Injunctions, Hold-Up, and Patent Royalties

Carl Shapiro, University of California at Berkeley

Send correspondence to: Professor Carl Shapiro, Haas School of Business, University of California, Berkeley, CA 94705, USA; Tel: 510-642-5905; E-mail: Shapiro@haas.berkeley.edu.

A simple model is developed to study royalty negotiations between a patent holder and a downstream firm whose product is more valuable if it includes a feature covered by the patent. The downstream firm must make specific investments to develop, design, and sell its product before patent validity and infringement will be determined. The hold-up component of the negotiated royalties is greatest for weak patents covering a minor feature of a product with a high margin between price and marginal cost. For weak patents, the hold-up component of negotiated royalties remains unchanged even if negotiations take place before the downstream firm designs its product. The analysis has implications for the use of injunctions in patent infringement cases. (JEL K20, O34)

1. Introduction

Patents are an increasingly important element of business strategy. In fiscal year 2009, 485,000 patent applications were filed with the U.S. Patent and Trademark Office (PTO), tripling over the past 20 years.1 Some 1.2 million patent applications were pending, roughly tripling over

---

the past decade. During fiscal year 2009, some 190,000 patents were issued, and over a 10-year period, some 1.8 million patents were issued, generating a large stock of in-force patents. Patent litigation has also risen significantly over the past decade, as reflected in the number of cases filed and the average size of the damages awarded for infringement.

Ultimately, patents have commercial force based on the remedies that are available to patent holders who prove in court that their patents are valid and infringed. Patent law provides that such prevailing patent holders will be awarded damages, based either on the profits they lost due to the infringement or the “reasonable royalties” they should have been paid by the infringing party. In addition, until recently, once a patent was found valid and infringed, the courts routinely issued injunctions requiring the infringing party to cease selling its infringing product.

While such permanent injunctions are fundamental to the property rights typically associated with patents, they have proven quite contentious, as is starkly illustrated by the widely publicized patent infringement case between NTP, Inc., and Research in Motion (RIM). NTP, a patent-holding company, claimed that RIM, the provider of the popular BlackBerry wireless e-mail device, had infringed several of NTP’s patents. After a jury found NTP’s patents valid and infringed by RIM, NTP asked the court to issue an injunction to stop RIM from selling infringing BlackBerry devices. As a result, RIM came under enormous pressure to settle the case to avoid a shutdown of the BlackBerry service, which could have resulted from a court injunction forcing RIM to stop infringing NTP’s patents. In March 2006, RIM paid $612.5 million to NTP to settle the case. To many observers, this payment reflected the strong bargaining position NTP enjoyed by virtue of its threat to shut down BlackBerry service, not the underlying value of NTP’s patented technology.

4. See Lemley and Shapiro (2005) for a general discussion of the empirical evidence regarding patent licensing and litigation.
Shortly after the settlement between NTP and RIM, the Supreme Court issued a landmark ruling regarding injunctions. In *eBay v. MercExchange*, a unanimous court struck down the approach taken by the Federal Circuit Court of Appeals, under which permanent injunctions were issued “absent exceptional circumstances.” The Supreme Court ruled unanimously that the district court has discretion whether to grant or deny injunctive relief based on traditional principles of equity, using a four-factor test.\(^6\)

The concurring opinion written by Justice Kennedy, joined by Justices Stephens, Souter, and Breyer, states:

> When the patented invention is but a small component of the product the companies seek to produce and the threat of an injunction is employed simply for undue leverage in negotiations, legal damages may well be sufficient to compensate for the infringement and an injunction may not serve the public interest. In addition injunctive relief may have different consequences for the burgeoning number of patents over business methods, which were not of much economic and legal significance in earlier times. The potential vagueness and suspect validity of some of these patents may affect the calculus under the four-factor test.

Precisely this fear of injunctions has led many leading companies in the information technology sector to complain about so-called “patent trolls” who, while responsible for little or no novel and non-obvious inventions, are able to obtain significant patent royalty payments from companies with revenue streams that can be put at risk in patent infringement cases.\(^7\)

At least prior to the *eBay* decision, once the defendant in a patent infringement case was found to be infringing a valid patent, the patent holder had a virtually automatic right to obtain a court-ordered injunction preventing the defendant from continuing to sell its infringing product. Such injunctions were routinely granted even if the patent covered only a minor feature of a complex, valuable, and popular product. With this rule, patent owners have been in a strong bargaining position, even the owners of weak patents covering only minor inventions. By obtaining an injunction, the owner of a patent who prevails in patent litigation, as a practical matter, 

---


has the power to stop the defendant from selling even a *non-infringing* version of the product, at least until the defendant can redesign its product and introduce a non-infringing version. The right to obtain an injunction thus gives the patent holder the power to hold up an infringing firm that has made specific investments to design, manufacture, and sell the infringing product. The prospect of such hold-up affects the negotiating strengths of the two parties prior to the onset of litigation.

Concerns about injunctions in patent cases are especially common in the information technology sector of the economy, including computer software, Internet business methods, semiconductors, and computer hardware and telecommunications products. First, there has been a surge of patenting of software and business methods over the past 10 years, as documented by Bessen and Hunt (2007). Second, there have been widespread complaints about patent quality and about vague and overly broad patents in this area, as reported by the Federal Trade Commission (2003) and the National Academies of Science (2004) and Jaffe and Lerner (2004). Third, software innovations tend to be incremental, with rapid sequential innovation; see, for example, Cohen and Lemley (2001). Fourth, software and hardware products tend to be complex, so a single product can potentially infringe many patents; see Heller and Eisenberg (1998), Hall and Ziedonis (2001) and Shapiro (2001). Fifth, software and hardware commonly are sold at prices well above marginal cost; the resulting margins commonly are necessary to provide a return on various investments, including R&D. Lastly, it can be costly and time consuming to redesign these products to avoid infringement claims. Hence, it is no coincidence that many firms in the information technology sector weighed in strongly in the *eBay* case and are pushing hard for patent reform.

This paper develops a model of licensing negotiations to show how injunctions affect the royalties that will be negotiated between patent

---

8. More recently, a number of commentators have argued that the patent system is out of balance and in many ways impedes rather than promotes innovation. See Burk and Lemley (2009), Bessen and Meurer (2008), and Heller (2008). Boldrin and Levine (2008) go so far as to argue for abolishing intellectual property altogether.

9. For example, see the positions on patent reform taken by the Coalition for Patent Fairness, www.patentfairness.org, which includes many firms in the information technology sector.
holders and technology users accused of infringement. Focusing on negotiated royalties is empirically justified since far more patents are licensed than are litigated to judgment. This paper takes as given the set of patents that are issued by the PTO, while recognizing that patents differ widely in their “strength,” i.e., the probability they will be held valid and infringed if litigated to judgment.

In the model developed here, downstream users who may be infringing a valid patent are subject to hold-up because they must make sunk investments that are specific to using the patented technology. The novel feature of the hold-up and opportunism problems identified here is that patents are probabilistic property rights, as recognized by Gallini (2002) and emphasized by Lemley and Shapiro (2005). Farrell and Shapiro (2008) explore the licensing of a probabilistic patent to a number of competing downstream firms. Assuming that each downstream firm can immediately and costlessly shift to a backstop technology if enjoined from using the patented technology, they focus on the form of licensing agreements and on the competitive interactions among the downstream firms. The current paper takes a complementary approach to the licensing of probabilistic patents. Here, we study licensing negotiations between a patent holder and a single downstream firm, focusing on the cost and disruption imposed on the downstream firm if it is precluded by a court order from using the patented technology.

The model of licensing negotiations developed here is designed to incorporate a number of key features that can give rise to hold-up and opportunism in patent infringement cases:

- **Probabilistic Patents**: Royalties are negotiated in the shadow of patent litigation. The relationship between patent strength and the level of negotiated royalties is derived.
- **Injunction Threat**: If patent litigation ensues, the downstream firm will continue producing and selling its product during the pendency of the patent litigation. However, if the patent is found valid and infringed, the court may issue an injunction forcing the downstream firm to withdraw its infringing product from the market. This pattern is very common in practice.
- **Patent Surprise**: The analysis includes the common case in which the downstream firm designs and begins selling its product before it is aware that it may be infringing the patent in question. The model shows how negotiated royalties in this case differ from royalties ne-
gotiated in situations where the downstream firm is fully aware of the patent holder’s assertions at the time it originally designs its product.

- **Patent on Minor Features:** The fraction of the total value created by the downstream product attributable to the patented invention is tracked as a parameter in the analysis. The model predicts that hold-up problems are greatest for patents covering a minor feature of a high-margin product.10

- **Redesign Time and Expense:** The cost that the downstream firm would have to bear to redesign its product to avoid using the patented technology and the time required to complete this redesign are tracked as parameters in the analysis. These two elements create the prospect of hold-up.

