedly appealing as a description of actual patent rights, the normative implications of the probabilistic view are far from being by all economists or legal scholars. Secondly, even if one were to subscribe to the probabilistic approach, it is not entirely clear what are the actual implications of this approach for patent settlements in general and PDF agreements in particular.

2.1. **First Theory of Harm: Are All Patent Settlements Objectionable?**

The review of patent applications by patent offices is necessarily imperfect, leaving a substantial portion of granted patents that are found invalid when further reviewed by the Courts. This situation does not necessarily reflect poorly on the performance of patent offices. As the patents that are challenged in Courts tend to be “those that matter”, it is actually optimal to save the cost of a true in depth examination on the vast majority of applications. In this sense, then, ensuring an effective review of patents by the Courts is important to the overall performance of our IPR systems. In spite of this however, it is widely believed that litigation settlements have a useful role to play as they provide faster and cheaper alternatives to legal disputes. Either one believes that this principle also applies to the special case of patent litigation or one does not. If one does not, then any patent litigation settlement gets in the ways of socially useful judicial review. Our first theory of harm would then logically imply that all patent settlements, not just PFD settlements, should be prohibited.

If, on the other hand, one believes that settlements have a role to play in patent litigation, then the question is how one would distinguish between “good” and “bad” settlement. Following the logic of the theory of harm, bad settlements should be those that involve patents that are likely to be overturned by the Court, i.e. “weaker” patents. The relevant question then is whether the presence of a transfer from the patent holder is a reliable indicator of the weakness of the patent involved in the litigation. The answer to that question is a qualified “no”. To see this, let us focus our attention on settlements that spell out a date of entry for the generic challenger as well as a possible transfer from the patent holder. Clearly the “overall package” offered to the generic must be more attractive if the patent is known to be weaker. This means that, if we compare two settlements with the same date of entry, one with a transfer and one without, one would generally believe that the settlement involving a transfer is associated with a weaker patent. On the other hand, we could not possibly draw any inference from the comparison of an agreement without payment and a given entry date and an agreement with payment that involves an earlier date of entry. So overall, when looking at a specific settlement, one simply cannot conclude that the presence of a payment implies that the patent at stake is weak. There no simple relationship between PFD and the strength of the underlying patent. Moreover, if one were to draw inferences from the combination of PFD and agreed upon entry date, one would face the following paradox: for a given size of transfer from the patent-holder to the generic entrant, a weak patent would lead to an earlier date of entry as the generic firm must be given a more attractive “package”. Clearly, a crackdown targeted at early entry agreements is not what Competition Authorities have in mind. Overall then, the first theory of harm does not seem to offer a sound basis for the singling out of PFD settlements.

2.2. **Second Theory of Harm #1: Probabilistic Patents**

It is important to clarify what economists mean when they refer to patents as “probabilistic” rights. There are essentially three “levels” of adherence to the probabilistic view.
**Level 1:** As a matter of positive analysis, the right to exclude granted by patents is without a doubt probabilistic as the patent-holder cannot be sure that the validity of the patent would be upheld if challenged and as, anyway, the precise coverage granted by the claims approved by the EPO remains quite uncertain until the construction of these claims has been further examined in Court. Finally, even if there was no inherent uncertainty in the IP right itself, Courts do make mistakes.

**Level 2:** As a matter of efficient design of a patent system, the probabilistic character of patent rights is actually desirable. As Ayres and Klemperer (1999) observe, making patent rights “more probabilistic” is similar to reducing what economists refer to as the “breadth” of the patent. From the work of Gilbert and Shapiro (1990), we know that, under rather general conditions, a patent design that trades-off breadth against length makes it possible to ensure a given reward to the innovator at least social cost. So, at this second, normative, level, the probabilistic aspect of IPRs is useful, as long as IP owners are properly compensated by adjustments to, less distortionary, dimensions of the patent right (such as length).

**Level 3:** The probabilistic right is all that the patent-holder is entitled to. The patent-holder cannot therefore take any action that would eliminate the probabilistic aspects of the right if this action is to the detriment of consumers. Concretely then, a patent settlement will be seen as anti-competitive if it leads to a level of consumer surplus that is lower than the surplus that consumers would have expected as a result of continuing litigation. Assume for example that consumers would get a surplus of 50 if the patent-holder prevailed and did therefore continue as a monopoly supplier until the end of the litigated patent but that this surplus would increase to 100 if the generic entrant prevailed (say if the patent was invalidated). If the ex ante probability of success of the patent-holder is $p$, then a settlement that leaves consumers with a surplus that is less than $50p + 100(1 – p)$ would be viewed as anti-competitive. This third, normative, view is what the second theory of harm relies on. While most economists have no problem with Levels 1 and 2, the same cannot be said about level 3. Among the potential objections are:

- **Consistency.** While they hold a valid patent, patent-owners who are not dominant are usually understood as facing no obligation to think about consumer welfare when acting within the scope of their patent. When entering a "normal" licensing agreement, for example, the terms of this agreement are properly set through bilateral negotiations without either party having to worry whether some alternative form of agreement would actually be better for consumers. Why then should such an obligation suddenly surface when an agreement – which might well involve licensing the technology – is reached as part of litigation against a potential generic entrant?

- **Practicality.** A rule that says that a patent-holder can use and defend her patent while remaining within the scope of this patent as long as the patent is currently valid is easy for economic agents to understand and easy to enforce. A world where patent-holders would have to evaluate every substantial action regarding the use of their IPRs by assessing the fundamental uncertainty of their rights would appear to lack the clarity and predictability for which competition law should strive. If patent-holders really need to ensure that they always leave consumers with at least as much surplus as would result from actually

---

4 On can take the view that the level 3 interpretation of probabilistic patents is indeed a consequence of level 2. It is because probabilistic patents might be part of an optimal patent system that it is important to hold companies to that standard. This of course assumes that the current patent system was indeed built in accordance with what economic theory tells us about the role of such a probabilistic dimension.
“drawing” the “lottery ticket” that is a patent, how will patent-holders know what will or will not be deemed to be anti-competitive?

• **Sources of Uncertainty.** There are further degrees of “purity” even within the adherents to the “third level” described above regarding what should be seen as forming part of the “legitimate” probabilistic nature of patents, i.e. those probabilistic aspects that the patent holder should see as given and inalterable. Should the possibility of judicial error be included into the probabilistic nature of patents or should patent holders have the right to protect themselves against such error? What about injunction risk? In pharmaceutical, failure to obtain an injunction can be catastrophic for the patent-holder: generic entry will lead to lower prices and it is practically impossible to restore the pre-entry price level later on even if the patent-holder ultimately prevails.

