# UC Berkeley
## Recent Work

**Title**
How Strong Are Weak Patents?

**Permalink**
https://escholarship.org/uc/item/8vg425vj

**Authors**
Farrell, Joseph
Shapiro, Carl

**Publication Date**
2007

# How Strong Are Weak Patents?[*]

## Joseph Farrell and Carl Shapiro[†]

## January 2007

**ABSTRACT**.  We analyze patent licensing by a patent holder to downstream technology users. We study how the structure and level of royalties depends on the patent's strength, i.e., the probability it would be upheld in court.  We examine the social value of determining patent validity before licensing, in terms of deadweight loss (*ex post*) and innovation incentives (*ex ante*).  When downstream users do not compete against each other or the patent holder, license fees approximate the license fee for an ironclad patent times the patent strength, and reviewing validity before licensing would be unproductive (in expected value).  But when downstream users compete, two-part tariffs for weak patents have high running royalty rates, combined with a negative fixed fee, and examining patent validity generates social benefits, both *ex post* and *ex ante*.  Even without negative fixed fees, rival downstream firms will accept relatively high running royalties, so determining patent validity prior to licensing is socially beneficial.

**Keywords**: probabilistic patents, weak patents, patent licensing, patent reform, oligopoly.

# 1. Introduction

Economists traditionally view a patent as an ironclad property right: others cannot use the patented technology without a license. In reality, however, to stop another party from using the patented technology, a patent holder typically must go to court and prove the patent is both valid and infringed, and this is by no means always clear. A patent is thus not a clear right to exclude but rather a right to sue for infringement and, if successful, the right to be awarded damages for proven infringement and typically to obtain an injunction against future infringement.

Because patent rights are probabilistic, the economic impact of an issued patent depends upon *patent strength*, i.e., the probability $\theta$ that it would be found valid and infringed if tested in court. We study how patent impact varies with patent strength, a relationship that is fundamental to evaluating the operation and potential reform of the patent system.

Evidence has mounted in recent years that many issued patents are questionable or "weak," and might well be found invalid if vigorously litigated.[1] Some observers argue that weak patents constitute undeserved monopolies, and that we should reform the patent process to weed them out. Optimists respond that, if enforced at all, weak patents are licensed at commensurately low royalty rates, because licenses are negotiated in the shadow of infringement litigation. Furthermore, as Mark Lemley (2001) stresses, it would be very costly for the PTO to scrutinize all patent applications as thoroughly as courts examine the relatively few litigated patents.

If a patent's validity will be tested in court before licensing, it has a market impact only if it turns out to be valid. Viewed at a date when it is still probabilistic, such a patent's expected impact, apart from litigation costs and effects during litigation, is thus proportional to its strength $\theta$.[2]

In fact, however, far more patents are licensed, either without litigation at all or to settle litigation, than are litigated to a final judgment.[3] When a patent is licensed in the shadow of

---

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

[2] In this paper, we abstract away from the interim social costs of probabilistic patents that arise during the pendency of patent litigation. We also assume the patent holder cannot behave opportunistically towards downstream firms that make investments specific to the patented technology. Shapiro (2006) studies patent hold-up.

litigation, how does its strength affect the terms on which it is licensed? That is the central topic of this paper. We use a simple game-theoretic model of the licensing of a probabilistic patent: the patent holder offers licenses to downstream firms, each of whom can accept the license, avoid using the patented technology, or infringe, prompting litigation. This naturally generalizes models of the licensing of ironclad patents, in which a downstream firm can either accept the offered license or avoid using the patented technology.[4] Our model assumes that litigation costs are zero, but nevertheless predicts licensing without litigation.

For patents licensed to downstream firms that do not compete against each other or against the patent holder, our model supports the optimistic perspective: weak patents generate little profits and little deadweight loss. But when downstream firms use the patented technology in competing against each other or against the patent holder, licensing interacts with that competition in two powerful ways. First, agreeing to per-unit royalties raises the joint profits of the patent holder and licensees by elevating the downstream price, moving it closer to the monopoly price. We show that this joint profit motive for high per-unit royalty rates prevails for weak patents if licenses can use unrestricted two-part tariffs. Second, a downstream firm's decision to litigate benefits other downstream firms as well as consumers, since the litigation may invalidate the patent. This force is dominant in the licensing of weak patents if the patent holder uses linear licenses or two-part tariffs with fixed fees that are restricted to be non-negative. As a result of this positive externality, incentives to challenge patents are sub-optimal, and downstream firms will accept surprisingly large per-unit royalties.

We focus on $r(\theta)$, the per-unit royalty rate at which a patent of strength $\theta$ will be licensed, because it governs licensees' marginal cost of using the patented technology and thus drives the deadweight loss associated with the patent. We show that the optimistic view requires that the per-unit royalty rate and the patent holder's profits are (roughly) proportional to patent strength,

---

[3] Lemley (2001) estimates that about 5% of all patents are either licensed without litigation or are litigated., and that only 0.1% of all patents are litigated to trial, so roughly 50 times as many patents are licensed (without litigation or to settle litigation) as are litigated to trial. Kimberly Moore (2000) reports that the percentage of patent cases going to trial has declined over time, to 3.3% by 1999 (Table 1). Jay Kesan and Gwendolyn Ball (2006) conclude that patent litigation is largely a settlement mechanism; about 10% of patent cases filed in 2000 led to rulings and verdicts (Table 6).