- **Reasonable Royalties:** The relationship between the level of “reasonable royalties” used to compute the patent damages and the level of negotiated royalties is derived. The level of reasonable royalties in a self-fulfilling equilibrium is calculated.

Section 2, “Royalty Negotiations,” presents the basic modeling elements. Section 3, “Surprise,” studies the case in which the downstream firm has already designed its product when faced with the patent holder’s infringement claims. The negotiated royalty is decomposed into a portion attributable to the patented invention and a portion attributable to hold-up. This section identifies the factors that determine the fraction of the patent holder’s overall payoff attributable to hold-up. Section 4, “Staying Permanent Injunctions to Permit Redesign,” shows that the portion of the patent holder’s payoff attributable to hold-up is reduced if the courts routinely stay injunctions to give infringing firms time to redesign their products and introduce non-infringing versions. Section 5, “Early Negotiations,” shows, somewhat surprisingly, that the patent holder can still profit from the prospect of hold-up even if the downstream firm is fully aware of the patent infringement claim against it when it initially designs its product. In fact, for weak patents, early knowledge of potential infringement is of no value at all to the downstream firm. Section 6, “Reasonable Royalties in Self-Fulfilling Equilibrium,” shows that the hold-up component of the patent holder’s payoff is even greater.

---

10. As stated by the General Counsel of Intel: “A fundamental invention deserves greater value than a relatively minor tweak to work that went before it. A broad application of the injunction remedy makes all patents “crucial,” whether they are or not.” See “Troll Call,” by Bruce Sewell, *Wall Street Journal*, March 6, 2006.
if reasonable royalties are determined endogenously in a self-fulfilling equilibrium. Section 7 discusses the implications of this analysis for legal rules regarding injunctions in patent cases and for reform of the patent system. Section 8 concludes.

2. Royalty Negotiations

A patent holder P owns a single patent. A downstream firm D produces a product that can incorporate a feature covered by the patent. The patent holder and the downstream firm are not competitors.11

2.1. Patented Feature

The patented feature increases the value of D’s product by \( v \geq 0 \) to all consumers, in comparison with the best non-infringing alternative.12 We call \( v \) the “value of the patented technology.” The downstream firm has a clear incentive to build the patented feature into its product: doing so allows it to charge a price that is higher by \( v \).

Let \( p \) denote the price per unit that D receives with the patented feature, and let \( c \) denote the marginal cost to D, apart from any royalty payments to P. We call D’s per-unit margin \( m = p - c \) for products incorporating the patented feature. D’s per-unit margin for products not incorporating the patented feature equals \( m - v > 0 \). In some cases, \( m - v \) will be large in comparison with \( v \). This occurs for “complex technologies” that incorporate many features or components, such as a complex piece of hardware like the BlackBerry handheld device or the Intel Pentium microprocessor or a sophisticated software product such as Microsoft Windows, Microsoft Office, Adobe Acrobat, or Adobe Photoshop. In other cases, \( m - v \) can be very small relative to \( v \). This occurs if D earns little margin on its product without the patented technology.

Let \( X \) denote the number of units produced by D per unit time. For simplicity, we treat D’s rate of sales per unit time, \( X \), as independent of

11. The analysis presented here would need to be modified to address cases in which the patent holder competes against the alleged infringer and thus can claim patent damages based on lost profits, not just “reasonable royalties.”

12. The value of the patented technology is measured net of any extra marginal costs caused by the patented feature. The analysis here is unchanged if the patented technology reduces the unit cost of production by \( v \).
whether or not D incorporates the patented feature into its product. The relevant patent lifetime is normalized to the period \([0, 1]\), so the total number of units sold during the lifetime of the patent equals \(X\). We assume no discounting.

2.2. Product Design and Redesign

The downstream firm makes an initial product design decision to either include or exclude the patented feature. We assume that the (fixed) initial product design costs borne by D are the same, whether or not D chooses to incorporate the patented feature into its product. Hence, we do not need to track these costs in our analysis.

If D initially incorporates the patented feature in its product, it is costly and time consuming for D to redesign its product later to avoid using this feature, as would be required for D to keep selling its product if P obtains an injunction against D and if the two parties do not sign a licensing agreement. We denote by \(F\) those (fixed) redesign costs. We denote by \(L \geq 0\) the lag from the time that D commits to incurring the redesign costs until the time when the redesigned product is ready for sale.

2.3. Patent Strength and Litigation Costs

If patent litigation occurs, there is a probability \(\theta\) that P’s patent will be held valid and infringed by D’s product, in which case we say that P wins the litigation. We call \(\theta\) the “patent strength.” With complementary probability \(1 - \theta\), the patent is ruled invalid or not infringed by D, in which case we say that D wins the litigation. Patent strength is common knowledge.

If the firms litigate, each must bear litigation costs, which we denote by \(C_P\) and \(C_D\), respectively. Our analysis focuses on the cases where these

---

13. Our analysis could be amended to account for circumstances in which the patented feature causes D to make extra sales. In that case, the analysis here would still apply to the sales that are not caused by the feature in question. Allowing sales to vary with time would not alter the basic analysis and could be accomplished simply by redefining the time variables in the analysis to reflect sales made as well as time passed.

14. Accounting for discounting would be straightforward. Each time variable in the analysis would just be redefined to measure the present discounted value of a constant annuity over that time period as a fraction of the present discounted value of a constant annuity lasting for the entire patent lifetime.
litigation costs are small enough relative to the stakes so that each party finds it worthwhile to incur its litigation costs rather than withdraw.

Litigation takes time $T < 1$. Since the patent lifetime is normalized as the period $[0, 1]$, $T$ is the duration of litigation as a fraction of the remaining lifetime of the patent. We define the “end” of litigation to be the time at which the ultimate winner of the litigation is determined. If $P$ wins the patent litigation, we assume that an injunction issues at this same point in time.\footnote{We use a highly simplified model of the litigation process. We assume that no preliminary injunction issues; in fact, such injunctions are rare. We also abstract away from intermediate rulings that cause the parties to update significantly their views on patent strength. We do not believe that our basic results are sensitive to this assumption. The analysis would be quite similar if one were to assume that an intermediate ruling does issue, so long as this ruling is highly accurate in terms of the ultimate disposition of the patent case, in which case one can think of the time from the intermediate ruling to the end of the litigation as a period during which the permanent injunction has been stayed, as analyzed below.} For simplicity, we assume that redesigning the product does not take as long at litigation, $L < T$, so the downstream firm that begins redesign work when sued for patent infringement can complete that work before facing a permanent injunction.

2.4. Nash Bargaining Over Royalties

We assume Nash Bargaining between $P$ and $D$, so they split any gains from trade available during any negotiations they have. We denote $P$’s bargaining skill by $\beta \in [0,1]$, so $P$ captures its disagreement payoff plus a fraction $\beta$ of the gains from trade. Likewise, $D$ captures its disagreement payoff plus a fraction $1 - \beta$ of the gains from trade. The disagreement payoffs used here are the payoffs that result if each party pursues its optimal credible strategy if negotiations are unsuccessful. In the text, we assume that bargaining occurs once prior to litigation and, if necessary, once again after the litigation is resolved. The Appendix shows that the hold-out component of the patent holder’s reward is at least as large if bargaining is ongoing.

Since we are studying Nash Bargaining in a model with symmetric information and since the combined payoffs of the two firms are larger under agreement (initial licensing) than under initial disagreement (no licensing), we know that the model must predict licensing, not litigation.
Therefore, this model should not be viewed as offering predictions about the likelihood of litigation. Rather, it informs the terms on which patent settlements, i.e., licensing, will occur. Even though the parties do not litigate in equilibrium, the rules regarding injunctions and damages do affect the equilibrium royalty rate because they affect the parties’ payoffs from litigation, and the parties negotiate a licensing agreement in the shadow of litigation.\textsuperscript{16}

2.5. Benchmark Royalty Rate

In order to isolate the effects of redesign costs and lags, we define a benchmark royalty rate that would result if the downstream firm could redesign its product without cost or delay, in which case there is no prospect for hold-up. In this simple and special case and if litigation costs are zero or neutral,\textsuperscript{17} the negotiated royalty rate would be equal to $\bar{r} \equiv \beta \theta v$. This benchmark reflects the value of the patented feature, $v$, discounted by the patent strength, $\theta$, and by the underlying bargaining skill of the patent holder, $\beta$. The corresponding benchmark level of profits for the patent holder is $\bar{\pi} = \bar{r} X = \beta \theta v X$. Comparing the negotiated royalty rate with this benchmark allows us to isolate the component of the negotiated royalties attributable to hold-up. Section 7 discusses this benchmark further.

2.6. Patent Damages and Reasonable Royalties

If the patent is valid and infringed, D will owe P damages for any past infringement. We denote by $s$ the “reasonably royalty” per unit that the court will require D to pay in this event.\textsuperscript{18} This variable is measured to

\textsuperscript{16} The vast majority of interactions between patent holders and alleged infringers result in licensing agreements. As reported in Lemley and Shapiro (2005), about 97\% of all filed patent cases settle. Furthermore, if we suppose that three to five patent disputes result in a licensing agreement for every one that leads to a patent suit even being filed, then there are more than one hundred patent licenses for every patent litigation that results in a final judgment.