### 2.3. Second Theory of Harm # 2: What Type of Settlement Should be Prohibited Under a Probabilistic Patent Approach?

Let us assume for the sake of discussion that we agree with the probabilistic patent benchmark: consumers should get at least what they would get if patent litigation was not settled. How do we know whether a given settlement satisfies such a criterion? This is where the work of Carl Shapiro becomes crucial as it is designed to provide us with a simple criterion, thereby addressing the “practicability” issue raised above. In a nutshell – and without getting into the myriad of possible variations on the model – when a settlement involves an agreed date of independent generic entry, a transfer that exceeds the expected costs of litigation of the originator implies that the originator believes that the agreed upon date of entry is later than the expected date of entry if litigation proceeded to the end. So, if one accepts the probabilistic patent benchmark, patent settlements involving such transfers can only be anti-competitive and should therefore be forbidden

While useful, this criterion is not fool-proof even if one accepts the probabilistic patent benchmark. In particular it does not apply with such simplicity if the two parties have different exposure to risk or have different attitudes to risk. This later possibility should not be ruled out too easily, especially when the generic entrant is under severe financial constraints.⁵

Moreover, the Shapiro criterion simply does not apply to more complicated settlement agreements that do not simply involve an agreed date of independent generic entry with or without transfer. As just one example, there can be settlement agreements in which generic entry takes place immediately, but in which the generic must purchase from the patent holder at an agreed transfer price or must pay an agreed royalty. Even if such agreements are accompanied by value transfers that exceed the expected future costs of litigation, it does not follow that such agreements necessarily reduce the expected welfare of the consumers of the affected drugs relative to expected consumer welfare if the parties had litigated. Since leaving consumers with at least the surplus that they could expect from continued litigation is what separates acceptable settlements from anti-competitive settlements under the probabilistic patent

---

⁵ The possibility of bankruptcy tends to make firms behave in a more “risk loving manner”. 
benchmark, the presence of a value transfer as part of such agreements simply cannot be seen as sufficient evidence that the agreement is anti-competitive.

In fact, it is easy to show that, for any date of generic entry that would be expected from the continuation of litigation, there is an immediate entry agreement with wholesale supply from the originator (or a royalty payment to the originator) that makes all parties – including consumers – better off. Moreover, in order to provide the generic with an incentive to enter this kind of welfare-enhancing agreement, a “reverse payment” from the patent holder to the generic will be required.6 It is therefore hard to see how the simple presence of a payment from the patent-holder to the generic could be used as evidence that any settlement involving immediate entry should be seen as anti-competitive.

2.4. A POLICY CONCERN: INNOVATION

The theories of harm presented above take an ex post view: the innovation covered by the patent has already been obtained, so there is no discussion of how antitrust enforcement might affect incentives to innovate. This is an important drawback. The patent system is designed to foster innovation and ensures the diffusion of knowledge to the eventual benefit of consumers. An analysis that ignores effects on innovation therefore takes an incomplete view of the effects of value transfer settlements on welfare.

There are two main issues here. Firstly, would banning value-transfer settlements actually hurt the profits of patent-holders in the pharmaceutical sector? This is not clear a priori. On the one hand, value transfer settlements enable the patent-holder to settle litigation at a lesser cost, so removing this possibility would hurt. On the other hand, generics also benefit from value transfers so they might be less willing to challenge the patent in the first place if the practice was removed. That effect would be beneficial to patent-holders. If the net effect is of banning value transfers is favourable to patent-holders, then we are effectively in a situation where patent-holders' own ex post rational use of value transfers militate against their ex ante interest. Is this really ground for pursing them under antitrust law? If on the other hand, the net effect of banning value transfers is to decrease the patent holder’s expected profits, then one must consider the feedback effect on innovation.

Let us first dispose of a red-herring. Authors on both sides of the debate have invoked the fact that “the patent system has been designed to balance a variety of effects optimally”, so one should not unduly tinker with it through competition law. On the “anti-transfer” side, the argument is that the patent system is optimal in an environment where value-transfer settlements are not allowed….so there is no need to compensate innovators if value transfer settlements are banned. On the other side, the (implicit) view is that, since opposition to value-transfer settlements is recent, one must assume that the patent system balanced effects under the assumption that all kind of settlements within the scope of the patent would be allowed. In this view, banning value transfer settlements would therefore have potentially serious effect on the balance of the IP system in pharmaceuticals. Such debate over original intent is useless.

6 The analysis supporting this statement can be found in appendix 1
If one wants to study rigorously the effect of value-transfer settlements within the patent system, one must follow the usual approach used in the patent design literature evoked above: take the level of reward to innovators as given and determine whether allowing for value transfer settlements makes it more or less costly to consumers to provide this level of reward to innovators. If one conducts that analysis, one actually finds that a value transfer settlements make it possible to provide a given reward to consumers at a lesser cost in terms of ex post consumer welfare. In other words, even within a patent system designed on the basis of a probabilistic view of patents, there would still be room for allowing for settlements that involve payments from the patent-holder to the generic.

The second issue is whether banning value-transfer settlements would actually address the two theories of harm described at the beginning of this note. In particular, it is far from clear that a prohibition on value-transfers would actually lead to a more efficient “weeding out” of bad patents. Just as the impact of a prohibition on innovation incentives was ambiguous, one cannot conclude that it would lead to more challenges working their way to a final litigation outcome: a ban on value transfers might increase the proportion of generic challenges making it all the way through litigation but, since it decreases the expected pay-offs of the challenging generic, it could also lead to fewer challenges in the first place.

3. THE EUROPEAN APPROACH

DG Comp has been pursuing a small number of cases. If we look across these cases, we notice that the Commission is relying on both article 101 and article 102, sometimes within the same case. Under Article 101, the Commission sees value-transfer settlements as “per object” infringements. The Commission’s approach under Article 102 is less clear. Still, while the Commission does present arguments relating to the therapeutic substitutability of the medicine at stake with other medicines within the same class, there appears to be a new emphasis on a definition of dominance that is arguably “tailored to the alleged abuse”. Under this approach, the very fact that the patent-holder had the power to exclude generic competition through value-transfer “bribes” and that it had the incentives to do so – since prices tend to fall abruptly when generics enter – suffices to establish dominance, irrespective of the extent of therapeutic substitutability within class, the intensity of non-price competition or the overall profit margins realised on the protected drug. The Commission has reached a decision in the Lundbeck case, imposing rather large fines on both Lundbeck and the generic companies involved in the PFD agreements.

---

7 The formal analysis underlying this point can be found in appendix 2.
8 A formal analysis of this point is presented in appendix 3.
3.1. **Collateral Damage: Difficulties in Applying Articles 101 and 102**

As mentioned above, the European Commission has at times relied on both Article 101 and Article 102 to pursue PFD settlements. Each approach presents its own difficulties.