[4] See Morton Kamien (1992) for a review of this literature and Debapriya Sen and Yair Tauman (forthcoming) for a more recent contribution.

i.e., $r(\theta) \approx \theta r(1)$ and $P(\theta) \approx \theta P(1)$, but we show that for weak patents licensed to downstream rivals $r(\theta)$ is a large multiple of $\theta r(1)$ and $P(\theta)$ is a large multiple of $\theta P(1)$.

This suggests that the benefit of more careful review at the U.S. Patent and Trademark Office (PTO) depends on competitive conditions among prospective licensees as well as on the patented technology. Such enhanced review can occur either before or after a patent issues; strengthening post-grant review of significant patents is a prominent part of patent reform proposals currently being considered in Congress. We confirm the role of competition, both *ex post* and *ex ante*.

*Ex post*, assume that the innovation has been made and that a patent of strength $\theta$ would be issued under normal review. How does further review affect the deadweight loss from the patent? For this analysis, $\theta$ need only be a commonly known probability; the relationship between patent strength and actual innovation does not matter. Unlike popular commentators, we do not assume that enhanced review would simply eliminate some "bad" patents without affecting others. Such a view is inconsistent with Bayesian statistics. Rather, given a set of patent applications, closer scrutiny brings out more information, inducing a mean-preserving spread on $\theta$: what does not kill a patent makes it stronger. For example, starting with $\theta = 0.4$, further scrutiny might lead with 50% probability to a patent of strength 0.7 and with 50% probability to a patent of strength 0.1. In the extreme, "ideal" PTO review that replicates judicial review would lead either to an ironclad patent (with probability $\theta$) or to no patent (with probability $1 - \theta$). The expected *ex post* benefit of such review is
$B(\theta) \equiv [\theta W(1) + (1-\theta)W(0)] - W(\theta)$, where $W(\theta)$ is the welfare resulting when a patent of strength $\theta$ is licensed. We show that $B(\theta)$ is small (and sometimes negative) for patents licensed to downstream firms that do not compete, but large and positive for weak patents licensed to downstream rivals.

*Ex ante*, the patent system seeks to encourage innovation; how does enhanced PTO review affect innovation incentives? We assume that courts truly assess patent validity, so that $\theta$ is the probability that the patent holder actually contributed to society the technology covered by the patent, rather than getting a wrongly issued or overly broad patent on prior art or obvious technology. We analyze the expected social contribution $K(\theta)$ made by the owner of a patent of strength $\theta$. If the downstream firms compete and licenses can use unrestricted two-part tariffs, we find that $K(\theta) < 0$ for weak patents, distorting *ex ante* incentives for research and patenting.

Even when $K(\theta) > 0$, *ex ante* incentives are distorted if $P(\theta) > K(\theta)$ and more generally if the ratio of private to social returns, $P(\theta)/K(\theta)$, varies substantially with $\theta$, as we show is typical for patents licensed to downstream rivals. For such patents, enhanced PTO review not only yields *ex post* benefits $B(\theta) > 0$ but also eliminates some distortions in *ex ante* incentives to engage in research and apply for patents.

Section 2 presents our licensing model for probabilistic patents. Section 3 establishes results for evaluating *ex post* and *ex ante* welfare effects of enhanced PTO review. For patents that are licensed to downstream firms that do not compete against each other or against the patent holder, Section 4 shows that $r(\theta) = 0$ for all $\theta$, so enhanced PTO review generates no *ex ante* or *ex post* social benefits. In contrast, for patents licensed to multiple downstream rivals, Section 5 shows that $r(\theta) \geq r(1)$, so $r(\theta)$ must exceed the benchmark level $\theta r(1)$, especially for weak patents; consequently ideal PTO review generates *ex ante* and *ex post* social benefits. When $\theta$ is small, downstream firms accept this high running royalty because it increases joint profits and the patent owner shares the increase with downstream firms through a *negative* fixed fee $F(\theta) < 0$.

If such negative fixed fees are not feasible, weak patents will be licensed using linear licenses consisting simply of a per-unit royalty, $s(\theta)$. Section 6 shows that $s(\theta)$ for weak patents far exceeds the benchmark $\theta s(1)$ when licensees compete. In Cournot competition, the ratio $s(\theta)/\theta s(1)$ is of order $N$, the number of downstream firms, for very weak patents. Section 7 generalizes our results the case of a vertically integrated patent holder that competes against its licensees, and explores the effects of relaxing some of our assumptions. Section 8 concludes.

## 2. Patent Licensing in the Shadow of Litigation

### A. *Technology and Licensing Game*

An upstream patent holder P offers licenses to *N* symmetric downstream firms. The patented technology lowers downstream firms' unit production costs by *v,* the *patent size*, relative to the best alternative, or *backstop*, technology. Equivalently, the technology makes each unit of the product worth an extra *v* to all customers. For now, we assume that the patent holder does not compete against the downstream firms.