\textsuperscript{17} Below, we show that litigation costs are neutral, i.e., they do not affect negotiated royalties if $\beta C_D = (1 - \beta) C_P$. This condition is met if the two parties have equal litigation costs and equal bargaining skill.

\textsuperscript{18} We are not studying the case in which P competes against D and thus asserts damages from D based on lost profits. Our analysis is confined to the case in which P’s damage claim only involves reasonable royalties.
include the possibility that D will be judged to have engaged in willful infringement, which can lead to a trebling of damages. To reduce the number of cases, we assume that the expected royalties are no greater than the value of the patented feature: $\theta s \leq v$. This inequality is satisfied if the “reasonable royalty” is no more than the value of the patented feature, i.e., $s \leq v$; this includes the benchmark case (see just below) where $s = \beta v$.

We initially treat $s$ as exogenous. This allows us to see how alternative rules governing the determination of patent damages affect negotiated royalty rates. However, after we calculate the equilibrium royalty rate for any given level of $s$, we perform much of our analysis assuming that $s = \beta v$. As explained below, this is a natural benchmark level for reasonable royalties under patent law. In particular, it is the royalty rate that would be negotiated between $P$ and $D$ if $D$ were aware of $P$’s patent when $D$ initially designed its product and if $P$’s patent were known to be valid and infringed by $D$. In Section 6, we show how $s$ can be determined endogenously in a fulfilled-expectations equilibrium.

2.7. Timing of Product Design, Negotiations, and Litigation

We consider two basic models which are similar in spirit but differ in their timing.

In the first model, “Surprise,” the downstream firm is unaware that $P$ may obtain a patent on the relevant feature at the time that $D$ makes its initial product design decision. Naturally, since the feature adds value, $D$ incorporates the patented feature into its product.19 The extensive form game begins when $P$ subsequently asserts its patent against $D$’s product and the two firms negotiate over patent royalties. If these initial licensing negotiations fail, $P$ decides whether or not to sue $D$ for patent infringement. If $P$ sues $D$, $D$ then decides whether or not to sell its allegedly infringing product. $D$ also can work on redesigning its product. If $P$ wins the litigation, the parties have another opportunity to negotiate a patent license subsequent to the court’s ruling.

19. This situation, which is common in the information technology sector, implicitly involves independent invention: $P$ and $D$ both “discover” the patented technology independently. As explained in Shapiro (2006) and Shapiro (2007), another, arguably superior way to deal with independent invention is to establish an independent invention defense, at least if $D$ uses the patented invention before the patent application is published or the patent issues.
The second model, “Early Negotiations,” is the same as the “Surprise” model except that D is aware of P’s patent, and the two parties have an opportunity to negotiate over royalties before D makes its initial product design decision.

3. Surprise

In this section, we study the situation in which D was unaware of P’s pending or issued patent at the time that D initially designed its product to incorporate the patented feature. This fact pattern occurs frequently, either because D designed its product before P’s patent application was published and before P’s patent issued or because D was simply unaware of P’s patent application or issued patent when D designed its product, even though D’s design efforts occurred after this information had become public.20

We are interested in understanding the factors that govern the negotiations between P and D over a patent license once P asserts its patent against D’s product. More specifically, our model is designed to explain the royalty rate likely to emerge from those negotiations. We pay particularly close attention to how injunctions and the rule by which reasonable royalties are determined affect the royalty rate negotiated between P and D.

As usual with Nash Bargaining, to determine the negotiated per-unit royalty rate, \( r \), we need to calculate each party’s payoff from agreeing on that royalty rate and each party’s disagreement payoff, which here involves patent litigation. The agreement payoffs are straightforward. The payoff to the downstream firm from accepting a license at rate \( r \) is given by \((m - r)X\). P’s payoff from this license is \( rX \). Their combined payoffs from licensing are simply \( mX \).21

---

20. The second of these possibilities need not imply that D was derelict or actively ignoring or evading or willfully infringing P’s patent, given the large number of patents, many of which have broad and vague claims.

21. We express the payment from D to P in terms of a uniform per-unit royalty rate, \( r \), which will apply for the lifetime of the patent. Allowing a fixed licensing fee or more generally a two-part tariff would not matter at all in our model, given our assumption that D sells a fixed number of units, whether or not D’s product incorporates P’s patented feature. Allowing royalty rates that vary with time also would not add anything to the analysis.
3.1. The Downstream Firm’s Threat Point

What happens if P and D do not reach an initial licensing agreement? In that event, P decides whether or not to initiate patent litigation against D. If P does not sue D, the game is over and P receives no royalties.\textsuperscript{22} We focus on the case where P’s litigation costs are small enough relative to the expected payoff from litigation that P’s threat to sue is credible and where D’s litigation costs are small enough relative to the value of staying in the market that D litigates rather than exiting.\textsuperscript{23} Since \( m > v \geq \theta s \), D finds it profitable to continue selling its product rather than withdraw it during the pendency of litigation, despite the possibility that D will subsequently be liable for patent damages. Our focus on these situations fits with our interest in valuable and complex products for which the patent involves only a relatively small component in the overall product. In these settings, unlike in pharmaceuticals, firms frequently find it optimal to keep selling their products in the face of patent infringement claims, even though doing so runs the risk of incurring liability for infringement.\textsuperscript{24}

The downstream firm also must decide whether to commit resources right away to redesigning its product. The analysis thus breaks into two cases, depending upon which of these strategies is optimal for D:

- **Redesign**: Develop a non-infringing version right away.
- **Do Not Redesign**: Do not develop a non-infringing version at this time.

The Appendix proves the following:

**Lemma 1.** The downstream firm’s optimal strategy in the absence of a licensing agreement is “Redesign” if and only if \( \theta > \frac{1}{b} \frac{F}{(m-v)L + F} = \theta^* \).\textsuperscript{25}

\textsuperscript{22} In a model where P could choose the date to initiate patent litigation, P would not find it optimal to wait and sue later. Delay gives D a chance to begin redesigning its product, brings the date at which the patent expires closer, and offers no advantage to P since D is not making any new investments specific to the patented technology.

\textsuperscript{23} The Appendix derives sufficient conditions on litigation costs such that litigation is credible for each party.

\textsuperscript{24} The patent holder is in an even stronger position if the downstream firm would withdraw its product during the pendency of patent litigation.

\textsuperscript{25} If \( \theta^* \geq 1 \), D never engages in redesign, regardless of patent strength, and the analysis simplifies to the case in which “Do Not Redesign” is optimal for D for all values of \( \theta \).
The downstream firm is more likely to redesign its product during litigation (a) the stronger is the patent, (b) the greater is the patent holder’s bargaining skill, (c) the larger are the total margins that D could earn without the patented feature which are at risk due to an injunction if P wins the patent litigation, \((m - v)XL\), and (d) the smaller are the redesign costs. Intuitively, if the patent is weak, D will not find it optimal to incur product redesign costs during the pendency of litigation, since it is relatively unlikely that the redesigned product will ever be needed.

3.2. Negotiated Royalties for Relatively Weak Patents

For \(\theta < \theta^*\), D’s threat point is “Do Not Redesign.” In this case, denoting the negotiated total royalties by \(\pi^*\) and the negotiated per-unit royalty rate by \(r^*\), the Appendix proves the following

**Theorem 1.** Suppose that reasonable royalties are set at their benchmark level, \(s = \beta v\). For relatively weak patents, i.e., \(\theta < \theta^*\), the patent holder’s payoff equals

\[
\pi^* \equiv r^*X = \beta \theta [vX + (m - v)LX + F] + [\beta C_D - (1 - \beta)C_P].
\]  

The first term, \(\beta \theta vX\), reflects the value of the patented feature, \(vX\), discounted by the probability that the patent is valid and infringed, \(\theta\), and by the patent holder’s bargaining skill, \(\beta\). The second term, \(\beta \theta (m - v)LX\), measures P’s ability to hold up D based on the lag time associated with design-around. This expression depends upon the value to D of using non-patented technology, \((m - v)\), as well as the redesign lag, \(L\). The third term, \(\beta \theta F\), measures P’s ability to hold up D based on the fixed costs associated with redesigning the product. The final term, \(\beta C_D - (1 - \beta)C_P\), reflects P’s net bargaining advantage associated purely with fact that failure to reach a licensing agreement will impose litigation costs on both parties. The net impact of this threat depends upon the two firms’ relative litigation costs and bargaining skill. This expression can be positive or negative. This term is larger, the greater is P’s bargaining skill, the greater are D’s litigation costs, and the smaller are P’s litigation costs. This term is zero in the neutral case in which \(\beta = 1/2\) and \(C_P = C_D\).
3.3. Negotiated Royalties for Relatively Strong Patents: \( \theta > \theta^* \)

For \( \theta > \theta^* \), D’s threat point is “Redesign.” In this case, the Appendix proves the following

**Theorem 2.** Suppose that reasonable royalties are set at their benchmark level, \( s = \beta v \). For relatively strong patents, i.e., \( \theta > \theta^* \), the patent holder’s payoff equals

\[
\pi^* \equiv \pi(X) = \beta \theta vX + \beta F + \frac{1}{2} \beta CD - (1 - \beta) CP.
\]

We have already discussed the first and last terms on the right-hand side of this expression. The new term, \( \beta F \), is the amount that P can extract per unit because D’s threat point involves redesign costs of \( F \). Note that this term is not discounted by patent strength because D’s threat point involves incurring the redesign costs whether or not the patent is valid.

