PRL
7/25/06
2:30 PM

How Strong Are Weak Patents?*

Joseph Farrell and Carl Shapiro†

October 2005

ABSTRACT. We analyze the licensing of patents that may be invalid, to licensees who compete in a downstream product market. If licenses involve two-part tariffs, the patentee and licensees will often – especially for weak patents – agree on a running royalty equal to the full cost savings from the patented technology. Patent weakness is reflected only in a negative fixed fee, which benefits downstream firms but not final consumers. Stronger patents are licensed with higher fixed fees and *lower* running royalties. Consumers would be better off if the patent were litigated and then licensed if upheld. We use these results to study two public policy questions: (1) Should negative fixed fees in patent licensing agreements be prohibited? and (2) Are there large benefits from reforming the patent system to improve patent quality and reduce the number of weak patents that are issued? We show that banning negative fixed fees benefits consumers and raises welfare in the short run. Taking account of innovation incentives, we also show that the optimal patent system prohibits negative fixed fees in patent licenses and that long-run welfare is improved by such a ban if patent lifetimes are optimally set. In the absence of negative fixed fees, running royalties qualitatively reflect patent strength, but for small, weak patents the running royalty still exceeds a natural normative benchmark, namely the cost savings from the patented technology times the probability the patent is valid. Indeed, if the downstream industry is a symmetric Cournot oligopoly with *N* firms with constant marginal costs, we show in a central case that the running royalty for a weak patent is *N* times this benchmark. Our analysis thus supports calls for patent reform: even if negative fixed fees are banned, weak patents have surprisingly large commercial effects, so there are significant benefits of improving patent quality.

---

* The authors thank participants at the Berkeley/Stanford I.O. Fest, the Berkeley I.O. Seminar, Northwestern University, and at the Conference in Tribute to Jean-Jacques Laffont for helpful suggestions. This paper and subsequent revisions will be available at Shapiro's web site: faculty.haas.berkeley.edu/shapiro.

† University of California at Berkeley. Contacts: farrell@econ.berkeley.edu and shapiro@haas.berkeley.edu.

1. Introduction

Economists usually view a patent as a property right that prevents others from practicing the patented technology without a license. In practice, however, U.S. patents are issued after only brief examination by the U.S. Patent and Trademark Office (PTO), and thus do not always meet statutory requirements for novelty and non-obviousness. Patents vary greatly in their strength, i.e., the probability that they will be found valid if fully tested through litigation.[1] Mounting evidence indicates that the patent office issues many "weak" or questionable patents.

Some observers argue that these weak patents constitute undeserved monopolies, and that it is crucial to reform the process by which patents are issued to improve patent quality and reduce the number of weak patents. Others respond that the threat of infringement or challenge ensures that, if enforced at all, weak patents are licensed at commensurately low royalty rates.[2]

To address this debate, which is central to patent reform efforts currently underway in Congress, we analyze patent licensing in the shadow of litigation over patent validity. We consider the common case where a patented technology is useful for multiple "downstream firms" that compete against one another. If a downstream firm rejects a license and challenges or infringes the patent, we assume this triggers patent litigation that determines whether or not the patent is valid. If a court finds the patent *invalid*, the challenger can freely use the technology; but, under Supreme Court precedent, so can its rivals.[3] Therefore, a successful challenge confers positive externalities on rivals and on consumers. However, if the patent is found *valid*, the unsuccessful challenger must either negotiate a license with the now stronger patent holder, or use the backstop technology (as the licensing literature has generally assumed in the first place).

---

[1] Many studies and articles support this view, including Federal Trade Commission (2003) and National Academies of Science (2004). For a recent overview, see Mark Lemley and Carl Shapiro (2005).

[2] See especially Mark Lemley (2001).

[3] The U.S. Supreme Court has ruled that if one challenger to a patent prevails on patent invalidity, this result can be relied upon by other users, who therefore need not pay royalties for that patent, even if they had previously agreed to do so in a license. See *Blonder-Tongue Labs, Inc. v. University of Illinois Foundation*, 402 U.S. 313, 350 (1971).

In our model, we show that a patentee can and will structure its licensing terms so that downstream firms will accept licenses rather than challenge the patent. This prediction fits well with the empirical evidence that very few patents are litigated to a final judgment. The key question is, of course, what royalties the licensees will pay. Strikingly, we find that weaker patents will tend to command *higher* running royalties (along with lower fixed fees). Since running royalties constitute marginal costs for downstream firms, weak patents have strong effects on downstream prices and on final consumers. Our analysis therefore supports the view that there are substantial benefits of reforming the patent system to improve patent quality.

Product-market competition among licensees drives price below the monopoly level in our model, so total profits are increasing in the downstream price. Running royalties elevate that price and thus increase total profits. Fixed fees are used to divide the resulting rents so the downstream firms will accept licenses. For weak patents, we show that the per-unit royalty will equal the full cost savings from the patented technology. To induce downstream firms to pay these per-unit royalties for a weak patent, the patent holder makes a lump-sum payment to each licensee. Patent weakness is reflected only in those negative fixed fees, which benefit downstream firms but not final consumers.

The Federal Trade Commission has brought several recent cases alleging that such "reverse payments," from patent holders to alleged infringers, violate antitrust laws, and some similar private cases have gone to court. In each of these cases, patent litigation between pharmaceutical patent-holders and potential generic entrants was settled on terms involving lump-sum payments from the patent holders to their would-be generic rivals. Courts are still grappling with the question of how to treat such payments; recently, the Eleventh Circuit reversed the Federal Trade Commission and found such payments by Schering-Plough to be legal.[4]

We study the economic impact of prohibiting the use of negative fixed fees in patent licenses. Our analysis strongly suggests that the *Schering* decision is not in the public interest. We show that the use of negative fixed fees increases deadweight loss and harms consumers *ex post* (given

---

[4] See the Decision by the Eleventh Circuit Court of Appeals in Schering-Plough Corp. and Upsher-Smith Laboratories vs. Federal Trade Commission, Case No. 04-10688, March 8, 2005. For discussion of the cases and antitrust analysis of such settlements, see Hovenkamp, Janis, and Lemley (2003) and Shapiro (2003a,b).

that the patented innovation exists). Although the option to use such fees increases the return to patents, we show that the *ex ante* optimal patent and antitrust system prohibits negative fixed fees. We argue that such a prohibition increases overall welfare if patent lifetimes are set optimally.

While negative fixed fees exacerbate the problems associated with weak patents, we show that, even without them, weak patents can still command royalties far out of proportion to their strength when licensed to downstream firms that compete vigorously with one another. As a result, even if antitrust rules prohibit negative fixed fees, there are still substantial benefits of improving patent quality.

Our analysis departs from the standard assumption in the patent licensing literature that patents are surely valid. We extend the small literature on the licensing of probabilistic patents. Meurer (1989) and Choi (1998) explore the licensing of probabilistic patents when the patent holder has private information about patent strength. They focus on signaling and information transmission, issues that do not arise in our framework, in which patent strength is common knowledge. Meurer also briefly considers the licensing of a patent for which patent strength is common knowledge. In a model with a patent holder and a single rival, he shows that settlement may not be possible if antitrust rules require any settlement to enable strong competition between the patent holder and the licensee. Joint profits are higher under litigation because it may lead to a monopoly (since the patented technology is assumed to be essential, again in contrast to our model). He does not, however, study the structure of licensing agreements, their impact on consumers, or interactions among multiple licensees, which are central to our analysis.[5]

Our paper is organized as follows. Section 2 introduces our model of patent licensing and litigation. Section 3 characterizes the equilibrium two-part tariff licenses. Section 4 proves that consumers are harmed by licensing using two-part tariffs including negative fixed fees. Section 5 shows that banning negative fixed fees promotes consumer welfare and reduces deadweight loss

---

[5] Choi (2002) studies patent pools with probabilistic patents, focusing on the incentives of one patentee to challenge another's patent rather than form a patent pool. Anton and Yao (2003) study how uncertainty about patent validity can encourage or discourage disclosure. Waterson (1990) discusses how uncertainty about patent infringement (patent scope) affects rivals' design decisions.

ex post. Section 6 complements this with an *ex ante* analysis, showing that the optimal patent and antitrust system prohibits negative fixed fees and that, if patent lifetime is roughly optimal, banning negative fixed fees improves welfare *ex ante*. Section 7 shows that weak patents can elevate prices to a surprising extent even if negative fixed fees are banned. Section 8 relaxes our standing assumption that per-unit royalties cannot exceed the cost savings associated with the patented technology, and shows that running royalties for weak patents are even higher and banning negative fixed fees is even more desirable in this case. Section 9 shows that a patent holder would voluntarily design its licensing agreements so as to relieve competing licensees of royalty obligations following an invalidity finding, even if this were not required under patent law. Section 10 summarizes our findings and discusses their policy implications.

2. Basic Model and Assumptions

A. Extensive-Form Game

We study licensing by a single patent holder to *N* symmetric downstream firms $i = 1, ..., N$.[6] The patented technology lets a downstream firm lower its unit costs by *v*, the *patent size*, in comparison with the best alternative, or *backstop*, technology.

1. The patent holder offers a single two-part tariff patent licensing offer (F, r) to all of the downstream firms.[7] The patent holder therefore has take-it-or-leave-it bargaining power.

[6] In future work, we plan to study the case in which the patent holder is vertically integrated and competes downstream against its licensees.

[7] In the Appendix, we provide sufficient conditions for this to be an equilibrium. We assume that an agreement signed between the patent holder and one licensee is observable to other licensees. McAfee and Schwartz (1994) show that an input supplier would indeed choose to commit to using observable contracts if possible. They show that the equilibrium with observable contracts obtains with unobservable contracts if each downstream firm believes that other firms will receive the same offer it receives, which McAfee and Schwartz call "symmetry beliefs." However, with unobservable contracts and "passive beliefs," under which a given firm does not adjust its beliefs about the contracts offered to other firms based on the offer it receives, McAfee and Schwartz show that the only equilibrium involves input prices at marginal cost, which translates to zero running royalties in our context. Segal and Whinston (2003) allow the input supplier to offer more complex menu contracts and prove a more general competitive convergence result for such "offer games." These results predict that with unobservable licensing agreements a patent will be licensed for a fixed fee with no running royalties. Reconciling these results with the empirical observation that per-unit royalties are frequently used is a challenge for future work on patent licensing. We simply assume that patent holders can commit to observable contracts, or that they can use most-favored customer clauses to obtain the same result.

2. Each downstream firm can: (a) accept the tariff that the patent holder has offered; (b) infringe and challenge the patent; or (c) avoid infringing by using the backstop technology.[8] The downstream firms make these choices simultaneously. Each firm's decision is then observed by the others.

3. If all downstream firms choose to accept a license or to use the backstop technology, then the game is effectively over. Each firm that accepts uses the patented technology but pays a fixed fee of F and a per-unit royalty of r to the patent holder. Each firm that uses the backstop technology incurs the costs associated with that technology. These firms, with their resulting costs, compete as downstream oligopolists. The resulting payoffs are discussed below.

4. If, instead, any downstream firm infringes and challenges the patent, then litigation ensues, leading to a court finding that the patent is valid or invalid.[9]

   A. The probability θ that the patent will be ruled valid, which we call the *patent strength*, is exogenous and is common knowledge.

   B. If the patent is ruled invalid, then all downstream firms can use the (formerly) patented technology free of charge: existing licenses are voided and licensees have no obligation to pay royalties to the patent holder. Downstream competition ensues.

   C. If the patent is ruled valid, existing licenses remain in force.[10] The patent holder offers a new two-part tariff to each firm without such a license. Each such firm then chooses to accept that offer or use the backstop technology. Downstream competition ensues.

---

[8] We assume that a downstream firm cannot sign a license and then challenge the patent. This is consistent with patent law, under which a licensee who is making royalty payments and thus not incurring potential liability for infringement lacks standing to challenge the patent. See *Medimmune, Inc. v. Centocor, Inc.*, Court of Appeals for the Federal Circuit, June 1, 2005.