In general, when an upstream monopolist sells an input to downstream firms that compete, complex multi-lateral contracting issues arise. The vertical contracting literature has shown that equilibrium depends heavily on the form of contracts allowed, on downstream firms' information, and on their beliefs about what they cannot observe.[5] If the upstream supplier can commit to arbitrary contingent contracts, it can organize a hub-and-spoke downstream cartel supporting the monopoly price downstream, even if it controls only a minor input.[6] At the other extreme, if contracts are private, the upstream supplier may be unable to charge any price above marginal cost for its input.[7] Finding those extreme outcomes unrealistic, we adopt a simple licensing model: the patent holder offers a two-part tariff $[F, r]$ to all downstream firms.[8] Such non-discriminatory offers are sometimes used in practice, are prominent in the ironclad patent licensing literature (see e.g. Morton Kamien (1992)), and are typically required for the licensing of patents incorporated into industry standards.[9]

With an ironclad patent, each downstream firm accepts the offered license or uses the backstop technology. Here, a downstream firm that rejects a license has another option: infringing the patent. In that case, we assume that the patent holder sues the infringer.[10] If the patent is held invalid, all downstream firms can use the technology free of charge.[11] Alternatively, if the patent

---

[5] See R. Preston McAfee and Marius Schwartz (1994) and Ilya Segal (1999).

[6] The upstream firm can sell its input at a price that supports the downstream monopoly price and threaten to subsidize all other downstream firms if any one downstream firm does not buy its input.

[7] This occurs if negotiations are private and each downstream firm has "passive beliefs"—does not adjust its beliefs about other firms' contracts when offered a new contract. In this case, the upstream firm negotiates the bilaterally efficient contract with each downstream firm, which involves a price equal to its marginal cost. See McAfee and Schwartz (1994) and Patrick Rey and Thibaud Vergé (2004).

[8] Restricting the number of licenses offered can be optimal for an ironclad patent; see Michael Katz and Shapiro (1986), Chun-Hsiung Liao and Debapriya Sen (2005), and Sen and Yair Tauman (forthcoming). However, this approach does not work as a licensing strategy for a probabilistic patent, since firms that do not receive licenses will infringe the patent. If P sues those firms, the equilibrium involves litigation (considered below). If P ignores infringing firms, downstream firms will be unwilling to pay for licenses.

[9] Benjamin Chiao, Josh Lerner and Jean Tirole (2006) study the rules of 59 standard-setting organizations; about 75% require essential patents to be licensed on fair, reasonable, and non-discriminatory terms. Licensing programs in standard-setting contexts that have attracted antitrust attention include those of Rambus (computer memory) and Qualcomm (mobile telephones).

[10] It is not always clear that P will want to sue an infringing firm, especially if others have signed licenses and litigating would put their royalty payments at risk. The patent holder might prefer, *ex post*, quietly to ignore an infringer. We discuss litigation credibility further in Section 7.

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

is ruled valid, we assume that any licenses already signed remain in force,[12] and that the patent holder can negotiate anew with the downstream firm(s) that lack licenses. Lastly, the downstream firms compete, given the licenses they have signed and the technologies they use.

## B. Downstream Oligopoly

The downstream oligopoly equilibrium depends on downstream firms' marginal costs. For ease of notation, we measure each firm's marginal cost relative to the cost $c$ that it would incur using the patented technology free of charge. With this notation, a firm that accepts a license with per-unit royalty $r$ has marginal cost $r$, and a firm using the backstop technology has marginal cost $v$. To analyze a symmetric equilibrium we need only consider the profits of one firm with costs $a$ when all other firms have costs $b$; write $x(a,b)$ for its output and $\pi(a,b)$ for its profits, net of running royalties but gross of any fixed fee. We assume $\pi(a,b)$ satisfies three mild conditions: (1) $\pi_1(a,b) < 0$: a firm's profits are decreasing in its own costs; (2) $\pi_2(a,b) \geq 0$: a firm's profits are non-decreasing in the other firms' costs; and (3) $\pi_1(a,a) + \pi_2(a,a) < 0$: each firm's profits fall if all firms' costs rise in parallel. Writing $p(a)$ for the price charged by each downstream firm if all firms have cost $a$, in a wide range of simple oligopoly models $p(a)$ is linear.[13] In the text, we assume this; the Appendix shows where we actually rely on this assumption.

## C. Optimal Two-Part Tariffs

Suppose that downstream firm D expects all its rivals to accept the offer $[F,r]$. If it too accepts, its payoff is $\pi(r,r) - F$. If it rejects the offer, infringes, is sued, and the patent is upheld, P would hold D down to its backstop payoff $\pi(v,r)$. Thus D's reservation payoff is $\theta\pi(v,r) + (1-\theta)\pi(0,0)$.[14] Figure 1 displays the game tree for this licensing game, simplified to focus on just one downstream firm. If P opts to avert litigation, it will set the largest $F$ such that it is a subgame equilibrium for all downstream firms to accept $[F,r]$: thus $F(\theta) = \pi(r,r) - \theta\pi(v,r) - (1-\theta)\pi(0,0)$. Writing total profits (divided by $N$) as

---

[12] We revisit this assumption in Section 7.

[13] Linear pass-through (meaning that price as a function of the royalty rate is a straight line, not necessarily through the origin) holds in Cournot oligopoly with linear demand or constant elasticity demand, in the standard Hotelling duopoly model, and in differentiated-product Bertrand oligopoly with linear demand, among others.