3.4. Negotiated Royalty Rates and Patent Strength

Figure 1 shows how the patent holder’s payoff varies with patent strength. For simplicity, we now introduce the standing assumption that the litigation cost term \( \beta CD - (1 - \beta) CP \) is zero.

The heavy straight line through the origin depicts Equation (1), which applies when \( \theta < \theta^* \). The flatter straight line in Figure 1, beginning at \( \pi^* = \beta F \) when \( \theta = 0 \), depicts Equation (2), which applies when \( \theta > \theta^* \). The equilibrium profits, \( \pi^* \), are depicted by the two heavier line segments in Figure 1. Note that the patent holder’s payoff drops discontinuously at \( \theta = \theta^* \). This reflects the idea that P’s payoff is discretely lower if D’s threat point is to redesign. Figure 1 also displays a line representing the benchmark level of royalties, \( \pi = \beta \theta vX \).

The two lines in Figure 1 depicting Equations (1) and (2) cross at \( \theta^{**} = \beta \theta^* \). For \( \theta \in (\theta^{**}, \theta^*) \), D would be better off if it could credibly commit to redesigning its product in the event no licensing agreement is reached. A credible threat to redesign the product would help D negotiate with P since redesign leaves P in a weaker negotiating position in the event that P wins the patent litigation.

3.5. Impact of Hold-Up for Relatively Weak Patents

We now compare the patent holder’s payoff with the benchmark level of \( \pi = \beta \theta vX \) that applies without redesign costs or lags. For relatively weak
patents, i.e., $\theta < \theta^*$, using Equation (1), the additional payoff to the patent holder due to hold-up is given by $
abla^* - \nabla = \beta \theta [(m - v)LX + F]$. Measured relative to the benchmark payoff, the hold-up term equals

$$\frac{\nabla^* - \nabla}{\nabla} = \frac{m-v}{v}L + \frac{F}{vX}. \quad (3)$$

The first term on the right-hand side of Equation (3) is the ratio of the value of the product without the patented feature to the value of that feature, times the percentage of the patent lifetime that is required for redesign. This term reflects the patent holder’s power based on the threat that an injunction will force the downstream firm’s product from the market during redesign. The second term is the ratio of the redesign costs to the total value of the patented improvement. Note that the expression in Equation (3) is independent of the patent strength and the patent holder’s bargaining skill.

**Corollary 1.** For relatively weak patents, i.e., $\theta < \theta^*$, the payoff to the patent holder due to hold-up relative to its payoff without hold-up is larger, the larger is the margin on a non-infringing product relative to the per-unit value of the patented feature, the time required to redesign the product to avoid infringement, and the ratio of the redesign costs to the total value of the patented feature.
We can illustrate the effects of hold-up using a numerical example. Suppose that the product sells for $40 without the patented feature and the marginal cost of producing the product is $30 per unit, so \( m - v = 10 \). Suppose the patented feature adds an extra $1 of value, so \( v = 1 \). This implies that \( (m - v)/v = 10 \). Suppose that redesign will take 1 year out of 10 years remaining in the patent lifetime, so \( L = 0.1 \). Then the first term on the right-hand side of Equation (3) equals \( 10 \times 0.1 \) or 1.0. The patent holder’s payoff due to hold-up is as large as its payoff based on the value of its patented feature. Put differently, the threat that \( P \) will obtain an injunction if its patent is proven valid and infringed causes the royalty rate to be twice as high as the benchmark level. In addition, redesign is costly. If the redesign costs equal half of the total value of the patented feature, \( F/vX = 0.5 \), causing the royalty rate to rise by an additional 50% above the benchmark level. The Appendix shows that \( \theta^* = 2/3 \) with these parameter values.

3.6. Impact of Hold-Up for Relatively Strong Patents

We can perform the same exercise for relatively strong patents, i.e., \( \theta > \theta^* \), in which case \( D \)’s threat point is “Redesign.” Using Equation (2), the additional payoff to the patent holder due to hold-up is given by \( \beta F \). Measured relative to the benchmark payoff, the hold-up term is now given by

\[
\frac{\pi^* - \pi}{\pi} = \frac{F}{vX \theta}
\]

(4)

The right-hand side is the product of two ratios: the ratio of the redesign costs to the total value of the patented invention to the downstream firm and the inverse of the patent strength.

**Corollary 2.** For relatively strong patents, i.e., \( \theta > \theta^* \), the payoff to the patent holder due to hold-up relative to its payoff without hold-up is larger, the larger is the ratio of the redesign costs to the total value of the patented feature and the weaker is the patent.

We can illustrate the effects of hold-up by modifying the numerical example given above so the marginal cost of production is $10, which
implies $m - v = 30$. If the patent strength is $\theta = 0.5$, then P’s threat to force D to redesign leads to a royalty rate twice the benchmark level.26

4. Staying Permanent Injunctions to Permit Redesign

We now show how licensing negotiations in our model are affected if the courts regularly stay the permanent injunctions that they grant and if these stays last long enough to give downstream firms the opportunity to complete their redesign efforts.27 If stays are routinely granted and reasonable royalties equal their benchmark level, $s = \beta v$, then D has no incentive to redesign its product prior to the resolution of litigation. D’s optimal strategy for all $\theta$ is “Do Not Redesign,” and the Appendix shows that patent holder’s payoff equals

$$\pi^{**} = \beta \theta [vX + F].$$

(5)

Figure 2 displays the heavy straight line through the origin which represents the patent holder’s payoff if injunctions are stayed, as given by Equation (5). For comparison purposes, the patent holder’s payoff without stays, from Figure 1, is also shown on Figure 2. Granting stays allows the downstream firm to delay its redesign efforts until it learns the outcome of the patent litigation. For this reason, stays are of no value for $\theta = 0$ or $\theta = 1$; for these extreme values of patent strength, there is no information to be learned. Stays are especially helpful for patents of intermediate strength, for which learning the outcome of the patent litigation is most informative.

**Theorem 3.** Routinely granting stays to permanent injunctions to provide infringing firms the time to design non-infringing products causes the hold-up component of the patent holder’s payoff to fall, moving the patent holder’s payoff closer to the benchmark level. Stays are most valuable to alleged infringers for patents of intermediate strength.

---

26. The Appendix shows that $\theta^* = 2/7$ with these parameter values, so “Redesign” is indeed the optimal strategy for D if $\theta = 1/2$.

27. We assume that redesign is more profitable for D than exiting the market. If not, then stays simply extend the time period in which D can use the patented technology in exchange for damages of reasonable royalties.
If stays to permanent injunctions are routinely granted, the hold-up component of the patent holder’s payoff is given by

\[
\frac{\pi^* - \bar{\pi}}{\bar{\pi}} = \frac{F}{vX}.
\]  

(6)

For relatively strong patents, this is a multiple \(\theta\) of the hold-up component of the patent holder’s payoff in the absence of stays. For these patents, the hold-up component of the patent holder’s payoff is based on the downstream firm incurring the fixed redesign costs, and stays allow the downstream firm to incur that cost only in the event the patent is found valid and infringed. Stays reduce but do not eliminate the hold-up component of the patent holder’s payoff.

In our previous numerical examples, we assumed that \(F/vX = 0.5\), i.e., the redesign costs were half of the total value of the patented feature. In that case, granting stays limits the hold-up component of the patent holder’s payoff to 50% of its benchmark payoff. In contrast, in our numerical example where “Do Not Redesign” was optimal, the threat of an injunction led to this 50% plus an additional 100% due to the lag associated with redesign.

5. Early Negotiations

We now consider how our analysis changes if the patent holder and the downstream firm negotiate before D initially designs its product. In particular, we now assume that P and D enter into licensing negotiations sufficiently early that D can design its product to include, or exclude, the
patented feature, at no extra design cost, and still have sufficient time to introduce its product as planned at time zero.

How are the negotiations between P and D affected if D has not yet designed its product? The only difference from the earlier model is that D has an additional option: D can design its product initially to avoid any chance of infringing P’s patent. We call this the “Design Around” strategy. D’s payoff from “Design Around” is \((m - v)X\). The Early Negotiations game differs from the Hold-Up game if and only if D’s payoff from “Design Around” is higher than D’s payoff from both “Do Not Redesign” and “Redesign.”