[9] Either (a) the downstream firm challenges the patent, seeking a declaratory judgment that the patent is invalid, as in Meurer (1989), or (b) the downstream firm simply uses the technology and the patent holder brings an infringement case. In the latter case, which is more common in practice, we do not explicitly study how the patent holder ensures that its threat to litigate is credible. One such mechanism is a provision that licensees are relieved of their obligation to pay royalties if the patent holder fails to use "best efforts" to pursue their rival who infringes the patent. More generally, litigation is credible if the patent holder expects its income from licensees to plummet if it ignores an infringer. This would occur if the infringing firm, with lower costs, captured significant market share from those who are paying royalties, or if failure to pursue one downstream firm would cause other downstream firms to cease making royalty payments . In separate work, we study strategies for ensuring litigation credibility.

[10] In practice, licensees typically make specific investments to use the patented technology and thus sign long-term licenses that are not subject to renegotiation in the event the patent is challenged and upheld. We restrict attention to two-part tariffs and do not explore more general contingent licenses, including licenses that specify adjusted royalty terms in the event the patent is challenged and upheld. In the extreme case, the patent holder could agree to subsidize all of its licensees if any other firm challenges its patent and loses. Such contingent licenses could (unrealistically) deter challenges at no (equilibrium) cost to the patent holder.

B. Reduced-Form Downstream Payoff Functions

The downstream oligopoly equilibrium depends on the N downstream firms' marginal costs; each firm's marginal cost depends on the technology it uses and any running royalties it pays. For ease of notation, we measure each firm's marginal cost relative to the marginal cost that it would incur if it could use the patented technology free of charge. In this notation, a firm that accepts a two-part tariff (F, r) has marginal cost r, and a firm that uses the backstop technology has marginal cost v.

We assume that, for any vector of marginal costs $\boldsymbol{r} = (r_1, r_2, ..., r_N)$ for the downstream firms, there is a unique downstream oligopoly equilibrium. We denote the equilibrium profits of firm i, net of running royalties but *not* net of any fixed fee, by $\pi^i(\boldsymbol{r})$. We assume that the N downstream firms are symmetric.[11]

Because we look for symmetric equilibrium, much of our analysis studies the profits of one firm when all other firms have equal costs. Therefore, simplifying notation, we denote by $\pi(a,b)$ the profits of a firm whose marginal cost is a when the marginal costs of all other firms are b. This firm's output level is denoted by $x(a,b)$. Per-firm profits if all firms have marginal cost c are $\pi(c,c)$; thus, if all N downstream firms accept the two-part tariff (F, r), as occurs in the equilibria we consider, each earns $\pi(r,r)$. All of these profit functions are gross of any fixed fees. If all its rivals are paying royalties of r, a firm that uses the backstop technology earns $\pi(v,r)$. If the patent is found invalid, all downstream firms can freely use the technology, so each earns $\pi(0,0)$.

We assume that the profit function $\pi(a,b)$ is "normal" in the following three senses: (1) $\pi_1(a,b) < 0$: a firm's profits are decreasing in its own costs; (2) $\pi_2(a,b) > 0$: a firm's profits are

[11] Fixing the number of competitors is, of course, a standard assumption in oligopoly theory. We flag it in part because a patent holder who structures licensing arrangements to cartelize the downstream industry may be constrained by the threat of entry, even if this threat would not otherwise be binding in the downstream oligopoly. As we shall see, the patent holder may well make lump-sum payments to downstream firms, i.e., use negative fixed fees, in conjunction with per-unit royalties that equal or exceed the cost savings from the innovation. Such licensing

increasing in its rivals' costs; and (3) $\pi_1(r,r)+\pi_2(r,r)\le 0$: each firm's profits fall if all firms' costs rise in parallel.[12] As a limiting case, we permit total downstream profits to be unchanged in r over some range.

3. Equilibrium Two-Part Tariff Licensing Agreements

We now solve for the optimal two-part tariff for the patent holder to offer to the downstream firms, assuming that the patent holder chooses a single tariff that will be accepted by all of them. In the Appendix, we show that the patent holder will indeed offer such a tariff, rather than inducing any firm to use the backstop technology, entering into asymmetric licensing agreements, or licensing to some firms and litigating with others. We use the term "licensing" for the outcome we study here, distinguishing it from "litigation," although litigation leading to a finding of validity will itself be followed by licensing of the (now ironclad) patent.

A. Licenses Must Remain Within the Scope of the Patent

A core principle of patent law is that a patent holder can only impose royalties for the use of the patented technology. If this rule is effectively enforced, the running royalty cannot exceed the patent size: $r \le v$. A downstream firm operating with a license with $r > v$ would find the backstop technology cheaper than the patented technology (since royalties cannot be imposed on production using the backstop technology), making the running royalty rate inoperative and irrelevant. Therefore, for most of our analysis we treat $r \le v$ as a constraint. However, joint profits are often higher if the patent holder can impose royalties on all of the downstream firm's production, not just production using the patented technology, in which case the $r \le v$ constraint no longer applies. So, neither the patent holder nor the downstream firms may have the incentive to enforce the rule limiting royalties to production that uses the patented technology. In Section 8 we explain how our analysis proceeds without the $r \le v$ constraint.

---

agreements encourage entry by raising incumbents' private marginal costs, especially if entrants expect to receive their own lump-sum payments from the patent holder.

[12] This condition depends on the form of oligopolistic interaction and the shapes of the cost and demand functions. Kimmel (1992) provides sufficient conditions for this to hold in Cournot oligopoly. See also Shapiro (1989).

B. Infringe and Challenge is the Binding Constraint

We next show that each downstream firm will be more tempted to infringe, challenging the patent, than to use the backstop technology.

If the patent is challenged and found invalid, then each downstream firm's payoff is $\pi(0,0)$. If the patent is found valid, the patent holder, with a now ironclad (certainly valid) patent, will make a new offer to the unsuccessful challenger. In contemplating a unilateral deviation to infringe, the potential challenger will expect that, if the challenge fails, its rivals will be paying running royalties of *r,* so its *ex post* reservation payoff will be $\pi(v,r)$ and the patent holder will hold it down to this payoff. Thus, the expected payoff from challenging the patent is:

$$\theta\pi(v,r)+(1-\theta)\pi(0,0).$$

By comparison, if the downstream firm uses the backstop technology, it earns $\pi(v,r)$. Since $r \le v$, normality of the profit function implies that $\pi(0,0) > \pi(v,v) \ge \pi(v,r)$. Therefore, challenging the patent is more attractive than using the backstop technology provided $\theta < 1$. Intuitively, the firm challenging the patent has a chance, θ, of ending up with the backstop payoff but also a chance, $1-\theta$, of invalidating the patent and earning the higher payoff $\pi(0,0)$. Thus, in evaluating the firm's willingness to pay for a license prior to litigation, we can ignore the backstop constraint and focus on the infringe/challenge constraint.

C. The Optimal Two-Part Tariff

The patent holder would always raise *F* if downstream firms would still accept licenses. In equilibrium, therefore, each downstream firm is indifferent between accepting the license and its next best alternative, which we just showed is infringing and challenging the patent. Its payoff from accepting the license, $\pi(r,r)-F$, must equal its expected payoff from challenging:

$$\pi(r,r)-F=\theta\pi(v,r)+(1-\theta)\pi(0,0).$$

Solving for *F* gives

$$F=\pi(r,r)-\theta\pi(v,r)-(1-\theta)\pi(0,0).$$

Substituting for *F*, the patent holder's total profit from each downstream firm is

$$G(r,\theta) \equiv rx(r,r) + F = rx(r,r) + \pi(r,r) - \theta\pi(v,r) - (1-\theta)\pi(0,0).$$

Define the total (upstream plus downstream) profits per downstream firm if all pay a running royalty of *r* as $T(r) \equiv rx(r,r) + \pi(r,r)$. We assume that $T(r)$ is quasi-concave in *r* and that $G(r,\theta)$ is single-peaked in *r*. Writing *m* for the running royalty that maximizes total profits $T(r)$ by supporting the full cartel outcome downstream, why would *r* not be set at *m*? There are two reasons. First, as described above, we have the constraint $r \le v$, and for a non-drastic innovation $v < m$, which we assume.[13] Second, the patent-holder's profit is not $T(r)$ (even up to a constant) but

$$G(r,\theta) = T(r) - \theta\pi(v,r) - (1-\theta)\pi(0,0).$$

$G(r,\theta)$ is maximized at a running royalty $r(\theta)$ that is below *m* and is *decreasing* in θ. Recalling that $r(\theta)$ is defined without considering the constraint $r \le v$, the patent holder sets $r(\theta) < m$ because it is more profitable to lower each licensee's reservation payoff by keeping its rivals' marginal costs *r* low. But this rent-shifting effect (Segal 1999) only arises through the probability that a challenger would actually face rivals with marginal cost *r*, which only happens if the patent is found valid, so its strength is proportional to θ, and $r(\theta) \to m > v$ as $\theta \to 0$. Overall (all proofs are in the Appendix):

Theorem 1: The running royalty is $r = \min[v, r(\theta)]$, which is decreasing in patent strength θ, and is equal to *v* for all sufficiently small θ.

Even – in fact, especially – a flimsy patent commands a running royalty equal to the full value-in-use of the patented technology. Perversely, because the rent-shifting effect is stronger, an ironclad patent may well be licensed at a lower running royalty (though with a higher return to

[13] Put differently, we are assuming that joint profits increase in the downstream price/cost margin, starting at the pre-patent margin. We are relying here on the presence of downstream competition. If licensees do not compete against each other, $m = 0$ (to avoid double marginalization). Therefore, with a single downstream firm, or with downstream firms that do not compete, running royalties will not be used, regardless of patent strength.

the patent-holder) than a weak patent. Define $r_I \equiv r(1)$ as the optimal running royalty for an ironclad patent if one ignores the constraint $r \le v$. If $r_I < v$, then Figure 1 applies. In this case, writing θ_V for the value of θ at which $r(\theta) = v$, we have $r = v$ for all $\theta \le \theta_V$. In contrast, if $r_I \ge v$, then Figure 2 applies, and $r = v$ for all θ. Each of these cases arises in reasonable models, as we show in the Appendix. With Cournot oligopoly downstream, $r_I < v$, so Figure 1 applies. With differentiated-product Bertrand oligopoly and perfectly inelastic total demand up to a high enough choke price, $r_I > v$, so Figure 2 applies.

If $r = v$, as is the case if $r_I \ge v$ and always for weak patents, consumers do not benefit from the patented technology: the benefits of innovations covered by weak patents are shared by the patent holder and the downstream firms. Consumers do benefit if $r < v$, as is the case for sufficiently *strong* patents if $r_I < v$. In that case, downstream firms are harmed by the patented technology, since as $\theta \to 1$ their reservation payoff approaches $\pi(v, r)$, which is less than $\pi(v, v)$, their payoff in the absence of the patented technology.

D. The Use of Negative Fixed Fees

For sufficiently weak patents, negative fixed fees will be used: from the expression above for *F*, if $r = v$, then $F = -(1-\theta)[\pi(0,0) - \pi(v,v)] < 0$. For weak patents, the patent holder sets $r = v$ and pays each downstream firm enough to dissuade it from challenging. For very weak patents, i.e., as $\theta \to 0$, any downstream firm could surely invalidate the patent and get $\pi(0,0)$, so the fixed fee must equal $F(0) = \pi(v,v) - \pi(0,0) < 0$. More generally:

Theorem 2. If allowed, negative fixed fees will be used for sufficiently weak patents. The fixed fee is an increasing function of patent strength.

Empirically, running royalties are common, perhaps suggesting that competition among licensees is a significant force in license design, although there are other explanations for running royalties.

E. The Patent Holder Prefers Licensing to Litigation

We now show that in a broad class of cases the patent holder prefers signing these licenses to litigating. Write $H(\theta)$ for the patent holder's maximized payoff from licensing, as above. If the patent is upheld in litigation, the patent holder's payoff is $H(1)$. If the patent is found invalid, the patent holder gets zero. Therefore, the patent holder prefers licensing to litigation if and only if $H(\theta) > \theta H(1)$. In the Appendix, we prove:

Theorem 3. The patent holder prefers licensing to litigating for weak patents, and for all patent strengths if the running royalty for an ironclad patent would be positive.

4. Patent Licensing Harms Consumers

We now ask how these patent settlements impact consumers. Specifically, we compare consumer welfare under licensing in the shadow of litigation versus consumer welfare from litigation that will be followed by licensing if the patent is upheld.

Theorem 4: Licensing that averts or settles litigation makes downstream firms better off, but consumers worse off, than if the patent is litigated to judgment and then licensed if upheld.