$T(r) \equiv rx(r,r) + \pi(r,r)$, let $r(\theta)$ maximize P's payoff per downstream firm,

$G(r,\theta) \equiv T(r) - \theta\pi(v,r) - (1-\theta)\pi(0,0)$, subject (we assume) to $0 \le r \le v$.[15]

Let $H(\theta) \equiv G(r(\theta),\theta)$ denote the patent holder's resulting no-litigation payoff. That will be the overall equilibrium if P prefers it to litigating the patent. Litigating gives P zero if the patent is declared invalid and $H(1)$ if it is upheld, for an expected payoff of $\theta H(1)$. Since P chooses between licensing and litigation, its payoff is $P(\theta) \equiv \max[H(\theta), \theta H(1)]$. Since $P(1) = H(1)$, we have $P(\theta) \ge \theta P(1)$: the patent owner can always get a fraction $\theta$ of the payoff from an ironclad patent by litigating.

## 3. Welfare Analysis of Probabilistic Patent Licenses

Before solving for the equilibrium two-part tariff in various settings, we develop welfare tools to evaluate the expected benefits of enhanced PTO review for patents that will be licensed rather than litigated.[16] Section 5 below shows that such licensing indeed arises in our model. Figure 2 displays a simplified game tree for licensing with ideal PTO review.

### A. *Ex Post Analysis*

Given that the innovation has been made and a patent of strength $\theta$ issued, *ex post* welfare $W(\theta)$ is the sum of the patent holder's licensing revenues, the downstream firms' profits, and the surplus enjoyed by final consumers. If all *N* downstream firms accept licenses with running royalty *r*, *ex post* welfare depends only on *r*, since fixed fees are just transfers: $W(\theta) = w(r(\theta))$,

---

[14] Using the backstop technology would yield $\pi(v,r) \le \pi(v,v) < \pi(0,0)$, which is less attractive than litigating.

[15] A downstream firm might accept a running royalty rate $r > v$ combined with $F < 0$. However, under patent law a license can impose royalties only for use of the patented technology. We assume that this rule is effectively enforced. This implies $r \le v$, since even a downstream firm that signed a license would use the backstop technology rather than pay $r > v$ to use the patented technology. If the rule is not well enforced, P can bribe each downstream firm with a negative fixed fee to accept a royalty rate $r > v$ (on all output) that supports the downstream monopoly price. In a previous version of this paper we showed that for sufficiently weak patents this is the equilibrium, and that $r \ge 0$ for weak patents even if $r < 0$ is allowed. We assume that both *T* and *G* are single-peaked in *r* on $[0,v]$.

[16] In our model, if the patent would be litigated, ideal PTO review has no benefit, since the patent's validity will be determined before licensing anyway. Outside our model, the PTO might hold a comparative advantage over the

where $w(r)$ denotes *ex post* welfare with royalty $r$; $w'(0) < 0$ with imperfect downstream

competition. Define $\lambda_{max} \equiv \max_{0 \le t \le v} |w'(t)|$, $\lambda_{min} \equiv \min_{0 \le t \le v} |w'(t)| > 0$, and $\mu \equiv \dfrac{\lambda_{max}}{\lambda_{min}}$.

Ideal PTO review (or litigation prior to licensing) gives expected welfare $\theta w(r(1)) + (1-\theta)w(0)$,

so its expected benefit is $B(\theta) = [\theta w(r(1)) + (1-\theta)w(0)] - w(r(\theta))$. The Appendix proves:

**Theorem 1.** $B(\theta) \ge [r(\theta) - \mu\theta r(1)]\lambda_{min}$ .

Theorem 1 implies that $B(\theta) > 0$ if $\dfrac{r(\theta)}{\theta r(1)} > \mu$. This justifies the intuitive benchmark $\theta r(1)$, in

that $B(\theta) > 0$ if $\mu \approx 1$, as is the case for small $v$, and $r(\theta)$ non-trivially exceeds the benchmark.

For larger innovations, $\mu$ may not be close to 1, but the Appendix shows that, for instance,

$\mu \le \dfrac{x(0,0)}{x(v,v)}\left[1 + \dfrac{v}{p(v) - v - c}\right]$ in Cournot oligopoly. By comparison, with $N \ge 2$, we identify

below cases where $\dfrac{r(\theta)}{\theta r(1)} \approx N$, so $B(\theta) > 0$ for a wide range of patent sizes.

Theorem 1 casts the *ex post* analysis in terms of total welfare, but our model ignores litigation

costs, which are borne by the patent holder and downstream firms. Given the PTO's review

standards, those parties can choose whether or not the patent is litigated, but consumers cannot,

so an externality-inspired approach would consider the effects of PTO review (or litigation) on

consumers. With our assumption that $p(r)$ is linear, the Appendix proves:

**Theorem 2**. **If $r(\theta) \ge \theta r(1)$ then consumers benefit from ideal PTO review.**

In the cases identified below where $r(\theta) > \theta r(1)$, the downstream firms have too little incentive

to challenge a weak patent.[17] In such cases, consumers will value the right to trigger patent re-

---

courts in patent review: for instance, PTO review could be faster, reducing social costs borne in the interim before a patent validity ruling.

[17] Jay Pil Choi (2002, 2005) argues that patent holders have weak incentives to challenge one another's patents if multiple weak patents are contributed to a patent pool. Our focus is instead on challenges by direct purchasers of the patented technology (downstream firms). Direct purchasers seem more likely to have legal standing, and although we are not aware of systematic evidence, we suspect that most patent licenses do not involve patent pools.

examination.[18]  Theorems 1 and 2 apply regardless of the mechanism that determines $r(\theta)$ and

$r(1)$, and do not depend on our specific licensing model.

## B. Ex Ante Analysis

A firm's private incentive to engage in the R&D and patenting activities that lead to a patent of

strength $\theta$ is $P(\theta)$.  How does this compare to the social contribution $K(\theta)$?