If the “Design Around” option is valuable to D, then D’s threat point payoff is \((m - v)X\), so the gains from trade equal \(vX\). Since P gets a fraction \(\beta\) of the gains from trade, the negotiated royalty rate must be \(\beta v\), which is above the benchmark level of \(\theta \beta v\) for all \(\theta < 1\). For sufficiently weak patents, Theorem #1 tells us that D can obtain a lower royalty than \(\beta v\) by threatening “Do Not Redesign.” This proves

**Theorem 4.** If early negotiations are of any benefit to the downstream firm, then the patent holder’s payoff equals \(\beta vX\), which exceeds the benchmark level of \(\beta \theta vX\). Early negotiations provide no benefit to the downstream firm for sufficiently weak patents.

The finding that, even with early negotiations, P and D will negotiate a royalty rate in excess of the benchmark level may be surprising, since the patent holder would not appear to have any ability to hold up the downstream firm in such early negotiations. However, when early negotiations are valuable to the downstream firm, D’s best threat, designing around the patent, is equivalent to conceding that the patent is valid and infringed without a fight. In this situation, the downstream firm does not get any reduction in royalties to reflect the probabilistic nature of the patent, so the royalty rate, \(\beta v\), is not discounted at all to reflect any weakness of the patent.

Another way to see why early negotiations are not valuable for weak patents is to consider how the downstream firm will respond when

---

28. D prefers the “Redesign” strategy to the “Design Around” strategy if the extra redesign costs, \(F\), are less than the expected benefits of selling a potentially infringing product during the pendency of litigation, \((v - \theta s)XT\).
approached early by the holder of a weak patent. Designing around the patent means giving up the value $v$ just to avoid the relatively small risk that the patent will prove valid. So the downstream firm’s credible threat is to proceed ahead and design the product to risk infringing the patent. But this means that the equilibrium in the early negotiations game is the same as in the hold-up game.

If the downstream firm does benefit from Early Negotiations, then the patent holder’s payoff is $\beta vX$, and the gap between the patent holder’s payoff and the benchmark payoff is given by

$$\frac{\pi^* - \bar{\pi}}{\bar{\pi}} = \frac{1 - \theta}{\theta}.$$  (7)

In this case, the proportional impact of the prospect of hold-up depends on the patent strength but not on any other variables. As an example, if the patent strength is $\theta = 0.5$, the threat of hold-up leads to a doubling in the patent holder’s payoff. Even for a rather strong patent, say $\theta = 0.8$, the threat of hold-up raises the patent holder’s payoff by 25% over the benchmark level.

6. Reasonable Royalties

We have assumed so far that the reasonable royalty rate was set at the benchmark level of $s = \beta v$. We now show that the hold-up problems just identified are magnified when $s$ is determined endogenously.

6.1. Reasonable Royalties in Self-Fulfilling Equilibrium

The law governing patent damages states that “the court shall award a claimant damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer.”29 Under established precedent, the reasonable royalty rate is usually defined to be the royalty rate that would be negotiated initially between the two parties if the patent were known to be valid and infringed and if they were willing and able to reach an agreement.30 In

29. 35 United States Code §284.

30. See, for example, Leonard and Stiroh (2005). The key case articulating this principle is Georgia-Pacific Corp. v. United States Plywood Corp., 446 F. 2d 295 (Second Circuit, 1971). For more recent Federal Circuit authority, see Rite-Hite Corp. v. Kelly Co., 56 F.3d 1538 (Fed. Cir. 1995) (en banc).
terms of the model presented here, the reasonable royalty should therefore equal $\beta v$, which is the benchmark level of $s$ used above.

In practice, however, even if the courts accept the principle that $s$ should equal $\beta v$, they face the problem that $\beta v$ may be difficult for juries to estimate with accuracy. Therefore, the courts have developed a series of factors that juries should consider when calculating reasonable royalties. In practice, the courts pay close attention to the royalties actually negotiated by the patent holder with other licensees for the patented technology. They also use the royalty rates negotiated for other “comparable” patents as proxies. In other words, in setting $s$, the courts rely on royalty rates that have actually been negotiated.

In terms of the variables in our model, the approach taken by the courts involves using the equilibrium licensing rate for a patent with strength $\theta = 1$ as a proxy for $s$. This approach necessarily introduces some degree of circularity into the definition of reasonable royalties, since $s$ depends upon the observed royalties that have actually been negotiated, and those royalties in turn depend upon $s$. We resolve this circularity by looking for a fulfilled-expectations equilibrium. The Appendix proves the following

**Theorem 5.** Suppose that court-determined reasonable royalties are based on the actual royalties negotiated between the patent holder and other similarly situated downstream firms. In a fulfilled-expectations equilibrium, the reasonable royalties are elevated above the benchmark level of $s = \beta v$ by the patent holder’s threat of hold-up.

The Appendix shows how these higher levels of reasonable royalties cause the level of negotiated royalties to exceed the levels shown above.

7. Implications for Patent Policy

The model developed here shows that patent holders gain a negotiating advantage based on hold-up when downstream firms—who are often themselves innovators—must make investments specific to the use of the patented technology prior to the resolution of uncertainty about patent validity and infringement. Given the large number of patents being issued, many with uncertain validity and scope, and given the lengthy
pendency of patent applications in the U.S. PTO, this situation appears to be fairly common, especially in the information and communications technology sector where many patents can read on a single product.

These observations have implications for the use of injunctions in patent cases. Naturally, one must exercise caution in drawing policy implications from any single model. The policy implications discussed here, like the model, are confined to situations in which the patent holder does not compete against the downstream firm and its damages claims are based on reasonable royalties, not lost profits. Following Lemley and Shapiro (2007a), I further confine the policy implications to situations in which the downstream firm has developed the technology independently rather than copying it from the patent holder.

The central implication of the model is that the courts can reduce or eliminate the hold-up component of negotiated patent royalties by selectively denying, or staying, permanent injunctions in patent cases involving non-competing patent holders whose damage claims are based on reasonable royalties. Since the gap between the benchmark royalties without hold-up and the negotiated royalties with hold-up is largest for weak patents covering a minor feature of a high-margin product, the model further suggests that those are the cases where denying or staying permanent injunctions will have the greatest impact. While it may not be practical for courts to determine which patents, having been litigated and found valid, had previously been weak, it does seem practical for the courts to distinguish cases based on the value of the patented feature in comparison with the margin between price and marginal cost on the infringing product. Since eliminating hold-up is generally desirable, it seems desirable for the courts to use the discretion granted to them under eBay to selectively limit the use of injunctions in this way.

A number of articles have challenged these policy implications on various grounds. See Elhauge (2008), Denicolò et al. (2008), Golden (2007), and Sidak (2008). Lemley and Shapiro (2007b) respond directly to Golden (2007), and the discussion in Shapiro (2007) regarding the patent holder’s social contribution in situations involving either complementary innovations or involving independent invention is directly relevant to this discussion. Here, I note a few of the main lines of attack and give very brief responses. For the purposes of this
discussion, I will refer to Theorem 1, which applies for relatively weak patents, and I will assume that litigation cost is neutral. In that case, using Equation (1), the patent holder’s payoff equals

$$\pi^* = \beta \theta [\nu X + (m-\nu)LX + F].$$

The first term on the right-hand side is the patent holder’s payoff without hold-up. The second term reflects the portion of the hold-up component attributable to the redesign delay, during which the downstream firm is unable to sell even non-infringing products. The third term reflects the portion of the hold-up component attributable to the fixed costs of redesign. In this case, Equation (3) expresses the patent holder’s payoff in comparison with the benchmark, non-holdup value $\pi = \beta \theta \nu X$ as:

$$\frac{\pi^* - \pi}{\pi} = \frac{m-\nu}{\nu} L + \frac{F}{\nu X}.$$  

**Objection 1.** The benchmark, no-holdup payoff to the patent holder is too low, lacks a normative basis, and is not suitable for policy purposes.

**Response 1.** The methodology used here is conventional in the field of economics. In particular, we isolate and study the effect of concern, which here is hold-up. The hold-up component of the patent holder’s payoff is not based on the value of the patented technology, $\nu$. There is little or no merit in increasing the patent holder’s payoff based on the costs and delays associated with downstream redesign. Hold-up discourages investments by downstream firms, who are often themselves innovators, and is a very poorly targeted way to reward patent holders.