Effectively, the patent holder shares with the downstream firms the extra profits that can be achieved through settlement. Put this way, the result is not surprising, since settlements generally are compromises. However, these profits come at the expense of consumers and cause greater deadweight loss (higher running royalties) than would arise (on average) from litigation. The motive for settlement here has nothing to do with litigation costs (which we assume away) and everything to do with keeping consumer prices high.

Theorem 4 tells us that the patent settlements studied above, which avert litigation over patent validity, harm consumers, at least when patent holders can use negative fixed fees. This finding is in sharp contrast to the common view in the courts that settlements are to be encouraged, including by allowing the use of such fees. To the contrary, we find that patent settlements impose a negative externality on consumers and are thus to be *discouraged*, relative to the private incentive to achieve them.

This externality formulation does not depend on the fact that our model omits litigation costs. Including large enough litigation costs would reverse the model's finding that litigation is more efficient than settlement, but would leave intact our prediction that testing of weak patents is *under*-supplied by private parties (assuming that they bear the litigation costs). There is a real social gain to finding out which patents are valid. This need not mean more litigation: alternatives could include more extensive PTO examination, or other means such as administrative opposition procedures. But voluntary litigation between the patent holder and competing downstream licensees to test weak patents is inadequate.[14]

In fact, consumers often benefit from greater litigation, even if final judgment is not reached:

Theorem 5: If the running royalty rate is concave in patent strength over the range $[\theta_V, 1]$ and if the downstream price is a linear function of the running royalty rate, then consumers are better off in expected value as a result of any news about patent strength; in particular they are better off on average the farther litigation progresses prior to settlement.[15]

Intuitively, Theorem 4 holds because the running royalty is concave in patent strength θ on the set $\{0, \theta, 1\}$ of patent strengths that can arise when the question is litigation to judgment versus none. Figures 1 and 2 show that the running royalty (including the zero that prevails once a patent is proven invalid) is broadly concave, suggesting a more general result; Theorem 5 makes this precise by observing that the running royalty rate is globally concave in patent strength θ if it is concave on $[\theta_V, 1]$. If $r_I \geq v$ then $\theta_V = 1$ and this concavity condition is surely met.

5. Treatment of Negative Fixed Fees

We have seen how the patent holder can use negative fixed fees to induce downstream firms to agree to running royalties of $r = v$ even for very weak patents. This suggests that banning the

---

[14] Miller (2004) discusses alternative ways to encourage the testing of weak patents.

[15] Although Theorem 5 ensures that partial information benefits consumers, Figures 1 and 2 suggest that a good deal of the consumer benefit arises through the chance that the patent is actually invalidated: the running royalty has a discontinuity at zero. In particular, settlements in the course of a trial that is going badly for the patent holder are especially harmful, and embarking on litigation is less beneficial to consumers if such settlement is likely.

use of negative fixed fees might be an attractive policy. As noted above, the FTC has argued for such a prohibition, but it was recently reversed by the Eleventh Circuit Court of Appeals.[16]

The effect of banning negative fixed fees depends, naturally, on how the patent holder and the downstream firms respond to such a ban. We explore two possibilities in turn: (1) the outcome will involve litigation; or (2) the parties will negotiate licenses with $F \geq 0$.

We find that banning fixed fees raises *ex post* welfare, benefits consumers, harms downstream firms, and reduces the payoff to the patent holder. These effects are qualitatively the same under the two approaches, but more pronounced if litigation is the alternative. Thus, the possibility that banning negative fixed fees would lead to litigation is a benefit, not a downside. Again, this result would be more nuanced if the model included litigation costs, but the fact that consumers benefit in expected value from a patent validity determination, such as litigation pursued to judgment, would survive.

Our findings thus contradict the view expressed by some courts that settlement is not only desirable but a key policy goal. Perhaps this view is an incorrect extrapolation from the case of private disputes whose settlement involves no significant externalities.

A. Effect of Banning Negative Fixed Fees: Litigation Alternative

We first consider the effect of banning negative fixed fees if such a ban would cause the parties to litigate, not sign different licenses. One interpretation of this approach is that litigation-averting licensing agreements might really not be reached if their terms are restricted. Another interpretation is that, whatever the counterfactual, one should evaluate whether the license agreement was anticompetitive, relative to no such agreement, even if some other, perhaps more pro-consumer, agreement was feasible. This approach is suggested by the FTC/DOJ *Antitrust Guidelines for the Licensing of Intellectual Property*, under which the antitrust agencies ask

[16] The court apparently believes that no harm to competition arises if weak patents are licensed as if they were ironclad, and that settlements are automatically preferable to litigation. Our analysis strongly suggests otherwise.

whether a license harms competition that would have occurred in the absence of the license.[17] Thus we call this the law-enforcement approach.

We already know from Theorem 4 that consumers prefer to litigation to licensing with negative fixed fees. The arguments contained in the proof of Theorem 4 also establish:

Theorem 6: Suppose that banning negative fixed fees would cause the patent to be litigated. Then the ban lowers the running royalty rate (whatever the outcome of the litigation), benefits consumers, and raises total welfare. The ban also reduces the payoffs to the patent holder and to the downstream firms.

Why do consumers benefit if a ban on negative fixed fees leads to litigation? Following litigation, a patent that is found valid will command a running royalty of $\min[r_I, v]$. Of course, if the patent is found invalid, the technology is freely available. Therefore, litigation leads to *ex post* licensing at a royalty of $\min[r_I, v]$ with probability θ, and royalty-free use of the technology with probability $1-\theta$. Since $r(\theta) \geq r_I$, royalties may be lower, and cannot be higher, from litigation, so consumers must prefer litigation to licensing with negative fixed fees.

In the language of the *Guidelines*, licensing that involves a negative fixed fee is anticompetitive in the sense that it leads to higher output prices and is worse for consumers than litigation followed (if the patent is upheld) by licensing of the resulting ironclad patent. Using the benchmark developed by Shapiro (2003), these settlements would be considered anti-competitive.

B. Effect of Banning Negative Fixed Fees: Licensing Alternative

We now consider the impact of banning negative fixed fees if the parties would respond to the ban by signing licenses not including negative fees. We can restrict attention to patent strengths for which the ban is binding, $\theta \in [0, \theta^*]$; otherwise it has no impact. If the ban is binding, then

---

[17] The *Guidelines* (§3.1) state: "The Agencies will not require the owner of intellectual property to create competition in its own technology. However, antitrust concerns may arise when a licensing arrangement harms competition among entities that would have been actual or likely potential competitors (a firm will be treated as a

the patent holder will set the fixed fee at zero and use uniform royalties. We now solve for the optimal uniform royalties.

The patent holder's payoff from a uniform royalty *r* accepted by all downstream firms is $rx(r,r)$. This is increasing in *r* for $r < m$, because $T'(r) = \frac{d}{dr}\pi(r,r) + \frac{d}{dr}[rx(r,r)] > 0$ in this range, but $\frac{d}{dr}\pi(r,r) < 0$, so $\frac{d}{dr}[rx(r,r)] > 0$. Therefore, unsurprisingly, the patent holder will set the highest *r* that will be accepted.

A downstream firm's payoff from accepting a linear royalty of r, if it expects its rivals to accept, is $\pi(r,r)$. Its payoff from challenging the patent is $\theta\pi(v,r) + (1-\theta)\pi(0,0)$. So it will accept if and only if $\pi(r,r) \geq \theta\pi(v,r) + (1-\theta)\pi(0,0)$. Thus the highest acceptable royalty rate satisfies:

$$\pi(s(\theta), s(\theta)) \equiv \theta\pi(v, s(\theta)) + (1-\theta)\pi(0,0).$$

We can now compare $s(\theta)$ versus $r(\theta)$. Totally differentiating with respect to θ gives

$$s'(\theta) = \frac{\pi(0,0) - \pi(v,s)}{\theta\pi_2(v,s) - \frac{d}{ds}\pi(s,s)}.$$

Since $\frac{d}{ds}\pi(s,s) < 0$, $\pi_2(v,s) > 0$, and $\pi(0,0) > \pi(v,s)$, we have $s'(\theta) > 0$. Therefore, for all $\theta \in [0, \theta^*)$, $s(\theta) < s(\theta^*) = r(\theta^*) \leq r(\theta)$. The ban therefore lowers running royalties and thus benefits consumers. It harms downstream firms, since each one's equilibrium payoff is its expected payoff from infringing, and the ban has lowered its rivals' running royalty rates. Of course, banning negative fixed fees reduces the patent holder's payoff. Summarizing, we have:

Theorem 7: Consider a patent that would be licensed with negative fixed fees if allowed. Suppose that banning such fees would lead to licensing, not litigation. The ban lowers running

---

likely potential competitor if there is evidence that entry by that firm is reasonably probable in the absence of the licensing arrangement) in a relevant market in the absence of the license (entities in a "horizontal relationship")."

royalties, and thus benefits consumers and raises total welfare. The ban reduces the payoffs of the patentee and the downstream firms.

Figure 3 shows how a ban on negative fixed fees lowers running royalties in the case $r_I < v$.

6. Negative Fixed Fees and Innovation Incentives

Courts that have permitted negative fixed fees have given two general policy arguments. First, they have argued that such payments encourage settlement and may thus avoid or end litigation. This is presumably true, but settlements may well be possible without negative fixed fees. Furthermore, as we showed above, settlements with negative fixed fees tend to generate negative externalities by harming consumers, so we do not see a strong basis for favoring such settlements, especially if the harm to consumers is large relative to litigation costs. Second, courts have noted that retaining the option of negative fixed fees raises a patent holder's payoff and thus encourages innovation.

This second argument often seems to bring reasoned policy discourse to a halt, for two reasons. First, innovation is often seen as sacrosanct and courts are rightly very wary of taking actions in the name of promoting competition that will impede innovation. Second, this argument seems to imply that the long-run net welfare effect of banning negative fixed fees depends on the elasticity of supply of inventions, concerning which we have little data. Yet we need not be so agnostic. While it is important to reward innovation, *there are better ways to do so* (even narrowly within the patent system). In particular, we show that the optimal patent and antitrust system prohibits negative fixed fees, *regardless* of the elasticity of supply of inventions.

Our analysis here follows the general approach of Kaplow (1984) and Gilbert and Shapiro (1990). Kaplow proposed a "ratio test" to assess whether patent holders should be allowed to engage in certain practices. Under this test, a practice should be permitted if and only if the ratio of patentee reward to monopoly loss is sufficiently high. Here, "patentee reward" is the reward to the patent holder resulting from using the practice in question, and "monopoly loss" is the deadweight loss resulting from using the practice in question. Kaplow showed that this approach fits well with standard economic principles. Similarly, Gilbert and Shapiro showed that a change in policy that lowers running royalties and extends patent life so as to leave present-value patentee profits (and hence R&D incentives) unchanged is generally an improvement.

In the same spirit, we find that negative fixed fees are an inefficient way to generate profits for patent holders: they create more deadweight loss, relative to patent-holder profits, than the corresponding ratio for extending patent life. This is because (as we have shown) they enable weak patents to be licensed with relatively high running royalties, generating relatively large deadweight loss yet relatively little in patent profits (because the patent holder must pay negative fixed fees to induce downstream firms to pay the high running royalties).

The key condition needed for this line of argument is that the ratio of deadweight loss to running royalty revenues be increasing in the royalty rate. Define the deadweight loss that results if all downstream firms pay a royalty of r as $D(r)$. (Deadweight loss is the difference between total welfare if the innovation were freely available to all downstream firms, and actual total welfare.) We assume that the ratio $D(r)/rx(r,r)$ is increasing in r for $r \le v$. This is a natural condition stating that the ratio of deadweight loss to running royalty revenues increases with price in this range. Gilbert and Shapiro (1990) show that a comparable condition holds with a homogeneous good if both profits and total welfare are concave in output; that condition in turn holds if v is small and demand is linear. In the Appendix we prove:

Theorem 8: If the ratio of deadweight loss to running royalty revenues, $D(r)/rx(r,r)$, is increasing in r for $r \le v$, then the optimal patent and antitrust system prohibits negative fixed fees. In particular, if the status quo permits negative fixed fees, one can construct an alternative system, in which such fees are banned and patent lifetimes are longer, that has lower deadweight loss and the same innovation incentives.