We assume that $\theta$ is the true probability that the patent holder contributed the patented

technology to society.[19]  Denoting by $\overline{W}$ the welfare that would result if the patented technology

were not available to society, $K(\theta) \equiv W(\theta) - [(1-\theta)W(0) + \theta\overline{W}]$.  The Appendix proves:

**Theorem 3: If** $r(\theta) > \theta v$ **then** $P(\theta) > K(\theta)$.

That is, if royalties exceed an intuitive benchmark, the patent holder's private return exceeds its

social contribution.  In expectation, the patent holder has inflicted a negative externality on

others; marginally profitable activities leading to such patents lower expected welfare.[20]

Even if $P(\theta) < K(\theta)$, the relative incentives to pursue patents of different strengths may be

biased.  Consider a firm allocating its R&D and patenting budget between two activities.  The

first activity is a "conventional" line of research that, if it succeeds technically, will produce a

useful but unsurprising technology, so there may already be prior art or a court may later deem

the invention obvious.  Thus, this activity generates patents of strength $\theta < 1$.  A second, more

creative line of research, if technically successful, will generate truly novel and non-obvious

results, leading to ironclad patents.  The firm will allocate its R&D budget based on the relative

---

[18] The Electronic Frontier Foundation has a "patent-busting project" (www.eff.org/patent) that seeks to overturn some Internet and software-related patents by gathering prior art and requesting re-examination by the PTO.  See Ian Austen, "Claiming a Threat to Innovation, Group Seeks to Overturn 10 Patents," *New York Times*, July 5, 2004. Joseph Miller (2004) advocates rewarding firms that successfully invalidate patents.

[19] Reiko Aoki and Jin-Li Hu (1999) analyze *ex ante* incentives of probabilistic patents, but focus entirely on the dilution (and other changes) of incentives for true innovation that will result only in a weak patent; James Anton and Dennis Yao (2003) have a similar focus.  Their approach assumes that when a patent is held invalid or not infringed, it is a court (or legal system) error. Alan Marco (2006) attempts to estimate the frequency of such errors if financial markets always get it right.  By contrast, we assume that when a patent is held invalid, it is because the court's thorough scrutiny shows that it truly did not represent a novel, useful, non-obvious contribution of the patent holder.

[20] This can happen even with ironclad patents: in our linear Cournot example $P(1)$ typically exceeds $K(1)$.

reward to the two kinds of patents, $P(\theta)/P(1)$. For social efficiency, the allocation should be based on the relative contributions, $K(\theta)/K(1)$. If $P(\theta)/P(1) > K(\theta)/K(1)$, the firm will devote too much of its budget to the conventional line of research. In an extreme case, the firm may do little actual R&D and devote most of its resources to applying for patents covering technologies that very likely are already known or are obvious, if such weak applications often yield valuable weak patents rather than rejections. Under ideal PTO review, weak applications mostly yield rejections and occasionally yield ironclad patents; the expected contribution and reward to a technical success in the conventional research line thus become $\theta K(1)$ and $\theta P(1)$ respectively, eliminating the bias. The Appendix proves:

**Theorem 4**. **If** $B(\theta) > 0$**, then** $P(\theta)/P(1) > K(\theta)/K(1)$.

Theorem 4 links *ex post* and *ex ante* analysis: if there are *ex post* benefits of ideal PTO review, then there is also an *ex ante* bias toward seeking weak patents, which such review eliminates.

## 4. Downstream Firms That Do Not Compete

Suppose the downstream firms operate in separate markets, so each firm's profits do not depend on others' costs: $\pi_2(a,b) \equiv 0$. Then running royalties would reduce profits through double marginalization and $T(r)$ is maximized at $r = 0$. Nor would running royalties help P extract a bigger share of joint profits, since, simplifying notation, a downstream firm's expected payoff from litigation is $\theta\pi(v) + (1-\theta)\pi(0)$, independent of $r$, so $G(r,\theta) = T(r) - [\theta\pi(v) + (1-\theta)\pi(0)]$. Thus, $r(\theta) = 0$ for all $\theta$. It follows that $B(\theta) = 0$: further PTO (or judicial) review of validity does not affect total welfare or consumer surplus. Moreover, $F(\theta) = \theta F(1)$, for otherwise either D or P would prefer litigation. Hence P's payoff is proportional to $\theta$: specifically, $P(\theta) = \theta[\pi(0) - \pi(v)]$. Indeed, nobody cares in expectation, *ex post*, whether the patent is licensed under uncertainty or further reviewed before licensing. *Ex ante,*

$K(\theta) \equiv \theta[W(0) - \overline{W}] > 0$, so $\dfrac{P(\theta)}{K(\theta)} = \dfrac{\pi(0) - \pi(v)}{W(0) - \overline{W}}$, which is between 0 and 1 and is independent

of $\theta$, so incentives are not biased among R&D and patenting strategies that lead to patents of different strengths. Summarizing, we have a reassuring benchmark:

**Theorem 5: If $r(\theta) = 0$ for all $\theta$, additional PTO review of patents generates no *ex ante* or *ex post* benefits. In our model, if downstream firms are not rivals, then $r(\theta) = 0$ for all $\theta$.**