When using article 101, the main difficulty comes from determining whether generic companies can be seen as “potential competitors” as long as the patent at stake is valid. Traditionally, firms that are barred from entering a market because of the presence of a valid patent have not been seen as potential competitors in this market. Since a patent is presumed valid until it is voided by a competent authority, a generic challenger who settles with the patent-owner cannot then be seen as a potential entrant since, if there is indeed infringement, at the time of the agreement entry could only occurs in violation of the patent. In a sense, then, the traditional view of a patent as being either “on” or “out” is mirrored by a dichotomous assessment of potential competition as “on” if entry does not violate a “on” patent and “off” if it does. This suggests that the pursuit of PFD cases under article 101 requires a redefinition of the notion of potential entry to fit the probabilistic theory of harm: if patents are thought to be probabilistic, then it would also make sense to consider potential entry as a probabilistic concept. In that view, a generic entrant would still be seen as a potential competitor if there is a sufficiently high probability that it would actually prevail in litigation and therefore be able to enter the market. An interesting implication of this view would be that article 101 could only be applied if there was sufficient evidence that the patent at stake is weak. However, the European Commission has carefully avoided any reference to the strength of the patents involved in PFD deals and has certainly not presented any evidence suggesting that those patents were weak. In my opinion, this is inconsistent with the need to redefine the notion of potential entrant in a probabilistic manner that fits the probabilistic nature of the Commission’s theory of harm.

The use of Article 102 raises two main issues. The first one is the traditional unease that some observers feel when abuse of dominance is used to get at an agreement between willing parties. The second relates to the manner in which dominance is established. We do not need to discuss here the general issue of how one assesses market power and dominance in “high sunk cost” industries such as pharmaceuticals, since this is not specific to value transfer settlements. However, the Commission’s approach to dominance seems, as mentioned above, to be tailored to the specific abuse that it pursues. In a nutshell, the Commission considers that the fact that generic production of a given medication leads to a collapse in the price of the drug, while generic entry into drugs that are good therapeutic substitutes does not, is evidence that, for the type of abuse considered, each drug is a market onto itself, regardless of how many close therapeutic substitutes are available. This seems to boil down to saying that any patent-holder who is the sole supplier of a drug that sells for a price that is substantially higher than its variable cost of production will be found to be dominant. In practice, such an approach implies that, in the context of alleged abuses regarding generic entry, the vast majority of existing drugs confer dominance on the relevant patent-holder.

The Commission’s approach raises two main issues. Firstly, should dominance be assessed solely in terms of price behaviour? Given that, in most health systems, doctors and patients have little incentive to consider the price of the medicines that they prescribe or use, the fact that a decrease in the price of a drug has little effect on the price or sales of another drug that is a close therapeutic substitute is hardly surprising. However, under European Law, dominance is defined broadly as the ability to behave to a substantial extent independently from other firms and consumers. This definition seems to imply that all forms of competition should matter. In particular there is intense rivalry between therapeutic substitutes in
terms of “share of voice” (i.e., medical profession advocacy), experimental studies and research. Disregarding these dimensions of competition to narrowly focus on a price rivalry that is inhibited by the rules of the health system seems hard to justify.

Secondly, finding a drug dominant whenever generic entry would lead to a substantial decrease in price amounts to evaluating the market power of the drug compared to a competitive benchmark where prices are equal to marginal (or at least variable) costs. This makes no sense in an industry with high sunk costs. In such industries, a much more natural competitive benchmark is the price at which the drug manufacturer breaks even over the lifetime of the product. Such a benchmark at least ensures that a firm producing a drug that does not even recover its initial R&D expenditure is not found “dominant” and hence endowed with the special responsibility that comes with dominance. Establishing dominance based on ex post marginal cost pricing makes a mockery of the whole intellectual property bargain whereby patent protection aims at rewarding inventors and hence at insuring that, on average at least, they can at a minimum recover their initial investment.

3.2. A BROADER POLICY VIEW

It is also interesting to ask what the likely impact of banning value-transfer settlements might be. We have already discussed what the potential effect on innovation might be. We now turn to the likely effect on settlements and a potential effect on the behaviour of generics companies.

After completing its review of the pharmaceutical sector in 2008, in which it indicated that it saw value transfer settlements as potentially problematic, the Commission decided to keep track of pharmaceutical settlements, classifying them in three categories (see table). The Commission concluded, with some satisfaction, that the proportion of settlements that imposed no restrictions on generic entry had increased. Moreover the proportion of cases that limited generic entry but without transfer payment increased compared to cases where entry was restricted and transfer payments were made.

The Commission concluded that its negative stance on value transfer settlements had not made it more difficult for firms to settle and had been effective in reducing the occurrence of the objectionable kind of settlement. The first point is of course not correct: the fact that there is still a large number of settlements tell us nothing without any information on the population of actual and potential litigation cases that these numbers refer to. As for the second conclusion, there is a bit of a sleight of hand. First, cases of settlements without limit on generic entry are almost all cases where the patent-holder had already essentially lost the case because of adverse preliminary rulings. Second – and most interestingly – every single case in the category involves a settlement where the generic entry was delayed until the end of the period of patent protection. In other words, over four years and more than 400 settlements, there was not a single example of the type of settlement where generic entry is allowed at a date that is supposed to reflect the parties' appraisal of their respective chances at trial.
### Table 1: Pharmaceutical Patent Settlements after the EU’s Pharma Review

<table>
<thead>
<tr>
<th>Total settlements</th>
<th>Settlements, by type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No limit on generic entry (A type)</td>
</tr>
<tr>
<td></td>
<td>Generic entry restricted, but no value transfer from originator to generic (B.I type)</td>
</tr>
<tr>
<td></td>
<td>Generic entry restricted, value transfer from originator to generic (B.II type)</td>
</tr>
<tr>
<td>207</td>
<td>104 (50%)</td>
</tr>
<tr>
<td></td>
<td>54 (26%)</td>
</tr>
<tr>
<td></td>
<td>46 (22%)</td>
</tr>
<tr>
<td>93</td>
<td>53 (57%)</td>
</tr>
<tr>
<td></td>
<td>31* (33%)</td>
</tr>
<tr>
<td></td>
<td>9 (10%)</td>
</tr>
<tr>
<td>89</td>
<td>54 (61%)</td>
</tr>
<tr>
<td></td>
<td>32* (36%)</td>
</tr>
<tr>
<td></td>
<td>3 (3%)</td>
</tr>
<tr>
<td>120</td>
<td>84 (70%)</td>
</tr>
<tr>
<td></td>
<td>23* (19%)</td>
</tr>
<tr>
<td></td>
<td>13 (11%)</td>
</tr>
</tbody>
</table>

This raises two questions. Firstly, where are the welfare benefits from the Commission’s expressed negative view of value transfer settlements then? Have private parties just gotten cleverer without any benefit for consumers? Secondly, does this raise question about how appropriate the counterfactual used in the “Shapiro” theory of harm? The counterfactual where parties agree on an entry date that reflects the strength of the patent is a fine theoretical benchmark, but is it a useful policy benchmark if it is never observed?