Theorem 8 does not even exploit the *meaning* of a patent's being weak. A weak patent is one that, according to a court enquiry far more thorough (and fair) than the cursory and somewhat one-sided PTO examination, is likely to be found invalid. By definition, a weak patent is one where it appears likely that the PTO missed relevant prior art or that the patented material is obvious. Many such patents, in other words, do not reflect inventions (by the patent-holder) that patent policy efficiently should (or is meant to) reward: many owners of weak patents are not innovators at all. Thus, allowing very much of the money from patent-related markups in product markets to flow to holders of weak patents is likely to be a highly inefficient way to reward invention, relative to sending more such funds to holders of strong patents.

Of course, unlike Congress, courts evaluating a settlement with negative fixed fees cannot extend patent lifetimes, even if they knew by how much to do so. However, if patent lifetimes are roughly optimal, then to first order it does not much matter for total welfare whether or not one increases patent lifetimes modestly. Since negative fixed fees are not ubiquitous, the overall increase required in Theorem 8 would surely be modest. Thus simply banning negative fixed fees would be a welfare improvement unless patent lifetimes are already too short in general (and, depending on innovation elasticities, maybe even if they are). Even in that case, however, rewarding holders of weak patents by permitting negative fixed fees is a poor way of correcting for the inadequate returns to innovation resulting from patent lifetimes being too short.

7. Do We Need Patent Reform?

We now show that weak patents can command surprisingly large running royalties even without the use of negative fixed fees. This finding suggests that it would be valuable to improve patent quality and reduce the number of weak patents that are issued, even if antitrust rules prohibit negative fixed fees in patent licenses.

If negative fixed fees are banned, weak patents are either litigated or licensed at the running royalty rate $s(\theta)$ derived above. Recall that $s(\theta)$ is defined by $\pi(s,s) \equiv \theta\pi(v,s)+(1-\theta)\pi(0,0)$. This equation implies that $s(0)=0$ and $s(1)=v$. We have already shown that $s'(\theta)>0$. We now study the $s(\theta)$ function more closely. We are especially interested in the properties of $s(\theta)$ for small values of θ, i.e., for very weak patents.

We find that $s(\theta)$ is much larger than a natural normative benchmark for very weak patents. In particular, we compare $s(\theta)$ with the benchmark of θv. Conventional analysis of patent licensing often compares per-unit royalty rates with a benchmark of the per-unit cost savings from the patented invention, *v*. After all, *v* measures the innovator's contribution and society's benefit (per unit of output) from the invention. Thus, if per-unit royalties are less than v, consumers benefit from the licensed innovation and the innovator is under-rewarded (holding aside any fixed fees and assuming that the innovation would otherwise never have occurred). Alternatively, if per-unit royalties exceed v, consumers are actually harmed by the patented invention (*not* just by the patent given the presence of the invention) and the innovator is over-rewarded (even assuming that the innovation would otherwise never have been made). Broadly, then, if patent strength proxies reasonably well for the probability that the patent holder truly *made* a non-obvious contribution, one would want to deflate the benchmark of v to θv. In addition, when a patent holder licenses at uniform royalties to a single downstream firm, the equilibrium running royalty is below θv.[18]

Further results about $s(\theta)$ depend on the downstream oligopoly payoff functions. However, we have some general results for very weak patents. Setting $\theta = 0$ in the expression for $s'(\theta)$, we have

$$s'(0) = \frac{\pi(0,0) - \pi(v,0)}{-\frac{d}{ds}\pi(s,s)\Big|_{s=0}}$$

For small v, we have $\frac{s'(0)}{v} = \frac{-\pi_1(0,0)}{-[\pi_1(0,0) + \pi_2(0,0)]} > 1$. Thus $s(\theta) > \theta v$ for small θ. Direct calculations provided in the Appendix prove the second part of:

Theorem 9: For small, very weak patents, the uniform royalty rate exceeds θv. With Cournot oligopoly, constant marginal costs, and inverse demand $p(X)$, $s(\theta) \approx N\theta v \frac{2}{2 + Xp''(X)/p(X)}$.

For small, very weak patents, Theorem 9 implies that $s(\theta) \approx N\theta v$ with linear demand and $s(\theta) > N\theta v$ with constant elasticity demand (and $\varepsilon > 1$, as needed for normality of the profit function).

[18] A single licensee will only accept a per-unit royalty of s if $\pi(s) \geq \theta\pi(v) + (1-\theta)\pi(0)$. Since profits are convex in the per-unit royalty, this requires that $s < \theta v$.

The reason that $s(\theta)$ exceeds θv for small, very weak patents can be gleaned from the role of π_2 in the equation $\frac{s'(0)}{v} = \frac{-\pi_1(0,0)}{-[\pi_1(0,0)+\pi_2(0,0)]}$. An increase in r that applies to an entire industry may be largely passed through to consumers. The downstream price rises and the profits of licensees may hardly be affected in equilibrium. For instance, with constant unit costs and nearly perfect competition, the licensees are almost completely unaffected by the (common) level of r, so $\pi_1(0,0)+\pi_2(0,0)$ is very close to zero, making $s'(0)/v$ very large. Theorem 9 shows that a firm's own costs can be far more important to that firm's profits than common costs even with imperfect competition.

Theorem 9 allows us to find sufficient conditions under which weak patents will be licensed rather than litigated if negative fixed fees are prohibited. As shown in the Appendix, the patent holder will license rather than litigate if $r_I \geq v$. If $0 < r_I < v$, we obtain the same result in the special case of linear demand Cournot oligopoly with constant marginal costs unless *N* is small.

We also can relate the key ratio here, $\frac{-\pi_1(0,0)}{-[\pi_1(0,0)+\pi_2(0,0)]}$, to more familiar properties of the underlying downstream oligopoly payoff functions. These relationships are easiest to see with a downstream duopoly with differentiated products. Define the Diversion Ratio as $\delta = -\frac{\partial x_2 / \partial p_1}{\partial x_1 / \partial p_1}$; this measures the fraction of sales lost by one firm when its price rises that are captured by the other firm. In a differentiated product Bertrand duopoly one can show that

$\frac{\pi_1}{\pi_1+\pi_2} = \frac{1-\delta[\frac{dp_2}{dc_1}]}{1-\delta[\frac{dp_2}{dc_1}+\frac{dp_2}{dc_2}]}$. The key variables driving this ratio are the Diversion Ratio, which reflects the closeness of substitutes of the two products, and the rates at which a firm's own costs, and the costs of its rival, are passed through into that firm's price.

Competition implements a *relative* performance scheme among competitors, so rewards (profits) are rather insensitive to common cost shocks if the industry is fairly competitive. In such an environment, the key driver of profits is the possibility of an idiosyncratic cost difference. Under the Supreme Court's *Blonder-Tongue* decision, a firm cannot gain an idiosyncratic cost *advantage* by successfully challenging a weak patent. But an idiosyncratic cost *disadvantage* awaits an unsuccessful challenger, since $s < v$. Downstream competition makes challenging unappealing compared to accepting the same terms as one's rivals.

We conclude that very weak patents can have surprisingly strong effects, even if negative fixed fees are prohibited. These effects arise because invaliding a patent provides a public good for the downstream industry, a fact that the patent holder can profitably exploit. Excessive payments for weak patents are especially likely to arise for patents that are licensed to a highly competitive downstream industry. Our analysis suggests very substantial benefits from reforming the patent system to improve patent quality and/or from encouraging challenges to issued patents.

8. Lax Treatment of Patent Scope

As discussed in Section 3 above, we have assumed that patent and antitrust law together impose the constraint $r \le v$ on licensing agreements. However, for weak patents, $r(\theta) > v$, so the holder of a weak patent could and would profitably induce each downstream firm to accept a two-part tariff (F, r) with $F < 0$ and $r = r(\theta) > v$ if such a contract were enforceable. In other words, relaxing the $r \le v$ constraint would exacerbate the adverse effects on consumers of negative fixed fees used with weak patents.

We have assumed thus far that such contracts are not enforceable.[19] We might call this the "strict patent-scope regime." In the alternative "lax patent-scope regime" in which $r > v$ is feasible, even larger negative fixed fees are used, and running royalties on weak patents are even higher. In the Appendix, we prove the analog of Theorem 1 in the lax patent-scope regime, and show that the upper bound on the royalty rate for weak patents is $r_0 > v$, defined by $\pi(v, r_0) = \pi(0,0)$. Here, r_0 is the highest royalty level such that a downstream prefers infringing and challenging to just using the backstop technology. We know that $r_0 > v$ because normality implies that $\pi(v, v) < \pi(0,0)$ and $\pi_2 > 0$.

[19] Customers and/or antitrust enforcement agencies may be needed to challenge such agreements, since by construction the fixed fee co-opts the licensee (and can be paid on a recurring basis).

Theorem 10: In the lax patent-scope regime with $r_I < r_0$, the per-unit royalty in a settlement license is $r = \min[r_0, r(\theta)]$.

Figure 4 shows the running royalty rate as a function of patent strength in the lax patent-scope regime. Like Figure 1, Figure 4 is drawn for the case where $r_I < v$.

In the lax patent-scope regime, patent licenses can be used to raise downstream prices above the level that prevailed prior to the patented invention, by charging running royalties above *v*, perversely making consumers worse off than if the patented innovation had never been made.[20] We find that weak patents will always be used in this way (so long as $m > v$, as we continue to assume). However, *strong* patents will not used in this way, at least if $r_I < v$ (recall that this holds if downstream competition is Cournot oligopoly). Negative fixed fees raise running royalties by more, and thus the *ex post* benefit of banning negative fixed fees is larger, in the lax patent-scope regime than in the strict regime.[21]

9. Will Licenses Survive a Finding of Invalidity? Blonder-Tongue

We have assumed so far that the patent holder cannot charge royalties to any downstream firm, even those that have signed licenses, if the patent is litigated and found invalid. This is the current state of U.S. law, under *Blonder-Tongue.* This has the attractive feature that it prevents downstream firms from having to pay royalties on a patent that has been found invalid.[22]

---

[20] Unfortunately, the lax patent-scope regime is quite realistic. Antitrust enforcement by third parties may be very difficult if royalties are also paid for trade secrets or hidden in overcharges for other inputs. Furthermore, in the language of antitrust law, the patent holder and the licensees might successfully argue that the "ancillary restraints" beyond the scope of the patent are necessary to make licensing feasible. They also might argue that the licensee *was in fact* using the patented technology (which it will if royalties are imposed regardless of the technology used, since the patented technology saves costs), so there has been no actual imposition of royalties outside the scope of the patent. Of course, this argument is incorrect, because the agreement closed off a technology, the backstop technology, that (had the agreement not imposed such royalties) the licensee would have used given the high running royalties charged for the patented technology.

[21] No licensee would accept $r > v$ without a negative fixed fee, so the strict and lax regimes are equivalent if negative fixed fees are prohibited.

[22] *Blonder-Tongue* is similar in some respects to the legal rule prohibiting royalties that extend beyond the lifetime of the patent. The Courts appear uncomfortable with agreements that involve royalties outside of the patent grant.

However, as we have seen, it also discourages patent challenges: a successful challenger is providing a public good for others, often including its product-market rivals.[23]

For this reason, in our model the patent holder would voluntarily write these same conditions into its licenses, even if not required to do so. In equilibrium no downstream firms challenge the patent, so foregoing royalties in the event of a finding of invalidity has zero *equilibrium* cost to the patent holder. But by ensuring that a challenger's rivals can freely use the patented technology after a finding of invalidity, voluntarily implementing *Blonder-Tongue* in licenses makes infringing less attractive and thus lets the patent holder impose higher royalties in equilibrium. [However, a patent holder who is licensing to firms in different downstream markets would only choose to implement *Blonder-Tongue* within downstream markets, not across them.][24]

10. Conclusions and Policy Implications

We have shown that weak patents, i.e., patents that may well be found invalid if tested by a court, can have surprisingly large effects when licensed to a group of downstream oligopolists.

The most striking effects arise when the patent holder is permitted to make lump-sum payments to the downstream firms to induce them to sign licenses. In this case, the patent holder can partially cartelize the downstream industry by offering licensing contracts of the following form: "I will pay you a lump sum if you agree to pay me a high per-unit royalty. If you challenge my patent rather than signing a license, you will be at a competitive disadvantage if you lose and the patent is found valid. But your upside from litigating is sharply limited: if you win and invalidate my patent, your rivals will automatically be relieved of any obligation to pay royalties." This last provision follows because U.S. patent law effectively relieves licensees of the obligation to pay royalties for a patent that has been declared invalid. Therefore,

---

(This relates to the strict patent-scope regime.) However, rules that limit royalties in this manner may have surprising and unintended effects on equilibrium licensing agreements.