The argument that hold-up is desirable because patent holders are generally under-compensated has no limiting principle. If patent holders really are systematically under-compensated, the patent system should be adjusted in other ways to increase their rewards, not by inefficiently enabling patent hold-up. Royalties based on hold-up are especially poorly targeted because the owners of weaker patents, i.e., the ones least likely to represent genuine innovation, benefit disproportionately from hold-up.
The critics assert that the proper benchmark is $\theta v$ rather than $\beta \theta v$, but this departs from the principle that reasonable royalties are based on \textit{ex ante} arms-length negotiations. Also, even if the patent holder’s (expected) social contribution were $\theta v$ per unit, it is not true that dynamic efficiency calls for the patent holder’s rewards to be equal to that amount because it is impossible (without subsidies) in the presence of complementary innovations for each innovator to appropriate 100% of its marginal contributions. With complementary innovations, the sum of the marginal contributions exceeds the total. Furthermore, the patent holder’s (expected) per-unit contribution to the downstream product is \textit{not} $\theta v$ in the case where the downstream firm independently develops the feature covered by the patent. In that case, which is the norm if the downstream firm did not copy the patented feature from the patent holder, the patent holder’s social contribution to this downstream product is nil. Shapiro (2007) elaborates on these points.

\textbf{Objection 2.} Denying injunctions to non-practicing patent holders is a form of discrimination against patent holders who adopt a particular business model.

\textbf{Response 2.} Patent damages law already distinguishes between infringement claims based on lost profits and those based on reasonable royalties. The analysis and policy recommendations here only apply to cases involving reasonable royalty claims. In those cases, the patent holder can be made whole on a forward-looking basis, without an injunction, by awarding reasonable royalties. This is the basic balancing of the equities noted by the Supreme Court in \textit{eBay}.

Furthermore, the trial court typically will already be determining reasonable royalties for the purpose of awarding retrospective damages. The forward-looking reasonable royalty rate is the same as the retrospective one, since they are based on the same hypothetical, arms-length \textit{ex ante} negotiation between the patent holder and the infringer. Determining lost profits on a forward-looking basis would be far more difficult, and this alone warrants distinct treatment.
Objection 3. The model is too simple to form the basis for policy recommendations, since it ignores asymmetric information and has only one patent and one downstream firm.

Response 3. The model is a useful step forward. The model identifies some fundamental aspects of hold-up with probabilistic patents. Further work would certainly be helpful to understand how some of the complicating factors listed affect royalty negotiations. But there is no reason to think that such factors fundamentally change the basic findings here. Plus, of course, patent policy must be based on empirical as well as theoretical findings.

Objection 4. The model is too simple to form the basis for policy recommendations because it does not include legal errors in the determination of reasonable royalties.

Response 4. Accounting for legal error can be important, but simply noting the possibility of legal error and invoking an error-cost framework does not alter the basic conclusions reached here.

The results of the model do not change in the presence of small or moderate errors in determining the reasonable royalty rate, so long as the court-determined royalties are unbiased. In the equilibrium of the model, there are gains from trade to be had between the patent holder and the downstream firm, and (by assumption) those gains will be achieved through successful negotiations. The court will not actually be called upon to determine the reasonable royalties. So long as the errors are small relative to the redesign costs and lags (see below), unbiased errors in the royalties that the court would determine do not matter.31 In reality, negotiations sometimes break down, but negotiations still take place in the shadow of litigation, so the expected value of the reasonable royalties is still what matters overall.

Sufficiently large errors in determining the reasonable royalty rate could favor the downstream firm, even if those errors are unbiased. See Denicolò et al. (2008). This can happen if these errors give the downstream firm a valuable option. The downstream firm could pay the court-determined

---

31. This statement assumes that P and D are risk neutral. Errors would create risk and thus disfavor the more risk-averse party. However, the financial markets provide numerous ways to share and spread the risk involved.
royalties when they are far too low and redesign the product when they are far too high. This potential problem might not arise, even for fairly large errors, for patents covering a small feature of a high-margin product: the downstream firm would pay greatly excessive royalties rather than withdraw its product from the market while engaging in the redesign.

**Objection 5.** Eliminating injunctions and awarding reasonable royalties destroys the “property” aspects on the patent system.

**Response 5.** First, the model and the associated policy recommendation to limit the use of injunctions are limited to cases involving reasonable royalty damage claims by a non-competing patent holder suing a downstream firm that did not copy the patented feature from the patent holder. As Lemley and Shapiro (2007) state rather firmly:

> While we strongly believe that the threat of holdup gives excessive reward to patent holders, especially in component industries, we consider the presumptive right to injunctive relief to be an important part of the patent law. In most cases, there will be no question as to the patentee’s entitlement to an injunction. To begin, we stress that our analysis in this Article is expressly limited to situations in which the patent holder’s predominant commercial interest in bringing a patent infringement case is to obtain licensing revenues.

In that case, restricting the use of injunctions does mean using a hybrid system involving liability elements once the patent is found valid and infringed. The Supreme Court has established that already in *ebay*. At issue here is the boundary between the property system and the liability system, i.e., between the use of injunctions and the awarding of reasonable royalties after a patent has been found valid and infringed.

There is no general presumption that a pure property system is superior to a hybrid system or a pure liability system. See Calebresi and Melamud (1972) and Kaplow and Shavell (1996). In particular, Kaplow and Shavell (p. 763) note that the two systems yield the same *ex post* allocation of resources if bargaining is always successful: the parties achieve the available gains from trade under either system. (The
systems differ in how the gains from trade are split). If bargaining is not always successful, the liability rule is superior in the current context. Ex post efficiency calls for the downstream firm to incorporate the patented feature into its product, so making it possible for the downstream firm to use the patented feature and then make the patent holder whole promotes ex post efficiency. If necessary, the downstream firm may be required to post a bond or pay in advance to insure that the patent holder really does receive the royalties it is owed. This objection thus reduces to the assertion that patent holders are systematically under-compensated, and reducing their ability to negotiate royalties based on hold-up is therefore undesirable. See Response #1.

The analysis here has implications for patent policy that go beyond the rules governing permanent injunctions in patent infringement cases. In particular, it supports the more general proposition that significant, adverse effects can arise when the PTO issues weak or vague patents. This same theme can be found in Shapiro (2007) and Farrell and Shapiro (2008). There are significant advantages to resolving patent validity, and clarifying patent scope at an early stage, before downstream implementers need to make designs regarding the development, design, and marketing of products that may later be judged infringing. The model developed here thus gives further support to the growing chorus of voices calling for policy changes that will improve patent quality and reduce patent pendency, e.g., by devoting greater resources to patent examinations. The model also gives further support for conducting additional post-grant reviews to weed out weak patents before they are licensed or litigated.

32. When bargaining is not always successful, Kaplow and Shavell suspect that the property system is superior because it avoids “undesirable takings when the owner values the thing more than the taker.” This makes good sense in their model, where the owner is assumed to generally value the thing more than the taker. However, in the patent infringement setting under study here, there are no such undesirable takings because there are always gains from trade associated with the use of the patented feature. This is a consequence of the fact that we are dealing with a non-competing patent owner. In that context, a liability rule reduces the possibility that bargaining breakdown will lead to inefficiency. Another problem with the liability system identified by Kaplow and Shavell, i.e., competition among takers for a single piece of property when damages are set too low, does not arise with intellectual property.
8. Conclusions

The model in this paper identifies some of the key factors governing the magnitude of the hold-up component of negotiated royalties for probabilistic patents. The model applies to situations in which the patent holder does not compete against the downstream firm and the patent holder’s damages claims are based on reasonable royalties, not lost profits.

The model is, by design, as simple as possible to illustrate the basic elements of hold-up with probabilistic patents. For weak patents covering a minor feature of a high-margin product that takes time to redesign, a large fraction of negotiated royalties can be attributable to hold-up, not to the value of the patented technology. This finding holds whether or not the downstream firm was aware of the patent prior to making its initial design decision.

The standard theory of hold-up and opportunism tells us that a downstream firm will be disadvantaged if it must negotiate for a patent license after it has made such specific investments. The contribution of this paper is to identify the key determinants of the hold-up component of patent royalties for probabilistic patents. One insight emerging from the model is that downstream firms can be subject to hold-up even if they are aware that they will be subject to a patent infringement suit before they make any specific investments.

The model supports the conclusion that the hold-up component of negotiated patent royalties will be reduced or eliminated if the courts, following the Supreme Court’s eBay decision, award reasonable royalties but do not issue injunctions in appropriate cases. Alternatively, the courts could grant stays on their injunctions, giving downstream firms time to redesign non-infringing versions of their products.

The model also addresses a circularity in the manner in which damages based on “reasonable royalties” are calculated in patent infringement cases. The circularity arises because reasonable royalties are often based on the royalties actually negotiated in the shadow of litigation, and these negotiated royalties depend in turn upon the magnitude of damages that courts are expected to award if the parties are unable to sign a licensing deal and instead engage in patent litigation. If this circularity is resolved in a fulfilled-expectations equilibrium, reasonable royalties exceed their benchmark level, increasing the hold-up component of patent royalties.
Further research is needed to understand how the effects identified here are altered when the patented feature significantly increases the sales made by the downstream firm, when the patent holder is a direct competitor of the allegedly infringing firm, and when multiple patents are asserted against the same product, either simultaneously or sequentially, by one or more patentees.