### 4. THE US SUPREME COURT ACTAVIS DECISION

One of the very few things that are clear in the recent Supreme Court Decision is that the Court could not follow an extreme “scope of the patent” approach that would have made impossible to guard against disguised “market sharing agreement. However, the Court failed to identify clear reasons why PFD might be considered to be anti-competitive. For example, while the Court recognises that patents need to be further tested in Court, it fails to explain why this might imply that some patent settlements are lawful while others are not. It seems also clear that the Supreme Court did not embrace the extreme “probabilistic patent” approach. While the Court clearly sees patents as probabilistic in our “level 1” sense and points out that reverse payments are puzzling, there are no references to the probabilistic benchmark (our “level three”) according to which consumers should get at least the level of welfare that they would expect from
continued litigation. Overall, then, the Court appears to have opted – not surprisingly – for a very traditional approach: the enforcement of IPRs, including settlements, is not a matter for patent law only (even if it allows for antitrust considerations) but is potentially fair game for antitrust authorities. It does not mean that enforcement will be simple however as the Court acknowledges that agreements that include reverse payments cannot be seen as “presumptively unlawful”. Interestingly, the Court seems to recognise the importance of patent strength in establishing whether or not an agreement is lawful under a rule of reason, going as far as pointing out that the relationship between the size of the payment and the implied strength of the patent would be one of several pieces of information that would make a detailed analysis of patent validity unnecessary. Clearly then the Court did not condone either an impervious “scope of the patent” approach or an extreme “probabilistic” view. In this sense, the decision is compatible with our previous discussion as neither of these views is a realistic basis for effective policy. The first one ignores the real concerns that patent settlement agreements can support market-sharing deals and the second one is both too extreme in its logical implications for other aspects of licensing behaviour and would be basically impossible to implement in all but the simplest cases. One could however interpret the Court’s current position in light of the two theories of harm that we have discussed. In terms of the first theory, the Court sees the need to ensure the review of patents, especially when these are likely to be weak. In terms of the second theory of harm, the Court seems to be most concerned by the fact that PFD settlements could be used to preserve the unjustified monopoly rents of a weak patent. The common denominator of the Court’s concerns is that they arise mostly when there are reasons to believe that the patents under litigation would be likely to be overturned if litigation proceeded to the dire end. In that sense, the Actavis decision seems to be at odds with the approach of regulators like the European Commission who have gone to great length to keep the notion of “patent strength” out of their line of argument.

Unfortunately, the Actavis decision does not tell us much more than that. The decision is particularly obscure in terms of burden of proof. While most of the language suggests that the burden of proof is essentially on the FTC, the Court also seems to leave the door open to a claim that a very large reverse payment would itself be presumptive evidence that would then need to be refuted by the defendant. Furthermore, if one were to actually read the decision as establishing a “rule of reason” approach, there is very little guidance as to the type of evidence that “reason” should look at. The decision is particularly ambiguous as to the role of patent “strength”. Is the demonstration that the patent could reasonably have been seen as strong at the time of the agreement always a legitimate defence or is it trumped anyway if the transfer from the patent-holder to the generic is judged to be “unreasonably” large anyway?

Overall then, even though the reasons for disagreement between value transfer “hawks” and value transfer “doves” are by now fairly clear, I fear that the Actavis decision has done little to bring about a quick resolution of the PFD debate.
5. CONCLUSION

Competition authorities have relied on two main theories of harm to pursue PFD settlements. The first theory states that such settlements unduly prevent the patents at stake from being properly re-evaluated by a Court. The main weakness of this theory of harm is that it fails to explain why PDF settlements should be seen as less desirable as any other type of patent settlement. The second theory of harm relies on the view of patents as probabilistic property rights. This theory of harm has two anchors. The first one is the claim that the holder of a probabilistic right should ensure that consumers enjoy a level of welfare that is at least as high as the level that they would expect from the completion of patent litigation. While this theory of harm is worth taking seriously, it has a number of weaknesses, including inconsistencies between the probabilistic view and traditional antitrust treatment of licensing, the fact that it cannot be applied to more complex agreements where payments are accompanied by immediate generic entry and the identification of the sources of uncertainty that are properly reflected in the “probabilistic” nature of the patent rights. The overall effect of policies banning PFD settlements on innovation is also a concern.

Turning to the on-going investigation of PFD agreements in Europe, I briefly discuss three sources of controversy. The first issue is what the proper definition of a “potential entrant” should become when one considers the patent rights themselves to be probabilistic. I argue that he logical approach would be to adopt a probabilistic definition of potential entry itself but that this also implies that only settlement of litigation relating to patents thought to be weak should be a concern. The second issue is the approach currently taken to determine dominance in PFD cases. I argue that this approach not only relies on the wrong competitive benchmark but it simply ignores the strong competitive constraints that therapeutic substitutes exercise on each other through non-price channels. Finally, tracking the evolution of settlements since the European Commission’s review of the pharmaceutical sector – where doubts about the legality of PFD settlements were first expressed – shows that the kind of settlement where firms agree on a date of generic entry without side payments actually do not arise. Since this type of settlement is the benchmark compared to which PDF agreements are thought to be abusive, this raises questions about the very foundations of the Commission’s theory of harm.

Finally we argue that the recent Actavis decision does not support either a pure “scope of the patent” approach or a pure “probabilistic patent” approach. Rather it seems to attempt to strike for a middle ground where the strength of the patents at stake would be an important element of the competitive appraisal of PFD settlements.
REFERENCES


K. Rockett, 2010, "Property Rights and Invention", Chapter 7 in B. Hall and N. Rosenberg (Eds.), Handbook of the Economics of Innovation (Vo.2), Elsevier.