[23] Miller (2004) stresses the public-good aspect and offers policy suggestions such as a bounty to successful patent challengers. Farrell and Merges (2004) argue that pass-through of running royalties exacerbates the problem.

successfully challenging a patent provides a public good for the downstream industry, an observation that is essential to our analysis.

These licensing agreements support a high downstream price, which increases total profits at the expense of consumers. The lump-sum payments, i.e., negative fixed fees, are designed to split these profits in a manner that is acceptable to downstream firms and thus avoids litigation. In this fashion, weak patents are not tested in court. In our main case, the "strict patent-scope regime," the only limit to this cartelizing mechanism for weak patents is the requirement that the running royalty not exceed the cost savings associated with the patented technology. In the lax patent-scope regime, even higher per-unit royalties are both possible and jointly profitable for weak patents.

In our model, prohibiting negative fixed fees, also known in the legal literature as "reverse payments" from the patent holder to its licensees, is a desirable antitrust policy. In the short run, a ban on negative fixed fees raises welfare and consumer surplus. In the long run, the optimal patent and antitrust policy prohibits negative fixed fees. In fact, using negative fixed fees benefits the owners of *weak* patents, whose social contributions are most tenuous, far more than the owners of strong patents, who are more likely to have actually innovated.

Although we obtain these results in a model that involves licensing by a patent holder to a set of downstream oligopolists, similar principles apply in the setting that has given rise to a number of recent court cases in which the legality of negative fixed fees has been tested: an incumbent supplier settles patent litigation by making a lump-sum payment to an actual or potential competitor as part of a patent licensing agreement. In particular, all of the court cases of which we are aware that test the legality of negative fixed fees under antitrust law arise in the pharmaceutical industry, where the Hatch-Waxman Act regrettably allows a deal between a patent-holder and one generic challenger to erect legal barriers (beyond simple patent protection)

[24] [Is this right? No Segal/relativity boost from doing so across markets, but it would help with litigation credibility, or would it?]

to entry by other generic producers.[25] In that setting, a lump-sum payment by the incumbent to the potential entrant, in exchange for an agreement to delay or abandon entry, is a horizontal agreement between rivals that is even more clearly harmful to competition than the vertical licensing agreements we have studied. While one can construct examples in which negative fixed fees are needed to settle patent litigation between an incumbent monopolist and a potential entrant (Willig and Bigelow (2004)), we support a rule under which such payments are presumptively anti-competitive.

Further research is needed to explore the implications of the prediction in our model that negative fixed fees will be widely used in licensing agreements between an upstream owner of a weak patent and downstream users of the patented technology. As an empirical matter, how widely used are such "reverse payments" in patent licensing agreements? Even if "naked" reverse payments are not common, are negative fixed fees frequently hidden in more complex licensing or cross-licensing agreements? Our findings could serve as a warning about the incentives to use negative fixed fees in the licensing of weak patents. With the recent surge of patenting, and with the growing chorus of concerns about patent quality, are difficult to detect "reverse payments" being used more and more frequently in licensing agreements?

Alternatively, our analysis of negative fixed fees could be seen as posing a puzzle: if negative fixed fees are not often used, why not? One possible explanation is that patent holders are deterred from using negative fixed fees by the risk of antitrust liability. If so, the usage of negative fixed fees might grow substantially if the recent court decisions permitting them are upheld and establish a broad precedent.[26] Another possible explanation is that high running royalties, the *quid pro quo* for the negative fixed fees, are difficult to sustain if licensing

[25] See FTC (2002). Title XI of the 2003 Medicare Prescription Drug Improvement and Modernization Act seeks to close some of the opportunities that the Hatch-Waxman Act offered for anticompetitive deals, and requires certain deals to be reported to the federal antitrust agencies.

[26] The Schering case departed from our model in at least two significant ways: the patent holder was integrated into the "downstream" market, and because of the vagaries of the Hatch-Waxman Act (see for instance FTC 2002 and Bulow 2004), there was effectively only one other downstream player. While we have not formally analyzed such a market structure, it seems that the Segal effect will not play the role that it does in our model. However, the language of the Schering decision would seem broad enough to cover markets to which our model would apply, making the decision troubling as general policy. (Separately, the settlement condoned in the Schering case appears anticompetitive for the reasons discussed by Shapiro 2003.)

agreements are secret. Or perhaps in some industries patent holders are deterred from using negative fixed fees by the threat of entry. To further understand the licensing of weak patents in these industries, our analysis could be extended to incorporate asymmetries among the existing downstream firms or entry by additional downstream firms.

Even if negative fixed fees are not used, we find that weak patents can command surprisingly high per-unit royalties when licensed to a downstream industry. Our analysis indicates that relying on traditional patent litigation to challenge weak patents is inadequate, especially if a number of firms who compete against one another downstream all benefit from using the patented technology. Consider a patent covering a technology that permits a cost saving of v per unit, and suppose that the patent is weak, i.e., valid only with a low probability, θ. If the owner of this patent licenses it using uniform royalties to N equally placed downstream oligopolists, our analysis indicates that the patent will command running royalties of roughly $N\theta v$. This is troubling, since the far smaller per-unit royalty of θv provides a natural normative benchmark: the cost savings associated with the patented technology times the probability that the patent holder truly discovered this technology (rather than obtaining a patent for a technology that was either obvious or already described in prior art). Even if antitrust law prohibits the use of negative fixed fees in patent licensing agreements, a weak patent can still command royalties far out of proportion to the patent holder's (expected) social contribution.

Since weak patents can have strong adverse effects on competition and on consumers, there are large benefits from reforming the patent system to improve patent quality and reduce the number of weak patents issued. There also are large benefits to establishing some new system by which weak patents can and will be challenged. Our findings caution that close attention must be paid to licensees' incentives to challenge weak patents, recognizing that the owners of those patents will seek to sign licenses precisely to avoid having them tested in court.

We believe that the analysis in this paper significantly advances our understanding of the commercial effects of probabilistic patents, especially weak patents. However, more empirical and theoretical work is needed to provide more reliable guidance for policy. What strategies do patent holders use to enhance the credibility of their threats to litigate against infringers, and how do these strategies affect equilibrium licensing agreements? How does the analysis differ in the presence of litigation costs, under different assumptions about how patentees and licenses negotiate, or in the presence of private information about patent validity? How are licenses structured if patent holders have difficulty committing to using observable licensing agreements or if licensees have difficulty enforcing most-favored customer clauses in licensing agreements? How is the analysis different if the patent holder is vertically integrated and competes directly against the downstream firms to whom it offers licenses? How are licenses structured if they are signed in the shadow of patent litigation over *infringement* rather than validity? How should a post-grant opposition system be designed so that important weak patents will be challenged? While our work does not answer these important economic and current policy questions, we believe our framework and findings provide useful building blocks to address them.

References

Anton, James, and Dennis Yao (2003), "Patents, Invalidity, and the Strategic Transmission of Enabling Information," *Journal of Economics and Management Strategy* 151-178.

Bulow, Jeremy, "The Gaming of Pharmaceutical Patents," pp. 145-187 in *Innovation Policy and the Economy*, volume 4, Adam Jaffe, Scott Stern, and Josh Lerner, editors, MIT Press 2004.

Choi, Jay Pil (1998), "Patent Litigation as an Information-Transmission Mechanism," *American Economic Review*, 88:5, 1249-1263.

Choi, Jay Pil (2002), "Patent Pools and Cross Licensing in the Shadow of Patent Litigation," available at http://www.msu.edu/~choijay.

Farrell, Joseph, and Robert Merges (2004), "Incentive to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help," *Berkeley Technology Law Journal*, 19:2, 943-970.

Farrell, Joseph, and Carl Shapiro, (1990), "Asset Ownership and Market Structure in Oligopoly," *Rand Journal of Economics*, 21:2, Summer, 275-292.

Federal Trade Commission (2002), *Generic Drug Entry Prior to Patent Expiration*, July.

Federal Trade Commission (2003), *To Promote Innovation: The Proper Balance Between Competition and Patent Law and Policy*, October.

Gilbert, Richard, and Carl Shapiro (1990), "Optimal Patent Length and Breadth," *Rand Journal of Economics*, 21:1, Spring, 106-112.

Hovenkamp, Herbert, Mark D. Janis and Mark A. Lemley (2003), "Anticompetitive Settlement of Intellectual Property Disputes." *Minnesota Law Review*. 87:6, pp. 1719-1766.

Kaplow, Louis (1984), "The Patent-Antitrust Intersection: A Reappraisal," *Harvard Law Review*, 97:8, June, 1815-1892.

Kimmel, Sheldon (1992), "Effects of Cost Changes on Oligopolists' Profits," *Journal of Industrial Economics*, 441.

Lemley, Mark (2001), "Rational Ignorance at the Patent Office," *Northwestern University Law Review* 95 (2001).

Lemley, Mark A., and Carl Shapiro (2005), "Probabilistic Patents," *Journal of Economic Perspectives*, 19:2, Spring, 75-98.

McAfee, Preston and Marius Schwartz (1994), "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity and Uniformity," *American Economic Review*, 84:1, March, 210-230.

Miller, Joseph Scott (2004), "Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents," *Berkeley Technology Law Journal*, 19:2, 667-739.

Meurer, Michael (1989), "The Settlement of Patent Litigation," *Rand Journal of Economics*, Spring, 20:1, 77-91.

National Academies of Science (2004), *A Patent System for the 21st Century*, http://www.nap.edu/books/0309089107/html.

Segal, Ilya (1999), "Contracting with Externalities," *Quarterly Journal of Economics*, 114:2, May, 337-388.

Segal, Ilya and Michael Whinston (2003), "Robust Predictions for Bilateral Contracting with Externalities," *Econometrica*, 71:3, May, 757-791.

Shapiro, Carl (1989), "Theories of Oligopoly Behavior," *Handbook of Industrial Organization*.

Shapiro, Carl (2003a), "Antitrust Limits to Patent Settlements," *Rand Journal of Economics*, 34:391-411.

Shapiro, Carl (2003b), "Antitrust Analysis of Patent Settlements Between Rivals," *Antitrust*, Summer, pp. 70-77.

Waterson, Michael (1990), "The Economics of Product Patents," *American Economic Review* 80:4 (September) 860-869.

Willig, Robert, and John Bigelow (2004), "Antitrust Policy Toward Agreements That Settle Patent Litigation." *Antitrust Bulletin*. 49:3.

Appendix

No-Discrimination Lemma

We show here that the patent holder optimally offers a single two-part tariff (F, r) to all of the downstream firms, and that (F, r) is chosen so that it is an equilibrium for all N downstream firms to accept (F, r). We present the argument in four lemmas.

Recall first that we limit the patent holder to making simultaneous public offers (F_i, r_i), and each downstream firm i chooses whether to accept or reject its offer (and, if the latter, whether to infringe or use the backstop technology); the offers are not contingent on what other downstream firms do. We denote by $(\mathbf{F}, \mathbf{r})$ the vector of offers (F_i, r_i). We say that $(\mathbf{F}, \mathbf{r})$ is symmetric if $(F_i, r_i) = (F, r)$ for all i.

Lemma A1. The patent holder will not make offers $(\mathbf{F}, \mathbf{r})$ such that any downstream firm(s) choose the backstop technology and produce a positive quantity.

Proof. Suppose that in response to offers $(\mathbf{F}, \mathbf{r})$, downstream firm i chooses the backstop technology and produces $x_i(\mathbf{r}) > 0$. If the patent holder changes $(\mathbf{F}, \mathbf{r})$ by replacing (F_i, r_i) by $(0, v)$, then (a) firm i 's marginal cost is v, as before, whether or not it accepts the new offer; (b) whether it accepts or rejects, it is an equilibrium for all other firms to make the same choices (both in licensing and in the product market) as before; (c) firm i will therefore be indifferent to accepting the new offer. For convenience, we assume that the downstream firm accepts the new offer if it is indifferent. Note that this argument uses the uniqueness of product-market equilibrium. In the case where some (or all) downstream firms are indifferent at equilibrium between two or more of their licensing strategies we assume that they would continue to make the same choices.