Finally, a more complex analysis, building on the model presented here, is required in situations where the patent holder negotiates with multiple licensees. In such cases, the patent holder has more at risk in patent litigation. A finding of invalidity will destroy the patent holder’s ability to collect royalties for the patent in question. On the other hand, a finding of validity will bolster the strength of the patent and allow the patent holder to obtain higher royalties in subsequent licensing negotiations. Just how these factors play out in situations where negotiations take place sequentially and the patent holder behaves strategically is a topic for future research.

Appendix

Disagreement Payoffs from the “Do Not Redesign” Strategy

We now compute D’s payoff from following the “Do Not Redesign” strategy if the initial licensing negotiations with P fail. This strategy entails litigation, so D incurs its litigation cost $C_D$. While litigation is ongoing, D continues selling its product, so D earns $mX$ per unit time during $[0, T]$, for a total of $mXT$, at which time litigation is completed. If D wins the litigation, which occurs with probability $(1 - \theta)$, its continuation payoff is $mX(1 - T)$.

If P wins the litigation, which occurs with probability $\theta$, then D owes damages to P equal to $sXT$, and P obtains an injunction against D. At that point, P and D can negotiate a license. Again, we assume Nash Bargaining between P and D. So again we need to calculate their payoffs from agreement and disagreement. At this point, both firms’ litigation costs have already been incurred, profits have already been earned on products already sold, and damages are already due based on those sales, so we can ignore those parts of the firm’s payoffs when considering the bargaining outcome going forward.
If P and D sign a licensing agreement, their combined prospective profits are \( mX(1 - T) \). If P and D do not reach an agreement, P gets nothing, and D is forced to either exit the market or incur the design-around costs and, after a lag, introduce a non-infringing product. While the redesign effort is underway, D must withdraw from the market.

There are two sub-cases, depending upon whether D is better off incurring the redesign costs or exiting the market. Exiting the market gives D a (prospective) payoff of zero, and redesign gives D a (prospective) payoff of \( (m - v)X(1 - T - L) - F \). We focus on the case in which redesign is more profitable. In this case, the downstream firm’s (prospective) profits of \( (m - v)X(1 - T - L) - F \) are equal to the combined disagreement profits of P and D. Subtracting this amount from the combined agreement profits of \( mX(1 - T) \) gives the gains from trade associated with reaching an agreement, which equal \( vX(1 - T) + (m - v)XL + F \). Under Nash Bargaining, these gains are split, so the payoff to P is

\[
\beta vX(1 - T) + \beta [(m - v)XL + F]. \tag{8}
\]

The prospective payoff to D if P wins is the combined prospective payoff, \( mX(1 - T) \), minus P’s prospective payoff, as just given in Equation (8). Simplifying, D’s prospective payoff equals

\[
(m - \beta v)X(1 - T) - \beta [(m - v)XL + F]. \tag{9}
\]

Now we are ready to work backwards to the initial negotiations. D’s expected payoff if those negotiations fail and D follows the “Do Not Redesign” strategy is equal to its prospective payoff, as shown in Equation (9), if P wins, plus the prospective payoff of \( mX(1 - T) \) if D wins, plus D’s expected profits over the period \([0, T]\), which are \((m - \theta s)XT\), minus D’s litigation costs, \( C_D \). Therefore, D’s threat point using the “Do Not Redesign” strategy equals

\[
(m - \theta s)XT + (1 - \theta)mX(1 - T) + \theta (m - \beta v)X(1 - T) - \theta \beta [(m - v)XL + F] - C_D.
\]

This expression can be written as

\[
mX - \theta X[\xi{T} + \beta v(1 - T)] - \theta \beta (m - v)XL - \theta \beta F - C_D. \tag{10}
\]

---

33. If D’s best option is to exit the market, then P and D split the gains from trade \( mX(1 - T) \), which involves a negotiated royalty rate of \( \beta m \). This case is more favorable to P than the case on which we focus.
We write this as $mX - E - CD$, where $E = \theta X[sT + \beta v(1 - T)] + \theta \beta (m - v)XL + \theta \beta F$ is the expected payment from D to P. With $s = \beta v$, $E = \theta \beta vX + \theta \beta (m - v)XL + \theta \beta F$ and D’s disagreement payoff is $mX - \theta \beta vX - \theta \beta (m - v)XL - \theta \beta F - CD$. D will litigate rather than exit if

$$CD \leq mX - \theta \beta vX - \theta \beta (m - v)XL - \theta \beta F.$$

P’s disagreement payoff equals $E - CP$. With $s = \beta v$, P’s threat to litigate is credible if and only if $CP \leq E$ or

$$CP \leq \theta \beta vX + \theta \beta [(m - v)XL + F].$$

**Disagreement Payoffs from the “Redesign” Strategy**

We now compute D’s payoff from following the “Redesign” strategy if the initial licensing negotiations with P fail. The “Redesign” strategy entails litigation, so D incurs its litigation cost $CD$. This strategy also involves redesign, so D incurs the redesign cost $F$. The payoff from the “Redesign” strategy differs only in a few terms from the payoff from the “Do Not Redesign” strategy just computed.

The only benefit that D enjoys from engaging in redesign immediately rather than waiting for the resolution of the patent litigation is the improved bargaining position D enjoys if P wins the litigation. In that event, which only arises with probability $\theta$, D saves $\beta (m - v)XL + \beta F$. The cost to D of improving its bargaining position is the redesign cost, $F$, which must be incurred before the outcome of the patent litigation is known. Therefore, D’s payoff from “Redesign” is equal to D’s payoff from “Do Not Redesign” plus $\theta \beta (m - v)XL + \theta \beta F - F$. So the payoff to D from the “Redesign” strategy equals

$$mX - \theta X[sT + \beta v(1 - T)] - F - CD.$$  \hspace{1cm} (11)

We write this as $mX - G - F - CD$, where $G = \theta X[sT + \beta v(1 - T)]$ is the expected payment from D to P. With $s = \beta v$, $G = \theta \beta vX$ and D’s disagreement payoff from “Redesign” equals $mX - \theta \beta vX - F - CD$. This strategy is better than exiting if $CD < mX - \theta \beta vX - F$. With $s = \beta v$, if “Redesign” is optimal for D, P’s threat to litigate is credible if and only if $CP \leq \theta \beta vX$. 


Proof of Lemma 1. As just explained, D will find it optimal to redesign its product immediately, rather than waiting for the outcome of the litigation, if and only if \(\theta \beta (m - v) XL + \theta \beta F > F\). Therefore, D’s optimal strategy is “Redesign” rather than “Do Not Redesign” if and only if

\[
\theta > \frac{1}{\beta (m-v) XL + F}.
\]  

Proof of Theorem 1. We now complete our analysis for the case in which D’s optimal threat point is to follow the “Do Not Redesign” strategy, as it will be for sufficiently weak patents. We showed above that the threat points in the initial negotiations when this is D’s optimal strategy are equal to \(mX - E - C_D\) for D and \(E - C_P\) for P.

Settlement allows the firms to save on litigation costs. Under Nash Bargaining, the two firms split these savings. Therefore, under Nash Bargaining, the initial negotiations give a payoff to P equal to its threat point, \(E - C_P\), plus its share, \(\beta\), of the gains from reaching agreement, \(C_P + C_D\). So P’s payoff from the initial licensing negotiations must equal \(E - C_P + \beta (C_P + C_D)\) or \(E + \beta C_D - (1 - \beta) C_P\). Since P receives no revenues other than the payment from D and incurs no costs, this expression must measure the total negotiated payment from D to P, which equals \(\pi^* \equiv r^*X\). Substituting for \(E\), using \(E = \theta X[sT + \beta v(1 + T)] + \theta \beta (m-v) XL + \theta \beta F\), the equilibrium payoff to the patent holder in this case is given by

\[
\pi^* \equiv r^*X = \theta [sT + \beta v(1-T)]X + \theta \beta (m-v) XL + \theta \beta F + [\beta C_D - (1 - \beta) C_P].
\]  

Substituting \(s = \beta v\) gives the expression for \(\pi^*\) in Theorem #1.

Proof of Theorem 2. We now complete our analysis for the case in which D’s optimal threat point is to follow the “Redesign” strategy. We showed above that the threat points in the initial negotiations when this is D’s optimal strategy are equal to \(mX - F - G - C_D\) for D and \(G - C_P\) for P.

Settlement allows the firms to save on litigation and redesign costs; under Nash Bargaining, the two firms split these savings. Therefore, under Nash Bargaining, the initial negotiations give a payoff to P equal to its threat
point, \( G - C_P \) plus its share, \( \beta \), of the gains from reaching agreement, \( C_D + C_P + F \). So P’s payoff from the initial licensing negotiations must equal \( G - C_P + \beta(C_P + C_D + F) \) or \( G + \beta F + \beta C_D - (1 - \beta)C_P \). As in the previous case, since P receives no revenues other than the payment from D and incurs no costs, this expression must measure the total negotiated payment from D to P, \( \pi^* \equiv r^* X \).