Appendix One: Agreements with Immediate Generic entry

The demand function on the downstream market where the patent-holder and the generic entrant would compete is given by

\[ P = 1 - Q = 1 - (Q_1 + Q_2) \]

Assuming constant per unit costs equal to \( c \), the corresponding monopoly output is

\[ Q^M = \frac{1 - c}{2} \]

And monopoly profits are

\[ \pi^M = \left( \frac{1 - c}{2} \right)^2 = \frac{(1 - c)^2}{4} \]

While consumer surplus is equal to

\[ CS^M = \frac{1}{2} \left( \frac{1 - c}{2} \right)^2 = \frac{(1 - c)^2}{8} \]

We now assume that both the patent holder (firm 1) and the generic entrant (firm 2) compete in the downstream market under a licensing agreement when firm 2 pays a royalty of \( r \) per unit of output to firm 1. The maximisation problem of the patent holder is:

\[ \text{Max}_{Q_1} Q_1 [1 - Q_1 - Q_2 - c] + rQ_2 \]

And the maximisation problem of the generic licenee is:

\[ \text{Max}_{Q_2} [1 - Q_1 - Q_2 - c - r]Q_2 \]

The corresponding first order conditions are:

\[ 1 - 2Q_1 - Q_2 - c = 0 \]
Adding up these two conditions gives us

\[ 1 - Q_1 - 2Q_2 - c - r = 0 \]

So that the total equilibrium output sold in the market is

\[ Q_1^* + Q_2^* = \frac{2 - 2c - r}{3} \]

While individual equilibrium outputs are given by

\[ Q_1^* = 1 - (Q_1^* + Q_2^*) - c = 1 - \frac{2 - 2c - r}{3} - c = \frac{1 - c + r}{3} \]

\[ Q_2^* = 1 - (Q_1^* + Q_2^*) - c - r = 1 - \frac{2 - 2c - r}{3} - c - r = \frac{1 - c - 2r}{3} \]

The resulting equilibrium profits and consumer surplus are

\[ \pi_1^* = \left( \frac{1 - c + r}{3} \right)^2 + \frac{1 - c - 2r}{3} \left( \frac{1 - c + r}{3} \right) + \frac{3r(1 - c - 2r)}{9} \]

\[ \pi_2^* = \left( \frac{1 - c - 2r}{3} \right)^2 = \frac{(1 - c - 2r)^2}{9} \]

\[ CS^* = \frac{1}{2} (Q_1^* + Q_2^*)^2 = \frac{(2c - r)^2}{18} \]

We can now determine the equilibrium lifetime profits and consumer surplus where \( K \) periods of monopoly are followed by entry without paying any per unit wholesale price or royalty to firm 1. \( T \) represents the remaining duration of patent validity.

\[ \pi^{tot}(K) = K \left( \frac{1 - c}{4} \right)^2 + (T - K) \left( \frac{2(1 - c)^2}{9} \right) \]
In the special case of immediate entry \((K = 0)\), we get:

\[
CS^{Tot}(K) = K \frac{(1 - c)^2}{8} + (T - K) \frac{2(1 - c)^2}{9}
\]

Comparing consumer surplus with immediate entry to consumer surplus we delayed entry, we get

\[
\pi^{Tot}(r) = T \frac{(1 - c + r)^2 + (1 - c - 2r)^2 + 3r(1 - c - 2r)}{9}
\]

\[
= T \frac{((1 - c) + r)^2 + ((1 - c) - 2r)^2 + 3r((1 - c) - 2r)}{9} = T \frac{2(1 - c)^2 + r(1 - c) - r^2}{9}
\]

\[
CS^{Tot}(r) = T \frac{(2 - 2c - r)^2}{18}
\]

Using the quadratic formula, we get the roots of the following second degree inequality

\[
T \frac{(2 - 2c - r)^2}{18} \geq K \frac{(1 - c)^2}{8} + (T - K) \frac{2(1 - c)^2}{9} \leftrightarrow
\]

\[
Tr[r - 4(1 - c)] \geq \frac{7}{4} K(1 - c)^2
\]

Using the quadratic formula, we get the roots of the following second degree inequality

\[
Tr^2 - 4Tr(1 - c) + \frac{7}{4} K(1 - c)^2 \geq 0
\]

as

\[
\tilde{r}_{CS} = \frac{4T(1 - c) - \sqrt{16T^2(1 - c)^2 - 4T \left( \frac{T}{4} K(1 - c)^2 \right)}}{2T} = \frac{2T(1 - c) - \sqrt{4T^2(1 - c)^2 - \frac{7KT(1 - c)^2}{4}}}{T}
\]
Thus $CS(r) - CS(K) \geq 0$ iff $r \leq (1 - c) \left[2 - \sqrt{4 - \frac{7K}{4T}}\right]$ or $r \geq (1 - c) \left[2 + \sqrt{4 - \frac{7K}{4T}}\right]$

$r_{cs}^+ = (1 - c) \left[2 + \sqrt{4 - \frac{7K}{4T}}\right]$ 

Thus $CS(r) - CS(K) \geq 0$ iff $r \leq (1 - c) \left[2 - \sqrt{4 - \frac{7K}{4T}}\right]$ or $r \geq (1 - c) \left[2 + \sqrt{4 - \frac{7K}{4T}}\right]$

$r_{cs,K=0}^- = (1 - c) \left[2 - \sqrt{4 - \frac{0}{4T}}\right] = 0 < \frac{(1 - c)}{2}$

$r_{cs,K=T}^- = (1 - c) \left[2 - \sqrt{4 - \frac{7T}{4T}}\right] = (1 - c) \left[2 - \sqrt{\frac{16 - 7}{4}}\right] = (2 - \frac{3}{2})(1 - c) = \frac{(1 - c)}{2}$

$r_{cs,K=0}^+ = (1 - c) \left[2 + \sqrt{4 - \frac{0}{4T}}\right] = (1 - c)[2 + 2] = 4(1 - c) > \frac{(1 - c)}{2}$

$r_{cs,K=T}^+ = (1 - c) \left[2 + \sqrt{4 - \frac{7T}{4T}}\right] = (1 - c) \left[2 + \sqrt{\frac{16 - 7}{4}}\right] = (2 + \frac{3}{2})(1 - c) > \frac{(1 - c)}{2}$

Therefore, for any $K \in [0, T)$ there will be $r \in [0, \frac{1-c}{2})$ such that $CS(r) - CS(K) \geq 0$.