Lemma A1 implies that all downstream firms either accept or infringe. For Lemma A2, we introduce quasi-concavity/convexity assumptions that we assume henceforth.

Quasi-Concave Total Profits: Total upstream plus downstream profits, $T(\mathbf{r}) \equiv \sum_i [r_i x_i + \pi^i(r_i, r_{-i})]$, are quasi-concave in $\mathbf{r}$.

Quasi-Convex Reservation Payoffs: The sum of the reservation payoffs of the downstream firms from using the backstop technology, $S(\mathbf{r}) \equiv \sum \pi^i(v, r_{-i})$, is quasi-convex in $\mathbf{r}$.

We show below that these assumptions hold in Cournot duopoly with linear demand and constant marginal cost, so long as both firms remain active, and we conjecture that they hold in many other sensible symmetric oligopoly models.

Lemma A2. With quasi-concave total profits and quasi-convex reservation payoffs, if $(\mathbf{F},\mathbf{r})$ induces all downstream firms to accept, then there exists a symmetric $(\hat{\mathbf{F}},\hat{\mathbf{r}})$ that is also accepted by all and that yields a higher profit to the patent holder than $(\mathbf{F},\mathbf{r})$.

Remark: With *strict* quasi-concavity/convexity, if $(\mathbf{F},\mathbf{r})$ is not symmetric, then the symmetric $(\hat{\mathbf{F}},\hat{\mathbf{r}})$ yields *strictly* higher profits.

Proof. We can work in $\mathbf{r}$ space and let $\mathbf{F}$ be defined implicitly by the incentive to accept. Thus we do not need to check the condition that firms accept the offers, but rather use it to determine the fixed fees.

Suppose that $\mathbf{r}$ is asymmetric. (If $\mathbf{r}$ is symmetric but $\mathbf{F}$ is not, the patent holder can easily find a symmetric $\mathbf{F}$ that is more profitable: just charge all downstream firms slightly less than the biggest F_i featured in $\mathbf{F}$. Since the oligopoly is symmetric and product-market behavior does not depend on F, all downstream firms will pay this F, since at least one was willing to pay it.) Set $\hat{r}_i = \frac{1}{N}\sum_j r_j$ for all *i*. Then $\hat{\mathbf{r}}$ is symmetric. Since the oligopoly game is symmetric, quasi-concavity of T implies that $T(\hat{\mathbf{r}}) \geq T(\mathbf{r})$. Similarly, symmetry together with quasi-convexity of S implies that $S(\hat{\mathbf{r}}) \leq S(\mathbf{r})$.

Now, if the backstop constraint binds, then i will accept paying a fixed fee of up to $F_i = \pi^i(\mathbf{r}) - \pi(v,\mathbf{r}_{-i})$. Hence the patent holder's profit from an offer whose variable royalties are $\mathbf{r}$, including fixed fees that (just) make the offer acceptable to all, is equal to $T(\mathbf{r}) - S(\mathbf{r})$. From the above, $T(\hat{\mathbf{r}}) - S(\hat{\mathbf{r}}) \geq T(\mathbf{r}) - S(\mathbf{r})$, so $\hat{\mathbf{r}}$ is more profitable than $\mathbf{r}$.

Likewise, if the infringe constraint binds, then i is willing to pay a fixed fee of up to $F_i = \pi^i(\mathbf{r}) - \theta\pi(v,\mathbf{r}_{-i}) - (1-\theta)\pi(0,0)$. Hence the patent holder's profit from an offer whose variable royalties are $\mathbf{r}$, including fixed fees that (just) make the offer acceptable to all, is equal to $T(\mathbf{r}) - \theta S(\mathbf{r}) - (1-\theta)\pi(0,0)$. From the above, $T(\hat{\mathbf{r}}) - \theta S(\hat{\mathbf{r}}) \geq T(\mathbf{r}) - \theta S(\mathbf{r})$, and of course $(1-\theta)\pi(0,0)$ is common, so again $\hat{\mathbf{r}}$ is more profitable than $\mathbf{r}$.

Remark: For very weak patents, i.e., for very small θ, the $S(\mathbf{r})$ function is given little weight, so the lemma holds for very weak patents so long as the $T(\mathbf{r})$ function is strictly quasi-concave.

Two final lemmas (A3 and A4) establish that the patent holder will indeed make offers that are accepted in equilibrium, rounding out the justification of the assumptions in the text:

Lemma A3. The patent holder will not make offers that are expected to induce any firms to infringe, without intending to litigate.

Proof. Assuming that the patent holder's decision not to litigate is predictable to downstream firms, such offers would be equivalent to making (accepted) offers of $(F_i, r_i) = (0,0)$ to (a) some or (b) all firms. (a) If it were only to some firms, the other firms would be accepting other offers.

(Lemma A1 shows that they are not choosing the backstop technology. If they are choosing to infringe and challenged the patent, Lemma A3 would follow immediately because the patent holder is litigating.) By Lemma A2, that could not be optimal for the patent holder. (b) It cannot be optimal for the patent holder to offer $(F_i, r_i) = (0,0)$ to all firms: that would make no money, but (for $\theta > 0$) the patent holder could make some money by picking on any one firm and demanding a small but strictly positive amount for a license, and that firm would pay some positive amount rather than face litigation.

Lemma A4. The patent holder will not adopt a hybrid strategy in which some downstream firms accept offers and it litigates against others.

Proof. Suppose that the patent holder signed licenses (F_i, r_i) with downstream firms $i = 1,...,k$ and then engaged in litigation. If the patent is ruled invalid, these licenses are void and hence have no effect on the patent holder's payoff. Therefore, we can focus attention on the state of nature in which the patent is ruled valid. In that event, suppose that the patent holder finds it optimal to sign licenses (F_i, r_i) with downstream firms $k+1,...,n$. (By the same logic as Lemma A1, the patent holder cannot find it optimal to induce any downstream firms to use the backstop technology after a finding of validity.) Then the payoff to the patent holder is precisely the same as if the patent holder had (a) not signed any licenses prior to litigation, then (b) engaged in litigation, and finally, (c) signed all of these same (F_i, r_i) after the finding of validity. This is because the running royalties are the same (each downstream firm is paying the same per-unit royalties in these two scenarios) and the fixed fees are the same (each downstream firm has the same reservation payoff in the two scenarios). But we know from Lemma A2 that this payoff is less than the payoff the patent holder could earn by offering the same two-part tariff to all downstream firms after a finding of validity. Therefore, the hybrid strategy cannot be optimal.

Cournot Duopoly Satisfies the Quasi-Concavity/Convexity Conditions

Total profits are $T = p(X)X - \sum_i c(x_i)$. With linear demand and constant costs, total profits are concave in total output X, which in turn is a linear function of the sum of firms' marginal costs, which is linear in **r**. Hence total profits are concave, hence quasi-concave in **r**.

The sum of the reservation payoffs as a function of r is in general quite complex, but with two firms it is $S(\boldsymbol{r}) = \pi(v, r_1) + \pi(v, r_2)$. Replacing each of r_1 and r_2 with their average will reduce the sum of the reservation payoffs if $\pi(v, r)$ is convex in r: the profits of one firm are convex in the other firm's royalty rate: $\partial^2 \pi^i / \partial r_j^{\,2} > 0$ for $i \neq j$. This holds with Cournot duopoly, linear demand, and constant marginal costs if both firms remain active.

Proof of Theorem 1

If $r(\theta) \le v$, then $r(\theta)$ is feasible and is therefore the optimal running royalty. Since *G* is single-peaked in *r*, if $r(\theta) \ge v$, then the patent holder sets *r* as high as allowed, namely $r = v$. This proves the first part of Theorem 1.

To prove the second statement of Theorem 1, let r_1 and r_2 be the optimal choices for given θ_1 and θ_2. By optimality, $G(r_1,\theta_1) \ge G(r_2,\theta_1)$ and $G(r_2,\theta_2) \ge G(r_1,\theta_2)$. Adding and rearranging, $G(r_1,\theta_1) - G(r_1,\theta_2) \ge G(r_2,\theta_1) - G(r_2,\theta_2)$. Using $G(r,\theta) \equiv T(r) + \theta(\pi(0,0) - \pi(v,r)) - \pi(0,0)$, this is equivalent to $(\theta_1 - \theta_2)[\pi(0,0) - \pi(v,r_1)] \ge (\theta_1 - \theta_2)[\pi(0,0) - \pi(v,r_2)]$, or $(\theta_1 - \theta_2)[\pi(v,r_2) - \pi(v,r_1)] \ge 0$. If $\theta_1 > \theta_2$, we must have $\pi(v,r_2) - \pi(v,r_1) \ge 0$, which requires $r_2 \le r_1$. An alternative proof (assuming differentiability) uses $\frac{\partial^2 G(r;\theta)}{\partial r \partial \theta} = -\pi_2(v,r) < 0$ to show $r'(\theta) < 0$.

Finally, the running royalty is the lower envelope of a decreasing function $r(\theta)$ and a constant function v, so it is weakly decreasing in patent strength on the whole range of θ, completing the proof of Theorem 1.

Ironclad Running Royalty versus Patent Size

For an ironclad patent ($\theta = 1$), the patent holder maximizes $G(r;1) = rx(r,r) + \pi(r,r) - \pi(v,r)$. Since $G(r;1)$ is single-peaked, $r_I < v$ if and only if $G(r;1)$ is declining in *r* at $r = v$. Differentiating $G(r;1)$ with respect to *r* and evaluating at *v* gives:

$$G_r(v;1) = x(v,v) + \pi_1(v,v) + v[x_1(v,v) + x_2(v,v)].$$

Since output declines when costs rise uniformly, the term in square brackets is negative. To sign the sum of the first two terms, note that the second term is the effect on profits of marginally higher own unit costs, π_1. We can decompose π_1 into a "direct" effect of higher costs on given output, which is just –*x*, canceling the first term, and an "indirect" effect on the firm's profits that arises through rivals' response to learning that the firm has higher costs. The sum of the first two terms is thus just that indirect effect.

Cournot competition. The indirect effect is negative: when rivals learn that (say) firm 1 has higher costs, they expect it to produce less output; as a result, rivals raise their own output, which reduces firm 1's profits. Therefore $G_r(v;1) < 0$ and so $r_I < v$. In fact, for small patents, i.e., for *v* small, the equilibrium running royalty rate can be negative. Of course, these licenses have positive fixed fees.

Indeed, $r_I < v$ whenever the indirect effect of an increase in own costs is either negative (as in Cournot) or else smaller than the negative effect of the decline in output when costs rise uniformly. However, we next display a reasonable oligopoly model where $r_I > v$.

Bertrand Oligopoly with Differentiated Products. With inelastic market demand in the range $r \leq v$, $\pi(r,r)$ is independent of *r* so long as $r \leq v$. Hence the "output effect" term in square brackets is zero. In differentiated product Bertrand oligopoly, the sum of the first two terms, i.e. the "indirect" profit effect of marginally higher own unit costs, is positive: when rivals learn that (say) firm 1 has higher costs, they expect it to set a higher price, and so they raise their own prices, increasing firm 1's profits. Hence $G_r(v;1) > 0$ and so $r_I > v$.

Proof of Theorem 2

Define θ^* as the solution to $\theta = \dfrac{\pi(0,0) - \pi(r(\theta), r(\theta))}{\pi(0,0) - \pi(v, r(\theta))}$. We show:

(a) if $r_I \geq v$, then for all θ, $F = -(1-\theta)[\pi(0,0) - \pi(v,v)] < 0$;

(b) if $r_I < v$ and $\theta \leq \theta_V$, then $F = -(1-\theta)[\pi(0,0) - \pi(v,v)] < 0$;

(c) if $r_I < v$ and $\theta > \theta_V$, then $F < 0$ if and only if $\theta < \theta^*$; and

(d) in all cases, *F* is an increasing function of θ.

(a) and **(b)** In these cases we know that $r = v$ and the fixed fee is set to make the downstream firm indifferent between accepting $r = v$ and deviating by infringing. This implies the indicated value of *F*, which is negative by the normality of the profit function. Using $F = -(1-\theta)[\pi(0,0) - \pi(v,v)] < 0$, $F'(\theta) = \pi(0,0) - \pi(v,v) > 0$, thus also proving (**d**) for these cases.