## 5. Downstream Firms That Compete

When downstream firms compete against one another, their prices fall below the joint profit maximizing level. As a result, total profits (including P's) rise when all downstream firms face a small positive running royalty $r$. If the downstream industry is reasonably competitive, the running royalty $m$ that supports the downstream monopoly price, maximizing total profits $T(r)$, can be large. We assume that $m \geq v$, so $r = v$ maximizes $T(r)$ in the feasible range $0 \leq r \leq v$.[21]

Since fixed fees allow P to capture joint profits $T(r)$ minus downstream firms' reservation payoffs, one might thus expect $r(\theta) = v$. That is correct for small $\theta$, although not in general, because P can lower each downstream firm's reservation payoff $\theta\pi(v, r) + (1-\theta)\pi(0,0)$ by lowering the running royalty rate $r$ to the firm's rivals.[22] P sets $r$ to maximize $G(r, \theta)$, not $T(r)$, and $G_r(r, \theta) = T'(r) - \theta\pi_2(v, r) < T'(r)$. But this rent-shifting effect is proportional to $\theta$, since an infringer faces rivals with marginal cost $r$ only if the patent is found valid. For weak patents, rent-shifting does not much modify joint profit maximization and sharing:[23] for $\theta \leq \theta_V \equiv T'(v)/\pi_2(v, v)$, we have $G_r(v, \theta) \geq 0$ and $r(\theta) = v$.[24] A downstream firm accepts this license because $F(\theta) = -(1-\theta)[\pi(0,0) - \pi(v,v)] < 0$.[25] As the Appendix shows, the patent holder strictly prefers such licensing to litigation. In sharp contrast with Theorem 5, we thus have:

---

[21] $m \geq v$ if and only if the downstream price charged by an integrated monopolist using the new technology is no lower than the oligopoly equilibrium price with the old technology: $p(m) \geq p(v)$. With even moderate downstream competition, this will hold for quite substantial innovations.

[22] This rent-shifting effect is recognized in the literature on the licensing of ironclad patents; see Sen and Tauman (forthcoming). Segal (1999) studies this effect much more generally.

[23] Rent-shifting would be a big impediment to cartelizing an industry without a patent, because each downstream firm might hope to be (very profitably) the only one outside the cartel. Inadvertently, Blonder-Tongue ensures that if such an outsider successfully challenges a weak patent for its own use, it also disrupts the cartel.

[24] Since $G_{r\theta}(r, \theta) = -\pi_2(v, r) < 0$, the optimal running royalty $r(\theta)$ is weakly decreasing in $\theta$.

[25] There exists $\theta^* \geq \theta_V$ such that negative fixed fees are optimal for all $\theta \leq \theta^*$, as shown in Figure 3.

**Theorem 6: For weak patents ($\theta \leq \theta_V$) licensed to downstream rivals using unrestricted two-part tariffs, $r(\theta) = v$ and $F(\theta) = -(1-\theta)[\pi(0,0) - \pi(v,v)] < 0$.**

Theorem 6 holds because high per-unit royalty rates maximize joint profits and this dominates royalty setting for weak patents. As a result, consumers gain nothing from the new technology, since each downstream firm's private marginal cost is the same as under the backstop technology. Ideal PTO review unambiguously benefits consumers and efficiency *ex post*. If the patent is ruled invalid, royalties drop to zero. If it is upheld, royalties become $r(1) \leq v$ rather than $r(\theta) = v$, so welfare will at worst be unchanged. The Appendix shows that if downstream competition involves strategic substitutes (as in Cournot competition), then $r(1) < v$ and welfare strictly improves even if the patent is upheld. Figure 3 displays $r(\theta)$ in the case where $r(1) < v$.

Many policies that benefit consumers and raise total welfare *ex post* also reduce patentees' payoffs, thereby worsening *ex ante* incentives. Here, however, the prospect of ideal PTO review, while reducing patentees' profits, strictly improves *ex ante* incentives in several respects. The Appendix proves

**Theorem 7: For weak patents ($\theta \leq \theta_V$) licensed to downstream rivals using unrestricted two-part tariffs, $r(\theta) = v$, $B(\theta) > 0$, and $P(\theta) > K(\theta)$. If also $\theta < \dfrac{w(0) - w(v)}{w(0) - w(v) + vx(v,v)}$, $K(\theta) < 0$. In the range where $K(\theta) > 0$, $P(\theta)/K(\theta)$ strictly decreases with $\theta$. Ideal PTO review ensures that the patent holder's social contribution is positive and that the ratio of profits to social contribution does not vary with patent strength.**

By equating $P(\theta)/K(\theta)$ across patent strengths, ideal PTO review eliminates profitable opportunities to do harm $(K(\theta) < 0)$ and eliminates a bias toward seeking weak patents,.

For Cournot oligopoly with linear demand and constant marginal costs, one can directly calculate $r(\theta)$, $F(\theta)$, $P(\theta)$, $B(\theta)$ and $K(\theta)$ in terms of $N$ and $v/A$,[26] where $A$ is the difference between the demand intercept and the production cost using the patented technology.[27] With $N = 5$ and

---

[26] The Supplementary Materials associated with this paper fully work out all of these functions in this special case, which is often used in the oligopoly and licensing literature (e.g., Kamien (1992); Sen and Tauman (forthcoming)).

[27] Alternatively, $2(v/A)$ approximates the proportionate increase in first-best welfare from the innovation.