We know compare profits with immediate entry to profits with delayed entry.

$$\pi(r) \geq \pi(K) \leftrightarrow$$
Using the quadratic formula, we get

\[
\frac{2(1-c)^2 + r(1-c) - r^2}{9} \geq \frac{K(1-c)^2}{4} + (T-K) \frac{2(1-c)^2}{9}
\]

\leftrightarrow r^2 - (1-c)r + \frac{K}{4T} (1-c)^2 \leq 0

Thus \( r_\pi^- = \frac{(1-c) - \sqrt{(1-c)^2 - \frac{K}{T} (1-c)^2}}{2} = \frac{(1-c)(1 - \sqrt{1 - \frac{K}{T}})}{2} \)


\[
\frac{1-c) + \sqrt{(1-c)^2 - \frac{K}{T} (1-c)^2}}{2} = \frac{(1-c)(1 + \sqrt{1 - \frac{K}{T}})}{2}
\]

Thus \( \pi(r) - \pi(K) \leq 0 \) iff \( r \geq \frac{(1-c)(1 - \sqrt{1 - \frac{K}{T}})}{2} \) or \( r \leq \frac{(1-c)(1 + \sqrt{1 - \frac{K}{T}})}{2} \)

\[
r_{\pi K=0}^- = \frac{(1-c)(1 - \sqrt{1 - \frac{K}{T}})}{2} = 0 < \frac{(1-c)}{2}
\]

\[
r_{\pi K=T}^- = \frac{(1-c)(1 - \sqrt{1 - \frac{T}{T}})}{2} = \frac{(1-c)}{2}
\]

\[
r_{\pi K=0}^+ = \frac{(1-c)(1 + \sqrt{1 - \frac{0}{T}})}{2} = (1-c) > \frac{(1-c)}{2}
\]

\[
r_{\pi K=T}^+ = \frac{(1-c)(1 + \sqrt{1 - \frac{T}{T}})}{2} = \frac{(1-c)}{2}
\]

For any \( K \in [0, T] \) there will be \( r \in [0, \frac{1-c}{2}] \) such that \( \pi(r) - \pi(K) \geq 0 \).
We now check that the overlap between the parameter range for which there is an immediate entry agreement that increases consumer surplus and the parameter range for which (joint) profits are also higher in such an agreement is not empty. For there to be an \( r \) for which both consumer surplus and profits to be higher with immediate settlement than with royalty free settlement we need

\[ r_{\pi}^- \leq r_{cs}^- \]

\[
\frac{(1 - c)(1 - \sqrt{1 - \frac{K}{T}})}{2} \leq (1 - c) \left[ 2 - \sqrt{4 - \frac{7K}{4T}} \right] \leftrightarrow \\
-\sqrt{1 - \frac{K}{T}} \leq 3 - 2 \sqrt{4 - \frac{7K}{4T}} \leftrightarrow \\
6 \left( \sqrt{1 - \frac{K}{T}} - \left( 1 - \frac{K}{T} \right) \right) \geq 0
\]

The above condition is satisfied for all \( K \in [0, T] \) as \( \sqrt{x} \geq x \) when \( 0 \leq x \leq 1 \).

Finally, we check the sign of the transfer required to make both firms better off with the immediate entry agreement.

Let Firm 1 pay Firm 2 a (positive or negative) side payment of \( b \frac{(1-c)^2}{9} \) (\( k \) times the per period per firm duopoly profit). This specification captures no transfer case when \( b = 0 \) and transfer from firm 2 to Firm 1 when \( b < 0 \).

\[
\pi_2(r) + b \frac{(1-c)^2}{9} \geq \pi_2(K) \leftrightarrow \\
T \frac{(1-c-2r)^2}{9} + b \frac{(1-c)^2}{9} \geq (T-K) \frac{(1-c)^2}{9} \leftrightarrow \\
4Tr^2 - 4T(1-c)r + (K + b)(1-c)^2 \geq 0
\]

The two roots are
We need \( r > r_1^- \) for industry profits to be higher with royalties and we need \( r < r_2^- \) for Firm 2’s profits to be higher with royalties. When

- \( b < 0 \) we have \( r_2^- > r_1^- \)
- \( b = 0 \) we have \( r_2^- = r_1^- \)
- \( b > 0 \) we have \( r_2^- < r_1^- \)

Thus we conclude that the immediate entry settlement which leads to higher CS and aggregate profits can only be achieved by a side payment from Firm 1 to Firm 2.

**Appendix 2: Could PFD Settlements be Part of an Optimal Patent Protection Design?**

We focus on the case where prohibiting settlements that involve transfers from the patent holder to the generic entrant would reduce the expected profits from the patent-holder, i.e. where the negative effect of prohibiting value transfer settlements on the profits of patent holders (conditional on a challenge being filed) is not outweighed by the negative effect that prohibiting value transfer settlements might have on the number of generic challenges (an effect that would be a benefit to patent holders). Under these conditions, prohibiting value transfer settlements reduces the expected profits of patent holders and therefore weakens incentives to innovate.

In order to restore incentives to innovate to their pre-ban level, policy makers must respond with a compensating adjustment elsewhere in the “innovation system”. We identify possible ways in which a
compensating adjustment might be implemented. We then consider whether consumer welfare would be higher in a world where value transfer settlements were allowed or in a world where they were prohibited and R&D incentives were preserved through these alternative policy measures.

In considering what kind of compensating adjustment could be adopted, pharmaceutical prices are a natural candidate. Even when it is protected by patent, a pharmaceutical company is usually not free to set its price as it wishes. Typically, it is constrained by a regulated price imposed by the local health authority or negotiated with local health insurers. The fact that prices are regulated directly or indirectly means that policy makers can use price as an instrument for increasing patent holder profits, if they so desire.

To assess the efficiency of an innovation system in which PFD settlements are permitted relative to one where such payments are prohibited but patent holders are allowed higher prices to preserve R&D incentives, note first that prohibiting value transfer settlements will lead to earlier entry by generics, conditional on the generic challenge being made. This is because earlier entry must be offered to (partially) compensate the generic challenger for the lack of transfer payment.

This observation about the effect of prohibiting value transfer settlements on market structure implies that we should be making the following comparison when considering whether consumer welfare would be higher in a world where value transfer settlements were allowed or in a world where they were prohibited and R&D incentives were preserved through the alternative policy measures, e.g. allowing higher prices during the period prior to generic entry:

1. A world where value transfer settlements are allowed, the patent holder can be expected to remain the sole provider for a proportion $\alpha$ of the remaining life of its patent, and the patent holder is allowed a regulated price $P_r$ (where $P_r$ is a price which, if collected over the life of the patent, will allow the patent holder to collect what health regulators consider risk-adjusted normal profits);

2. A world where value transfer settlements are prohibited, the patent holder is the sole provider for a proportion $\beta < \alpha$ of the remaining life of its patent, and (in order to maintain patent holder profits and hence innovation incentives) the health regulator allows the patent holder to collect a price of $P_r + X$ during the period prior to generic entry (when regulation rather than competition determines price).