(c) Since the infringe constraint binds, $\pi(r(\theta), r(\theta)) - F(\theta) = \theta\pi(v, r(\theta)) + (1-\theta)\pi(0,0)$. Totally differentiating with respect to θ gives $F'(\theta) = [\pi(0,0) - \pi(v,r)] + r'(\theta)[\pi_1(r,r) + \pi_2(r,r) - \theta\pi_2(v,r)]$. Because $r \leq v$ and the profit function is normal, the first expression in brackets is strictly positive. Since $\pi_1(r,r) + \pi_2(r,r) < 0$ and $\pi_2 > 0$, the second expression in brackets is strictly negative. We have already shown that $r'(\theta) < 0$, so $F'(\theta) > 0$, thus also proving (**d**) for this case. Setting $F(\theta) = 0$ gives the defining expression for θ^*. If $r_I < v$ then $\theta^* < 1$.

Proof of Theorem 3

Observe that $H(\theta) = \max_r G(r;\theta)$ is the upper envelope of linear functions of θ; hence it is convex in general and linear where r does not vary with θ.

(1) If $r_I \geq v$ then $r = v$ for all $\theta > 0$ so H is linear in θ. Also, $\lim_{\theta \to 0} H(\theta) = T(v) - \pi(0,0) = T(v) - T(0) > 0$ (since T is increasing up to *m*), reflecting the opportunities for partial cartelization under the cover of (even) a flimsy patent. Therefore, the licensing payoff $H(\theta)$ is a straight line that starts above 0 and ends up at $H(1)$. The litigation payoff is a straight line starting at 0 and also ending up at $H(1)$. So the payoff from licensing is greater than the payoff from litigation (strictly for $\theta < 1$, and they are equal at $\theta = 1$).

(2) If $r_I < v$ then in the range $0 < \theta \leq \theta_V$, $H(\theta)$ is as discussed in part (1). For $\theta \geq \theta_V$, r varies, so $H(\theta)$ is a convex function of θ on $(\theta_V, 1]$. Therefore, if the $H(\theta)$ curve lies above the $\theta H(1)$ line as $\theta \to 1$, where the two meet, then $H(\theta) > \theta H(1)$ for all θ. But, since it is convex and begins above the line, the $H(\theta)$ curve lies above the $\theta H(1)$ line near $\theta = 1$ if and only if $H'(1) < H(1)$. Now $H'(1) = \pi(0,0) - \pi(v, r_I)$ and $H(1) = T(r_I) - \pi(v, r_I)$. So $H'(1) < H(1)$ if and only if $\pi(0,0) = T(0) < T(r_I)$. Since $T(r)$ is increasing in r for $r < m$, $T(r_I) > T(0)$ if and only if $r_I > 0$. If the optimal running royalty for an ironclad patent is negative, as it can be in the simple Cournot case, the patent holder would prefer to litigate a sufficiently strong patent rather than license without litigation. Weaker patents, however, are more profitably licensed without litigation, as are all patents if the running royalty for an ironclad patent is positive.

Proof of Theorem 4

Consumers: First consider the case $r_I \geq v$. In this case, if the patent is licensed without being fully litigated, the running royalty rate will be *v*. If the patent is litigated to judgment, the running royalty rate will be *v* if the patent is upheld, i.e., with probability θ, and zero with probability $1-\theta$. Litigation thus either does not affect running royalties or else sends them to zero; it thus is strictly better for consumers so long as $\theta < 1$. Next, consider the case $r_I < v$. If the patent is licensed under uncertainty (prior to a court ruling), the running royalty rate is $\min[v, r(\theta)]$. If the patent is litigated to judgment, the running royalty rate will be r_I if the patent is upheld and zero otherwise. Either way, litigation lowers running royalties and thus makes consumers better off.

Downstream Firms: A downstream firm gets its litigation payoff even in the licensing outcome, because the patent holder holds it to this reservation payoff. Thus each downstream firm's preference between licensing and litigation outcomes is determined by the level of its rivals' running royalties if the patent were to be litigated and upheld (actually in litigation, counterfactually in licensing). In the litigation outcome this is $\min[r_I, v]$, while in the licensing outcome it is $\min[r(\theta), v]$. Since $r_I < r(\theta)$, the downstream firm never prefers litigation; it will be indifferent if and only if $r_I \geq v$, since then its rivals will be paying v in either case.

Proof of Theorem 5

If $r_I \geq v$, then the running royalty is concave in patent strength: it is zero if $\theta = 0$ and then always equal to v for $\theta > 0$. If $r_I < v$, then the running royalty is zero if $\theta = 0$, equal to v in the range $0 < \theta \leq \theta_V$, and decreasing in θ for $\theta > \theta_V$. So if $r(\theta)$ is (weakly) concave over the range $[\theta_V, 1]$, it is weakly concave over all of [0,1].

Consumer surplus is a decreasing, convex function of the downstream price. If the downstream price is linear in the running royalty rate (as in Cournot oligopoly with constant marginal costs), consumer surplus $S(r)$ is a decreasing convex function of the running royalty r. If $r(\theta)$ is concave and $S(r)$ is convex and decreasing, then $S(r(\theta))$ is convex. To see this, consider any θ_1 and θ_2, and $\lambda \in (0,1)$; write $\theta_\lambda \equiv \lambda\theta_1 + (1-\lambda)\theta_2$. Since $r(\theta)$ is concave, $r(\theta_\lambda) \geq \lambda r(\theta_1) + (1-\lambda) r(\theta_2)$. Since $S(r)$ is decreasing, it follows that $S(r(\theta_\lambda)) \leq S(\lambda r(\theta_1) + (1-\lambda) r(\theta_2))$. Meanwhile, since $S(r)$ is convex, $S(\lambda r(\theta_1) + (1-\lambda) r(\theta_2)) \leq \lambda S(r(\theta_1)) + (1-\lambda) S(r(\theta_2))$. Putting these together, $S(r(\theta_\lambda)) \leq \lambda S(r(\theta_1)) + (1-\lambda) S(r(\theta_2))$, so $S(r(\theta))$ is convex in θ.

If consumer surplus is a convex function of patent strength, consumers are risk loving in patent strength. But ongoing litigation that reveals information about patent validity can be seen as inducing a mean-preserving spread in θ. (This mean-preserving spread goes all the way to the two-point distribution where $\theta \in \{0,1\}$ if the litigation proceeds to judgment.) Therefore, if the running royalty rate is a concave function of patent strength over the whole interval [0,1], consumers (expect to) benefit more and more as additional information about patent validity is revealed.

Proof of Theorem 8

Recall that negative fixed fees are only used over the range $[0, \theta^*]$, where θ^* was defined in the proof of Theorem 2. In this proof, we distinguish between patents for which negative fixed fees are used if permitted, $\theta < \theta^*$, and those for which negative fixed fees are not used even if permitted, $\theta \geq \theta^*$.

Profits and Deadweight Loss

If negative fixed fees are permitted, the patent holder's maximized payoff from licensing is $H(\theta) = r(\theta) x(r(\theta), r(\theta)) + F(\theta)$ where $F(\theta) = \pi(r(\theta), r(\theta)) - \theta\pi(v, r(\theta)) - (1-\theta)\pi(0,0)$. The deadweight loss is $D(r(\theta))$.

Define the patent holder's maximized payoff from licensing if negative fixed fees are banned as $J(\theta)$. For $\theta \geq \theta^*$, such fees are never used anyway, so $J(\theta) = H(\theta)$ and deadweight loss is $D(r(\theta))$. For $\theta < \theta^*$ the ban makes uniform royalties $s(\theta)$ the most profitable form of license, so $J(\theta) = s(\theta)x(s(\theta), s(\theta))$ and the deadweight loss is $D(s(\theta))$.

Constructing the Longer Patent Lifetime

We are now prepared to consider patent lifetimes. For convenience in this proof, we assume stationary market conditions. Write T for the status quo patent lifetime. The strategy of the proof is to construct an extended patent life $\tilde{T} > T$ such that banning negative fixed fees and extending the patent lifetime to $\tilde{T}$ leaves unchanged the overall expected return to obtaining a patent and lowers the overall expected deadweight loss.

In order to construct $\tilde{T}$, we need some new notation. Define $H(T_W, T_S)$ as the expected (present value) payoff to a patent if negative fixed fees are permitted, patents with $\theta < \theta^*$ have lifetime T_W, and patents with $\theta \geq \theta^*$ have lifetime T_S. In this notation, the expected present-value profits to a patent holder under the status quo are $H(T,T)$. Introducing $L(T) \equiv \int_0^T e^{-it}dt$ and $H_W \equiv \int_0^{\theta^*} H(\theta)dK(\theta)$, and $H_S \equiv \int_{\theta^*}^1 H(\theta)dK(\theta)$, where $K(\theta)$ is the *a priori* cumulative distribution function for patent strength, we have $H(T_W, T_S) = L(T_W)H_W + L(T_S)H_S$. Likewise, write $J(T_W, T_S)$ for the expected payoff if negative fixed fees are banned, and define J_W and J_S in the obvious way. Note that $H(\theta) = J(\theta)$ for $\theta \geq \theta^*$, so $H_S = J_S$.

Using these functions, we construct $\tilde{T} > T$, along with an auxiliary variable, $\hat{T} < T$, according to the following two equations: $H(T,T) = H(\hat{T}, \tilde{T})$ and $H(\hat{T}, \tilde{T}) = J(\tilde{T}, \tilde{T})$. The first equation states that overall patent profits are unchanged if (continuing to allow negative fixed fees) we extend the patent lifetime from T to $\tilde{T}$ for patents with $\theta > \theta^*$ but shorten it from T to $\hat{T}$ for patents with $\theta < \theta^*$. The second equation states that overall patent profits are unchanged if we then extend the patent lifetime from $\hat{T}$ to $\tilde{T}$ for patents in the range $\theta < \theta^*$, but ban negative fixed fees.

The first equation can be written as $L(T)H_W + L(T)H_S = L(\hat{T})H_W + L(\tilde{T})H_S$, which is equivalent to $[L(T) - L(\hat{T})]H_W = [L(\tilde{T}) - L(T)]H_S$. The second equation can be written as $L(\hat{T})H_W + L(\tilde{T})H_S = L(\tilde{T})J_W + L(\tilde{T})J_S$. Since $H_S = J_S$, this simplifies to $L(\hat{T})H_W = L(\tilde{T})J_W$. These are two linear equations in $L(\hat{T})$ and $L(\tilde{T})$. It is simple to check that they are non-collinear, so they give unique values for $\hat{T}$ and $\tilde{T}$.

Deadweight Loss is Reduced

We now show that if the patent lifetime is extended to $\tilde{T}$ and negative fixed fees are banned, deadweight loss is lower than in the status quo. Since patentee profits are, by construction, unchanged, this will complete the proof.

Define the flow deadweight loss from patents in the range $[0,\theta^*]$ if negative fixed fees are allowed as $D_W \equiv \int_0^{\theta^*} D(r(\theta))dK(\theta)$. Likewise, define the flow deadweight loss from patents in the range $[\theta^*,1]$, whether or not negative fixed fees are allowed, as $D_W \equiv \int_0^{\theta^*} D(r(\theta))dK(\theta)$. And define the flow deadweight loss from patents in the range $[0,\theta^*]$ if negative fixed fees are banned as $E_W \equiv \int_0^{\theta^*} D(s(\theta))dK(\theta)$. With these definitions, the deadweight loss in the status quo is given by $L(T)D_W + L(T)D_S$, and the deadweight loss if patent lifetime is extended and negative fixed fees are banned is $L(\tilde{T})E_W + L(\tilde{T})D_S$. Our goal is to show that $L(T)D_W + L(T)D_S > L(\tilde{T})E_W + L(\tilde{T})D_S$.

To prove this, we calculate the deadweight loss associated with an artificial intermediate regime, which involves extending the patent lifetime from T to $\tilde{T}$ for patents in the range $\theta > \theta^*$ but shortening the patent lifetime from T to $\hat{T}$ for patents in the range $\theta < \theta^*$ while still allowing negative fixed fees. The deadweight loss in this intermediate regime is $L(\hat{T})D_W + L(\tilde{T})D_S$. We now show that $L(T)D_W + L(T)D_S$ is larger than $L(\hat{T})D_W + L(\tilde{T})D_S$, which in turn is larger than $L(\tilde{T})E_W + L(\tilde{T})D_S$, completing the proof.