$v/A = 0.1$, $r = v$ for $\theta \leq 0.41$, $F(\theta) < 0$ for $\theta < 0.48$, $K(\theta) < 0$ for $\theta < 0.18$, and $P(\theta)/K(\theta) > 2$ for $\theta \leq 0.48$.

## 6. Negative Fixed Fees Not Feasible

Section 5's results involve negative fixed fees, but we do not know how often such fees are feasible or used in practice.[28] Large negative fixed fees may induce entry, and may carry antitrust risk.[29] If a patent for which $F(\theta) < 0$ when feasible is licensed when negative fixed fees are not feasible, it will be licensed with no fixed fee (since $G$ is single-peaked in $r$).

If all downstream firms pay a pure running royalty $s$, P's income per downstream firm is $R(s) \equiv sx(s,s) = T(s) - \pi(s,s)$. In the range $0 \leq s \leq v$, $T(s)$ increases with $s$, and $\pi(s,s)$ falls with $s$, so $R(s)$ increases with $s$ and, unless it prefers to litigate, P will license at the highest royalty that downstream firms will accept rather than litigate. That is, $s = s(\theta)$, defined by $\pi(s(\theta), s(\theta)) = \theta\pi(v,s) + (1-\theta)\pi(0,0)$.[30] The Appendix gives conditions under which P prefers such linear licensing to litigation in our model.[31]

How does $s(\theta)$ compare to our benchmarks? If D litigates and loses, it will be at a cost disadvantage $v - s$ relative to its licensed rivals, so its downside from litigating is proportional to $v - s$ and to $|\pi_1(s,s)|$. In contrast, if D litigates and *wins* it will not gain any competitive

---

[28] We are unaware of any systematic empirical evidence on how often licenses contain negative fixed fees. Bharat Anand and Tarun Khanna (2000) assemble a sizeable data base of licensing contracts but lack sufficient information on the use of running royalties vs. fixed fees to reach reliable conclusions. The Federal Trade Commission (2002, 2005) reports on the use of negative fixed fees in certain pharmaceutical patent agreements.

[29] The Federal Trade Commission has brought several antitrust cases challenging negative fixed fees, known as "reverse payments" in antitrust circles, in agreements between vertically integrated patent holders (branded pharmaceutical suppliers) and would-be generic competitors. See Jeremy Bulow (2004), Herbert Hovenkamp, Mark Janis, and Lemley (2003), Shapiro (2003), and Robert Willig and John Bigelow (2004). The patent holder may be able to disguise negative fixed fees (for example, it might transfer know-how to the licensee, or agree to a side deal).

[30] D's payoff if it litigates and loses is $\pi(v,s)$ if P would then hold D to its backstop payoff. P will indeed do so if it can charge a positive fixed fee or, if licenses are constrained to be linear, if it would optimally charge a running royalty of $v$. The supplementary materials show that this is optimal in the linear Cournot example for small values of $v$, and we assume it below.

[31] In the linear Cournot case with $N = 5$, $v/A = 0.1$, and $\theta = 0.2$ the patent holder's expected payoff from licensing is about twice as large as from litigating. With $N = 10$, the ratio is about three to one.

advantage over its rivals. Rather, it will have lowered industry-wide costs from $s$ to zero, raising its profits from $\pi(s,s)$ to $\pi(0,0)$. Its upside is thus proportional to $s$ and to $\left|\pi_1(s,s)+\pi_2(s,s)\right|$.

We thus define the oligopoly's *relativity* coefficient $\rho \equiv \dfrac{\left|\pi_1(0,0)\right|}{\left|\pi_1(0,0)+\pi_2(0,0)\right|}$ as the relative

importance to a firm of small changes in its own costs versus small changes in industry-wide costs (evaluated at $s=0$). The extent to which $\rho>1$ measures the strength of downstream competition. For example, if downstream firms are symmetric Cournot oligopolists with constant marginal costs, the Appendix shows that $\rho \geq N$ for linear or constant-elasticity demand (and gives a more general expression for $\rho$). If the downstream industry is a Bertrand duopoly with differentiated products, $\rho$ is higher, the closer substitutes are the two downstream products. Using $\pi(s,s)=\theta\pi(v,s)+(1-\theta)\pi(0,0)$, the Appendix proves:

**Theorem 8: For small $v$, $s(\theta) \approx \theta v \dfrac{\rho}{1+(\rho-1)\theta}$.**

If $\rho>1$, Theorem 8's approximation for $s(\theta)$ exceeds $\theta v$ for $\theta \in (0,1)$ and is in turn approximately $\rho\theta v$ for small $\theta$. As $\rho\to 1$, $s(\theta)\to\theta v$, confirming that it is now relativity that enables the running royalty to exceed the *ex ante* benchmark level $\theta v$ (in Theorem 3). In contrast, in Theorem 6, the mechanism was joint profit maximization (tempered for stronger patents by rent shifting). Figure 3 displays $s(\theta)$ and $r(\theta)$ in the case where $r(1)<v$.