Both scenarios (by construction) provide the patent holder with the same profits. Can we say anything about whether one scenario is better than the other from the perspective of consumer welfare in the market directly affected? The answer to this question comes from the economic literature on optimal patent design. There are different instruments that can be used within a patent system to affect returns to innovation. One instrument is the length of patent protection. Another instrument is “patent breadth”. As used by economists, patent breadth refers to anything that makes it harder for rivals to introduce a competing product or technology without infringing the patent. In practice, this boils down to the number, type and precise wording of the claims that are allowed by the patent examiner. If patents are interpreted broadly, this can provide the patent holder with additional protection from substitute products, thereby allowing the patent holder to charge higher prices during the period of patent protection. From the

---

9 See Regibeau and Rockett (2007) and Rockett (2010) for a review.
perspective of a patent holder, patent length and breadth are substitute instruments. A patent holder would be indifferent between long narrow patents and short broad patents, provided the profits to the patent holder are kept the same.

One of the main findings in the literature on optimal patent design is that, even if patent holders might be indifferent between long narrow and short broad patents, consumers should not be. This literature shows that, under some broad conditions, increasing the reward obtained by patent holders by a given amount by increasing patent breadth becomes increasingly costly in terms of consumer welfare. By contrast, increasing the reward to patent holders by increasing the length of protection increases profits and consumer losses in the same proportion. This means that, all else equal, it is more efficient to ensure a given reward to innovators by granting long patents with rather narrow cover than by granting broad patents for a shorter period of time.10

There is a clear relationship between this finding in the literature on optimal patent design and the question posed above – i.e. whether consumer welfare would be higher in a world where value transfer settlements were allowed or in a world where they were prohibited and R&D incentives were preserved through the alternative policy measures such as allowing higher prices during the period prior to generic entry. A world where value transfer settlements are allowed (but prices are regulated to allow no more than “normal” profits) is analogous to a world with long narrow patents. By contrast, a world where value transfer settlements are not allowed, generic entry takes place sooner and regulated prices are increased during the period prior to entry to preserve patent holder profits is analogous to a world with short broad patents. The literature on optimal patent design therefore implies that, between these two alternatives, the preferred solution from the perspective of consumer welfare is the scenario analogous to long narrow patents – i.e. a world where value transfer settlements are allowed.

Appendix 3: Can a Reduction in the Private Benefits of Settlements Lead to Fewer Patents Being Reviewed by the Courts?

Reducing the private gains from settlement has two opposing effects on the number of patents that are ultimately tested in court. By reducing the size of the “settlement pie”, it will increase the percentage of filed cases that reach judicial decision (rather than settling). However, by reducing the value to a generic challenger of the “option to settle”, it may also reduce the number of patent challenges that are filed. The net effect on the number of patents that get tested in court may therefore be positive or negative.

This appendix presents a simple economic model that illustrates these conflicting effects and which confirms that the net effect of prohibiting value transfer settlements on the number of patents that get tested in court can be positive or negative. For ease of exposition, the model is based on various simplifying assumptions. These assumptions can be dropped without affecting the qualitative nature of the results.

**Probability assessments**

The patent holder is certain that his patent is valid and remains certain to the end.\(^{11}\) The potential challenger, by contrast, believes that the validity of the patent is uncertain. Ex ante, the potential challenger believes that the patent is invalid with a probability \(p\). This probability \(p\) is drawn randomly from a uniform distribution over \([0,1]\).

There is a cost \(C\) of starting the challenge. Once the challenge has been started, the challenger gets more information about the patent. This additional information may or may not lead to a revision in the challenger’s assessment of the probability that the patent is invalid. More precisely, there is a probability \((1 – q)\) that the challenger will receive a new signal \(p’\) (where \(p’\sim U[0,1]\) and is independent of \(p\)). Hence, the ex post belief of the challenger as to the probability of success at trial is its prior \(p\) with probability \(q\), and a new independent draw \(p’\) with probability \((1 – q)\).

Litigation costs are \(L\) for either party. The discrepancy between the entrant’s beliefs and the beliefs of the patent holders will be the key mechanism that prevents the parties from always settling, so that some cases actually make it to trial. In other words, this model belongs to the “divergent expectations” tradition in the literature on the economics of litigation settlement as opposed to models where settlements arise because of asymmetric information. Interestingly, Waldfogel (1998) finds that the predictions of “divergent expectations” models are better aligned with empirical evidence that those based on asymmetric information models.\(^{12}\)

**Settle or litigate?**

Let us assume that the challenger gets a payoff of zero if it loses (excluding litigation costs) and gets \(W\) if it wins. What we want to know is whether improving the settlement pay-off of the challenger would in fact decrease the number of cases going to a judicial decision. In order to answer this question as simply as possible, rather than modelling how settlement values are determined (e.g. by assuming a particular bargaining model), we “blackbox” the settlement process and assume that the challenger gets an expected pay-off of \(X\) from settling. Implicit in this assumption is that \(X\) is not so high that the patent holder would refuse to settle.

If the challenger decides to pursue the litigation to its conclusion, then its expected pay-off is:

\[
zW - L
\]

where \(z\) equals the challenger’s belief of its probability of success at trial (i.e. ex ante, \(z=p\), and ex post, \(z=p\) with probability \(q\) and \(z=p’\) with probability \((1-q)\)).

The challenger would therefore settle if and only if:

\[
zW - L \leq X
\]

or equivalently:

\[
zW' - L' \leq X'
\]

---

\(^{11}\) This is a simplifying assumption to avoid updating the probability assessments of the patent holder.

\(^{12}\) Waldfogel, J., 1998, Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation, NBER Working Paper 6409. Hay and Spier (1997), supra note Error! Bookmark not defined., also conclude that “[t]he existence of divergent party expectations concerning trial remains the most influential account of why cases may fail to settle” (page 5).
The decision to challenge

The patent will be challenged only if the expected payoff from challenging, taking into account the costs of launching the challenge (C), is greater than zero. In order to determine whether this is the case, we must work backwards, and consider the payoffs once the challenge has been launched.

This requires working through the different possible scenarios. There are two ranges of parameters to examine – where \( p \leq \frac{X+L}{W} \) and the challenger initially expects to settle; and where \( p > \frac{X+L}{W} \) and the challenger initially expects to litigate.

Case 1: \( p \leq \frac{X+L}{W} \) (“opportunistic challenge”)

In this case, the challenger would settle unless it receives information that causes it to reassess the probability of invalidity and this revised probability of invalidity is sufficiently high.

Working backwards, if the challenger chooses to challenge the patent:

With probability \( q \), the challenger does not receive a new signal. Thus its belief of probability of success at trial remains at \( p \); there is settlement and the challenger gets \( X \).