For this purpose, we make three observations about the ratios of the deadweight loss to profits. First, over the range $[0,\theta^*]$ the royalty rate, and therefore also deadweight loss, is non-increasing, but profits are increasing with θ, so the ratio of deadweight loss to profits, $D(r(\theta))/H(\theta)$, declines with θ in this range. Therefore, writing $\rho \equiv D(r(\theta^*))/H(\theta^*)$, we know that $D(r(\theta))/H(\theta) \ge \rho$ for all $\theta \le \theta^*$. Therefore, $D_W > \rho H_W$.

Second, for $\theta \le \theta^*$, $D(s(\theta))/J(\theta) = D(s(\theta))/s(\theta)x(s(\theta),s(\theta))$ is increasing in θ, since by assumption this ratio increases with the per-unit royalty rate $s(\theta)$, which increases with θ. Recall that $H(\theta^*) = r(\theta^*)x(r(\theta^*),r(\theta^*)) = J(\theta^*)$ since $F(\theta^*) = 0$ by definition. Therefore, $\rho = D(r(\theta^*))/J(\theta^*)$, so for all $\theta \le \theta^*$, $D(s(\theta))/J(\theta) \le \rho$. Therefore, $E_W < \rho J_W$.

Third, for $\theta \ge \theta^*$, $r(\theta) \le r(\theta^*)$ and $r(\theta)$ decreases with patent strength, so $D(r(\theta)) \le D(r(\theta^*))$ and deadweight loss declines with patent strength. Since the patent holder's profits increase with patent strength, the ratio of deadweight loss to patentee profits declines with patent strength. In particular, we must have $D(r(\theta))/H(\theta) < D(r(\theta^*))/H(\theta^*) = \rho$. Therefore, $D_S < \rho H_S$.

(1) Proof that $L(T)D_W + L(T)D_S$ is larger than $L(\hat{T})D_W + L(\tilde{T})D_S$. Rewriting, this is equivalent to $[L(T) - L(\hat{T})]D_W > [L(\tilde{T}) - L(T)]D_S$. Using $D_W > \rho H_W$, we have $[L(T) - L(\hat{T})]D_W > [L(\tilde{T}) - L(T)]\rho H_W$. Using $[L(T) - L(\hat{T})]H_W = [L(\tilde{T}) - L(T)]H_S$, this implies $[L(T) - L(\hat{T})]D_W > \rho[L(\tilde{T}) - L(T)]H_S$. Using $D_S < \rho H_S$, this in turn implies $[L(T) - L(\hat{T})]D_W > [L(\tilde{T}) - L(T)]D_S$.

(2) Proof that $L(\hat{T})D_W + L(\tilde{T})D_S > L(\tilde{T})E_W + L(\tilde{T})D_S$. Simplifying, this is equivalent to $L(\hat{T})D_W > L(\tilde{T})E_W$. Using $D_W > \rho H_W$, we have $L(\hat{T})D_W > \rho L(\hat{T})H_W$. Using $L(\hat{T})H_W = L(\tilde{T})J_W$, this implies that $L(\hat{T})D_W > \rho L(\tilde{T})J_W$. Using $E_W < \rho J_W$, this in turn implies that $L(\hat{T})D_W > L(\tilde{T})E_W$.

Proof of Theorem 9

Here we study $\frac{\pi_1}{\pi_1 + \pi_2}$ in the case of Cournot competition. This is the key ratio to explore, since we already know that $\frac{s'(0)}{v} = \frac{\pi_1(0,0)}{\pi_1(0,0) + \pi_2(0,0)}$.

With constant marginal costs and Cournot oligopoly, the first-order condition for firm *i* output choice is $p(X) + x_i p'(X) - c_i = 0$. Totally differentiating this, we get $[p'(X) + x_i p''(X)]dX + p'(X)dx_i - dc_i = 0$. Following the notation from Farrell and Shapiro (1990), we define $\lambda_i \equiv \frac{-p'(X) - x_i p''(X)}{-p'(X)}$, so with $dc_i = dr_i$ we have $\mathrm{d}x_i = -\lambda_i \mathrm{d}X + \frac{\mathrm{d}r_i}{p'(X)}$.
Writing $\Lambda \equiv \sum_i \lambda_i$ and adding up across all firms gives $\frac{\mathrm{d}X}{\mathrm{d}r_1} = \frac{1}{[1+\Lambda]p'(X)}$. Substituting for $\mathrm{d}X$ using this expression, we get $\frac{\mathrm{d}x_1}{\mathrm{d}r_1} = \frac{1+\Lambda-\lambda_1}{[1+\Lambda]p'(X)}$ and $\frac{\mathrm{d}x_j}{\mathrm{d}r_1} = \frac{-\lambda_j}{[1+\Lambda]p'(X)}$, $j \neq 1$.

For each firm $j \neq 1$, by the envelope theorem, the profit impact of a small increase in firm 1's running royalty is given by that firm's equilibrium output x_j times the change in price resulting from the equilibrium change in output by all *other* firms, $\mathrm{d}X - \mathrm{d}x_j$. This price change is given by $p'(X)[\mathrm{d}X - \mathrm{d}x_j]$, which equals $\frac{1+\lambda_j}{1+\Lambda}\mathrm{d}r_1$. Since this expression does not contain any parameters specific to firm 1, the effect on firm *j*'s profits of a small increase d*r* in *all other* *firms'* running royalties is given by $(N-1)x_j \frac{1+\lambda_j}{1+\Lambda}\mathrm{d}r$. Returning to our main notation, we therefore have $\pi_2 = (N-1)x_j \frac{1+\lambda_j}{1+\Lambda}$.

Similarly the effect on firm 1's profits of a small increase $\mathrm{d}r_1$ in its own running royalty is equal to the direct cost effect, $-x_1\mathrm{d}r_1$, plus the effect of the price change caused by other firms' output changes, $x_1 p'(X)[\mathrm{d}X - \mathrm{d}x_1] = -\frac{\Lambda-\lambda_1}{1+\Lambda}\mathrm{d}r_1$. Therefore $\pi_1 = -x_1 \frac{1+2\Lambda-\lambda_1}{1+\Lambda}$.

Putting these together, starting at a symmetric equilibrium where each $\lambda_i = \lambda$ and $x_i = x_j$, and simplifying, we get

$$\frac{\pi_1}{\pi_1+\pi_2} = \frac{1+(2N-1)\lambda}{2+N\lambda-N} = N\left[1+\frac{N(1-\lambda)}{2-N(1-\lambda)}\right].$$

In a symmetric equilibrium, we also have $\lambda_i \equiv \frac{-p'(X)-Xp''(X)/N}{-p'(X)} = 1+Xp''(X)/Np'(X)$. Writing $E \equiv -Xp''(X)/p'(X)$ for the elasticity of the slope of the inverse demand curve, we have $\lambda = 1-E/N$ or $E = N(1-\lambda)$. See Shapiro (1989) for a further discussion of the role of "Seade's" E in Cournot comparative statics. Hence, we obtain $\frac{\pi_1}{\pi_1+\pi_2} = N\frac{2}{2-E}$, or equivalently,

$$\frac{\pi_1}{\pi_1+\pi_2} = N\frac{2}{2+Xp''(X)/p'(X)}.$$

Note that if demand is linear or convex, $p''(X) \geq 0$, then $E \geq 0$ and $\frac{\pi_1}{\pi_1+\pi_2} \geq N$. For linear demand, $E=0$, so $\frac{\pi_1}{\pi_1+\pi_2} = N$. When demand has constant elasticity $\varepsilon > 1$ ($\varepsilon > 1$ is the regularity condition for $\pi_1+\pi_2 < 0$), we have $E = 1+\frac{1}{\varepsilon}$, so $\frac{\pi_1}{\pi_1+\pi_2} > N$.

Licensing vs. Litigation with Weak Patents if Negative Fixed Fees Are Banned

Licensing gives the patent holder a payoff of $s(\theta)x(s(\theta), s(\theta))$ for a patent that would be licensed using negative fixed fees if permitted.

If $r_I \geq v$, then litigation gives the patent holder $\theta vx(v,v)$. Licensing is better for the patent holder if and only if $s(\theta)x(s(\theta),s(\theta)) > \theta vx(v,v)$, which can be written as $s(\theta)/\theta v > x(v,v)/x(s(\theta),s(\theta))$. Since output declines with cost, the right-hand side is less than unity, while Theorem 9 tells us that $s(\theta)/\theta v > 1$. So licensing is preferred to litigation in this case. In fact, in the Cournot case in Theorem 9, licensing gives a payoff that is about *N* times as large as litigation, suggesting a considerable margin in this argument.

If $r_I < v$, then litigation gives the patent holder $\theta[\pi(r_I,r_I) + r_I x(r_I,r_I) - \pi(v,r_I)]$. Writing $s(\theta) = k\theta v$, licensing is preferred to litigation if and only if $kvx(s,s) > \pi(r_I,r_I) + r_I x(r_I,r_I) - \pi(v,r_I)$. Rearranging, and using the fact that $r_I > s$ so $x(r_I,r_I) < x(s,s)$, a sufficient condition for this to hold is that $(k-1)vx(r_I,r_I) > \pi(r_I,r_I) - \pi(v,r_I)$. Using the intermediate value theorem, the right-hand side of this expression equals $-(v-r_I)\pi_1(z,r_I)$ for some $z \in [r_I,v]$. In the special case of Cournot with linear demand and constant marginal costs, $\pi_1(z,r_I) = -2x(z,r_I)N/(N+1)$, so the sufficient condition becomes $(k-1)vx(r_I,r_I) > 2(v-r_I)x(z,r_I)N/(N+1)$. Since $k \approx N$ in this case, this condition can be approximated by $(N-1) > 2\dfrac{(v-r_I)}{v}\dfrac{x(z,r_I)}{x(r_I,r_I)}\dfrac{N}{N+1}$ which must be met if $N \geq 3$.

Proof of Theorem 10

We characterize here the equilibrium two-part tariff offered by the patent holder when negative fixed fees are allowed and we assume the "lax regime" so that the patent holder and licensee can agree on royalties outside the scope of the patent. Analytically, this changes two things. Directly, it implies that we no longer impose the constraint that $r \leq v$. Less obviously, it implies that a downstream firm may prefer using the backstop technology to infringing and challenging the patent.

If its rivals accept licenses with running royalty *r*, then a downstream firm who rejects a license gets $\pi(v,r)$ by using the backstop technology, and $\theta\pi(v,r) + (1-\theta)\pi(0,0)$ by infringing and triggering litigation. The two are equal if $\theta = 1$; if $\theta < 1$ then backstop gives a higher payoff than infringe if and only if $\pi(v,r) > \pi(0,0)$, or if and only if $r > r_0$ where r_0 is defined by $\pi(v,r_0) = \pi(0,0)$. Observe that (by normality) $r_0 > v$.

Because the two reservation payoffs are equal if $\theta = 1$, the patent holder's maximization problem if backstop is the only binding alternative is the same as it would be with an ironclad patent. (Since litigation will never arise in this case, this is not surprising.) Consequently, if a downstream firm would strictly prefer backstop versus infringing in equilibrium, the patent holder maximizes $G(r,1)$, which involves setting $r = r_I$. Thus if the equilibrium running royalty strictly exceeds r_0, it must be equal to r_I.

Indeed, if $r_I \geq r_0$, then $r = r_I$ for all θ. This is analogous to Figure 2 in the strict patent-scope regime. If we considered an alternative candidate *r* at which only infringe binds (that is, $r < r_0$), the patent holder would prefer to change *r* locally unless $r = r(\theta)$; but recall that $r(\theta) \geq r_I$.

Suppose then that $r_I < r_0$. Then no $r > r_0$ can be optimal, since locally the patent holder would be optimizing against the backstop constraint and would thus want to move *r* towards $r_I < r_0$. Hence the infringe constraint binds in equilibrium. Since $r = r(\theta)$ maximizes the patent holder's payoff given that constraint, it is chosen unless it exceeds r_0, in which case r_0 is chosen.

Summarizing, if $r_I < r_0$ then $r = \min[r(\theta), r_0]$; if $r_I \geq r_0$ then $r = r_I$.

Note that for very weak patents, i.e., as $\theta \to 0$, we have $r(\theta) > r_0$ and so $r = \max[r_0, r_I]$. In other words, the running royalty rate for very weak patents is at least as large as the running royalty rate for an ironclad patent, and can readily be larger.