Substituting for \( G \) using \( G = \theta X(sT + \beta v(1 + T)) \), the equilibrium payoff to the patent holder in this case equals

\[
\pi^* \equiv r^* X = \theta[sT + \beta v(1-T)]X + \beta F + [\beta C_D - (1-\beta)C_P].
\]

Substituting \( s = \beta v \) gives the expression for \( \pi^* \) in Theorem #2.

**Numerical Examples: “Do Not Redesign” vs. “Redesign”**

Lemma #1 establishes that D does better adopting the “Do Not Redesign” strategy rather than the “Redesign” strategy if and only if \( \theta \beta < \frac{F}{(m-v)XL} \), which can be written as \( \theta \beta < \frac{F}{XL + F/vX} \). In our numerical example in which “Do Not Redesign” is optimal for D, we have \( F/vX = 0.5 \), \( (m - v)/v = 10 \), and \( L = 0.1 \), so this inequality becomes \( \theta \beta < \frac{0.5}{1 + 0.5} = 1/3 \). With equal bargaining skill, \( \beta = 1/2 \), this condition in turn becomes \( \theta < 2/3 \). In our numerical example in which “Redesign” is optimal for D, we have \( F/vX = 0.5 \), \( (m - v)/v = 30 \), and \( L = 0.1 \), so this inequality becomes \( \theta \beta < \frac{0.5}{3 + 0.5} = \frac{1}{7} \). With equal bargaining skill, \( \beta = 1/2 \), this condition in turn becomes \( \theta < 2/7 \).

**Proof of Theorem 3.** Working backward as usual, we ask what payoffs result if P wins but the injunction is stayed while D redesigns its product. In this situation, D’s prospective payoff after P wins, and if no license agreement is then reached, is given by \((m - \beta v)XL + (m - v)X(1 - T - L) - F\). P’s prospective payoff if no license is signed is \( \beta v XL \). The gains from reaching agreement are \( v X(1 - T - L) + F \), which represents the use of the patented technology after the injunction would go into force plus D’s redesign costs. D’s prospective payoff under the negotiated license is therefore equal to \((m - \beta v)XL + (m - v)X(1 - T - L) - F + (1 - \beta)(v X(1 - T - L) + F)\) which simplifies to \((m - \beta v)X(1 - T) - \beta F\).
D’s overall payoff from the “Do Not Redesign” strategy thus equals 
\[(m - \theta \beta v)X(1 - T) - \theta \beta v X(1 - T) + \beta L - C_D.\]
Simplifying, this expression becomes \[mX - \theta \beta v X - \theta \beta F - C_D.\]
Under the initial licensing agreement, D gets this payoff plus its share, \(1 - \beta\), of the gains from trade, \(C_D + C_P\), or \[mX - \theta \beta v X - \theta \beta F - [\beta C_D - (1 - \beta)C_P].\]
With \(\beta C_D - (1 - \beta)C_P = 0\), this implies that \[\pi^* = r^* X = \theta \beta v X + \theta \beta F.\]

**Proof of Theorem 5.**

**“Do Not Redesign” Optimal for D**

If \(\theta < \theta^*\) so “Do Not Redesign” is D’s optimal threat point, the patent holder’s payoff is given by Equation (13). Assuming \(\beta C_D - (1 - \beta)C_P = 0\) and putting \(\theta = 1\) into that equation gives \[r^*(1) = sT + \beta v (1 - T) + \beta (m - v) L + \beta F/X.\]
The condition defining a self-fulfilling equilibrium is \(s = r^*(1)\). Using this relationship, we can eliminate \(r^*(1)\) from the previous equation to get \[s = sT + \beta v (1 - T) + \beta (m - v) L + \beta F/X.\]
Solving for \(s\) gives 
\[s = \beta v + \frac{\beta}{1 - T} [(m - v) L + \frac{F}{X}].\]

This expression tells us that the reasonable royalty rate will itself be greater than the benchmark level of \(\beta v\) because it is influenced by the threat of hold-up, as reflected by the term in brackets on the right-hand side.

This expression for \(s\) cannot hold if \(\theta s > v\), in which case D would choose not to continue to sell its product during the pendency of litigation. In addition, for this expression for \(s\) to hold, D must prefer redesigning its product to exiting the market after losing the litigation. If, instead, D’s threat point after losing the patent litigation is simply to withdraw from the market, then \(s = \beta m\).

If the expression above for \(s\) does apply, we can use Equation (13) to derive a new equation for the payoff to the patent holder:
\[\pi^* = r^* X = \theta \beta v X + \frac{\theta \beta}{1 - T} [(m - v) L + F].\]
When this equation applies, the hold-up term in the fulfilled expectations equilibrium is magnified by the factor \(1/(1 - T)\), in comparison with the case where \(s = \beta v\).
“Redesign” Optimal for D

If $\theta > \theta^*$ so “Redesign” is D’s optimal threat point, the patent holder’s payoff is given by Equation (14). Substituting using $s = r^*(1)$ and solving for $s$, we get

$$s = \beta v + \frac{\beta F}{X(1-T)}.$$  

As in the previous sub-section, this equation cannot apply as $T$ approaches unity because it would violate our standing assumption that $\theta s \leq v$. It also cannot apply if it implies that $s > \beta m$. So long as $T$ is not too large, however, this expression is valid and leads to a revised expression for the patent holder’s payoff:

$$\frac{\pi^* - \pi}{\pi} = \frac{1}{1-T} \frac{F}{vX} \frac{1}{\theta}.$$  

Again, the hold-up term is magnified by the factor $1/(1-T)$.

Multiple Rounds of Bargaining

The text modeled bargaining as taking place during two discrete episodes, one prior to litigation and, if necessary, another after the litigation is resolved. We now explain how negotiated royalties are affected if bargaining is instead modeled as an ongoing process.\(^3^4\)

The ability of the parties to return to the bargaining table can change the negotiated outcome by affecting the parties’ threat points. Binmore et al. (1989) distinguish between bargaining “breakdown” and bargaining “impasse.” Breakdown refers to the situation where at least one party walks away from the bargaining table and invokes its outside option. In contrast, impasse refers to the situation where no agreement is reached but both parties continue to negotiate, with neither invoking its outside option. The notion of impasse arises in a game with multiple rounds of bargaining where at least one party must take some

\(^3^4\) We continue to assume that the two parties negotiate over a long-term license that will cover the remainder of the patent lifetime. Alternatively, one could assume that if the parties fail to negotiate a long-term license in any given round of bargaining, they can still sign a short-term license covering the time period until they next sit down to bargain. Allowing the parties to reach such short-term agreements would not change our results here.
irreversible action to invoke its outside option. Binmore et al. argue that a threat to walk away from the bargaining table is not credible and will not affect the negotiated outcome if invoking it would give the party walking away less than it could get by continuing to bargain without that threat.

“Redesign”

Consider the bargaining that takes place after D has redesigned its product and P has won the patent litigation. In this subgame, neither party is threatening to take an irreversible step if the licensing negotiations fail. P earns a flow payoff of zero and has no outside option to invoke. D sells the non-infringing version of its products, but this does not require that D take any irreversible step. The negotiated royalty of $\beta v$ would be unchanged if the parties could continue to negotiate after an initial impasse.

“Do Not Redesign”

The analysis is more complex in the post-litigation subgame where D has not redesigned its product. If negotiations fail, again P earns a flow payoff or zero and has no outside option to invoke. But now D does have an outside option: to redesign its product. The text assumed that D would invoke this option, giving D its breakdown payoff of $(m - v)X(1 - T - L) - F > 0$. In contrast, D’s impasse payoff is zero: D earns no profits while continuing to bargain with the injunction in force. Since D’s impasse payoff is less than its breakdown payoff and P’s breakdown and impasse payoffs are the same, D’s payoff must be no larger under the Binmore et al. approach than under the conventional Nash bargaining approach.

We can identify the conditions under which ongoing bargaining gives a lower payoff to D. If $F > (m - v)X(1 - T - L)$, D would exit rather than redesign its product, so redesign cannot be credible, whether or not bargaining is ongoing, and the negotiated royalty rate is $\beta m$. If $F < \beta(m - v)X(1 - T - L)$, D saves enough money by redesigning to make the threat to redesign its product credible, even with multiple rounds of bargaining. In that case, the negotiated royalty is the same as in the text. If $\beta(m - v)X(1 - T - L) < F < (m - v)X(1 - T - L)$, then D’s threat to redesign is not credible with multiple rounds of bargaining. In that case, D’s payoff from redesign
still serves as a constraint on the payoff that D will accept. Therefore, D’s prospective payoff equals its payoff from redesign, namely \((m - v)X(1 - T - L) - F\). This implies that P’s payoff equals \(vX(1 - T) + [(m - v)XL + F]\). This is the same as the payoff shown in Equation (8), namely \(\beta vX(1 - T) + \beta(m - v)XL + F\), except that the entire expression is not discounted by \(\beta\).

References