With probability \( 1 - q \), the challenger receives a new signal \( p' \). Then:

- If \( p' \leq \frac{X+L}{W} \) the challenger still settles and still gets \( X \); this occurs with probability \( \frac{X+L}{W} \) since \( p' \sim U[0,1] \);
- If \( p' > \frac{X+L}{W} \), the litigation continues; this occurs with probability \( 1 - \frac{X+L}{W} \). The challenger’s expected payoff can be computed as \( \frac{W + X + L}{2} - L \).

Denote \( \Pi_L \) the expected pay-off excluding litigation costs. (The subscript “L” denotes “low” values of \( p \), i.e. \( p \leq \frac{X+L}{W} \), as opposed to the “high” case below.) So, the expected pay-off considered by the challenger when deciding whether or not to launch a challenge is:

\[
\Pi_L = qX + (1-q) \left\{ \frac{X+L}{W} \cdot X + \left( 1 - \frac{X+L}{W} \right) \left( \frac{W+X+L}{2} - L \right) \right\}
\]

which can be written as

\[
\Pi_L = qX + \left( \frac{1-q}{W} \right) \left\{ X(X+L) + \frac{1}{2} [W^2 - (X+L)^2 - L(W-X-L)] \right\}
\]

The challenge is launched if \( \Pi_L > C \).

Notice that \( \Pi_L \) is independent of \( p \). This means that there is a minimum value of \( X \), \( X_{min} \), such that \( \Pi_L \) is greater than \( C \) (the costs of launching a challenge). Thus, all patents where the challenger’s initial assessment of the probability of invalidity (\( p \)) would imply a settlement if a challenge were launched will in fact be challenged if and only if \( X \geq X_{min} \).

\[13\]

Since \( E(p') > (X+L)/W \) and the challenger’s expected payoff is \( E(p'W - L | p' > (X+L)/W) \).
This result implies that a decrease in $X$ that causes it to fall below the $X_{\text{min}}$ threshold discretely reduces the fraction of patents that get challenged from 1 to 0 for patents where the challenger’s initial assessment of the probability of invalidity ($p$) would imply a settlement if a challenge were launched.

Intuitively, this result is obtained because – for $p \leq (X + L)/W$ – the challenge is launched with the ex ante expectation that it will be settled; and even if it does reach litigation, $p$ does not affect the expected litigation payoff, since $p'$ is independent of $p$.

Although the challenger’s ex ante expectation is to settle, some cases are in fact litigated, since with probability $(1-q)$ a new signal $p'$ is received, and if $p'$ is sufficiently high the challenger will prefer to litigate rather than settle.

Since a proportion of these cases will be litigated, a reduction in the settlement payoff of the challenger below $X_{\text{min}}$ would unambiguously reduce the number of patents that are challenged and hence will reduce the number of patents that are the object of a judicial decision. Decreases in $X$ that do not cause $X$ to fall below $X_{\text{min}}$ would, on the other hand, increase the absolute number of cases actually scrutinised by the courts.

One might be suspicious that this result might be wholly dependent on the “discrete” information structure that we have assumed. To dispel such doubt, let us then look at the second range of values of $p$.

**Case 2: $p > (X + L)/W$ (“intent to litigate”)**

In this case, a firm that decides to launch a challenge does so with the initial expectation that it will not settle and will instead litigate to a judicial conclusion. The challenger will stick to litigation unless it relies on a new signal and this signal is sufficiently low. That is, in the case that the challenger chooses to challenge the patent, it does so in the expectation that it will litigate rather than settle; it will choose not to litigate only if $p'$ is sufficiently low (i.e. if $p' \leq (X + L)/W$).

Working backwards, if the challenger chooses to challenge the patent:

*With probability $q$, the challenger does not receive a new signal. Thus its belief of probability of success at trial remains at $p$; there is litigation and the challenger’s expected payoff is $pW - L$.\n
With probability $(1-q)$, the challenger receives a new signal $p'$. Then:

- if $p' \leq (X + L)/W$ the challenger settles and gets $X$; this occurs with probability $(X + L)/W$ since $p' \sim U[0,1]$;
- if $p' > (X + L)/W$, the litigation continues; this occurs with probability $1 - (X + L)/W$. The challenger’s expected payoff can be computed as $(W + X + L)/2 - L$ (see footnote 13 above).

Therefore the expected pay-off of the challenger when deciding whether to challenge the patent is:

$$
\Pi_H = q(pW - L) + (1 - q) \left\{ \left( \frac{X + L}{W} \right) \cdot X + \left( \frac{1 - X + L}{W} \right) \left( \frac{W + X + L}{2} - L \right) \right\}
$$

which can be written as

$$
\Pi_H = q(pW - L) + \left( \frac{1 - q}{W} \right) \left\{ X(X + L) + \frac{1}{2} W^2 - (X + L)^2 - L(W - X - L) \right\}
$$

$\Pi_H = C$ defines a critical value of $p$, that we will call $p_c$, such that the patent will only be challenged if $p \geq p_c$. Assume that $p_c > (X + L)/W$, so that we begin with a world in which patents are challenged. An increase in $p_c$ reduces the number of cases that are challenged in the first place.
We can now compute the effect of $X$ on $p_c$:

$$\frac{\partial p_c}{\partial X} = -\frac{(2 - q)(X + L)}{qW^2} < 0$$

This equation implies that a reduction in the settlement payoff $X$ will increase $p_c$, the minimum ex ante assessment of patent invalidity required to motivate incurring the costs of launching a challenge when, if a patent is challenged, the challenger expects initially to litigate to a decision (rather than settling). An increase in $p_c$ will reduce the number of challenges (and hence the number of cases that have the chance of going to a judicial decision).

To confirm that a decrease in $X$ reduces the number of patent challenges, we now define the mass $K$ of patents that are challenged and go all the way to trial:

$$K = \frac{(1 - p_c)(1 - q)(W - X - L)}{W}$$

So that

$$\frac{\partial K}{\partial X} = \left[\frac{1 - q}{W^2}\right] \frac{1 - q}{W} (X + L)(W - L - X) - (1 - p_c)$$

Notice that we must have $W - L > X$ to have any case ever go to litigation. Hence the first term of the expression in the second set of square brackets is positive, while the second term is negative. However, the second term becomes ever smaller as $p_c$ gets closer to 1. Hence, even in this “case 2”, there must be ranges of parameters for which a reduction in the settlement payoff $X$ unambiguously reduces the total number of patents going to trial.

Intuitively, there are two opposing forces. A reduction in $X$ making settlement less attractive reduces the number of cases challenged (since $dP_c/dX < 0$). A proportion of these cases would have ended in trial. On the other hand, a reduction in the settlement payoff $X$ means that fewer patents which have been challenged will settle and more will go to final trial.