If a patent is linearly licensed as in Theorem 8, Theorem 1 (recalling $r(1)\leq v$) implies

$B(\theta) \geq \theta v \left[ \dfrac{\rho}{1+(\rho-1)\theta} - \mu \right] \lambda_{\min}$. For weak patents, the expression in brackets is robustly

positive; for instance, with Cournot oligopoly, linear demand, and constant costs, $\rho = N$ and

$\mu \leq \dfrac{x(0,0)}{x(v,v)} \left[ 1 + \dfrac{v}{p-c-v} \right]$. For $N=5$ and $v/A=0.1$, linear licenses are used for $\theta<0.48$. In

this range, $s(\theta)/\theta v$ declines from nearly 4 near $\theta=0$ to 1.7. While $K(\theta)>0$ for all $\theta>0$, $P(\theta)/K(\theta)$ declines with $\theta$ from over 7 near $\theta=0$ to 2.0 at $\theta=0.48$; for an ironclad patent, $P(1)/K(1)=1.13$. The ratio $B(\theta)/[B(\theta)+K(\theta)]$ declines with $\theta$ but for very weak patents is

near 0.5, meaning that ideal PTO review contributes roughly as much to society as (in expectation) did the patent holder without such further PTO review.

## 7. Variations and Extensions

### A. *Vertically Integrated Patent Holder*

Our analysis extends easily to the case in which the patent holder is vertically integrated, competing downstream with $N$ other downstream firms. Define $\pi^I(a,b)$ as the profits of a downstream firm with cost $a$, given that the other downstream firms have cost $b$ and the patent holder competes using the patented technology, and write $\phi(r)$ for P's product market profits if all rivals pay royalty $r$. We assume that $\pi^I(a,b)$ satisfies the three conditions assumed above for $\pi(a,b)$, and that $\phi'(r) \geq 0$. The Appendix shows how our analysis and main results carry over. With unrestricted two-part tariffs, for weak patents $r(\theta) = v$ and $K(\theta) < 0$, and for all patents $B(\theta) > 0$. If the patent holder only faces one downstream rival, then $r(\theta) = v$ for all $\theta$ and $F(\theta) < 0$ for all $\theta < 1$. If no downstream firm using the backstop technology could profitably compete against the patent holder, then again $r(\theta) = v$ for all $\theta$ and $F(\theta) < 0$ for all $\theta < 1$: the patent holder pays each downstream firm to agree not to infringe or challenge the patent, which is tantamount to exit.[32] If negative fixed fees are not feasible, then as in Theorem 8, $s(\theta) \approx \rho^I \theta v$

where now $\rho^I = \dfrac{\left|\pi^I_{\phantom{I}1}(0,0)\right|}{\left|\pi^I_{\phantom{I}1}(0,0) + \pi^I_{\phantom{I}2}(0,0)\right|}$. In the linear Cournot example, $\rho^I = (N+1)/2$, so again

$B(\theta) > 0$ for all $\theta$.

### B. *Short-Term vs. Long-Term Licenses*

We assumed above that a downstream firm's license remains in force if the patent is upheld after P litigates it with another downstream firm. Outside our model, licensees often make specific investments to use the patented technology, which provides an efficiency reason to design licenses that way. The Appendix shows that our results grow stronger if the patent holder can

---

[32] If P's downstream division were less efficient than downstream firms, these outcomes would not maximize joint profits. If feasible, P would prefer to commit to shut down its downstream division. Technically, $T$ would not be increasing in $r$, as we are assuming.

offer "short-term" licenses that do not survive a finding of validity. With unrestricted two-part tariffs, $r(\theta) = v$ for all $\theta$, so the *ex post* welfare analysis is the same as it was above if $\theta_v = 1$. *Ex ante,* for all $\theta < 1$, $P(\theta)$ is higher than we derived above, and $K(\theta)$ is unchanged or lower, so the bias resulting from $P(\theta)/K(\theta) > P(1)/K(1)$ is stronger than above.

## C. *Linear Licenses*

As Kamien (1992) noted, running royalties appear to be common. It is our impression that this is true (contrary to Theorem 5) even if licensees do not compete; the reasons are presumably outside our model, such as risk aversion, asymmetric information, and moral hazard.[33]

When licensees do not compete and pass-through is linear, the Appendix shows that consumers' risk preferences between a certain royalty $s(\theta)$ and the uncertain result of PTO review are reflected in each downstream firm's similar preferences. Since a downstream firm will infringe if $s(\theta)$ is too high, replicating through litigation the validity gamble of PTO review, consumers are protected by downstream firms as their agents against royalties that hurt them relative to first determining patent validity. Since consumers are risk-loving in price, this implies that $s(\theta) < \theta s(1) = \theta v$; however, to first order, $s(\theta) = \theta v$.

When licensees compete, the equilibrium running royalty is the much higher $s(\theta)$ calculated in Theorem 8 for all $\theta$, not just for $\theta \leq \theta_V$. Since $s(\theta) > \theta v$, $s(\theta)x(s(\theta),s(\theta)) > \theta vx(v,v)$ so the patent holder prefers licensing to litigation. The Appendix shows that

$$\frac{B(\theta)}{B(\theta)+K(\theta)} \geq \left[\frac{s(\theta)}{\theta v} - \mu\right]p'(s).$$ For small $\theta$, $\frac{s(\theta)}{\theta v} \approx \rho$ so $\frac{B(\theta)}{B(\theta)+K(\theta)} \geq [\rho - \mu]p'(s)$. This

can easily exceed unity, in which case $K(\theta) < 0$.

## D. *Enhanced Review of Patents*

Short of "ideal" review, more realistic "enhanced" patent review (by the PTO or in litigation that settles before final judgment) uncovers some additional information about patent validity, inducing a mean-preserving spread on patent strength. Enhanced review thus increases

---

[33] For example, Sugato Bhattacharyya and Francine Lafontaine (1995) construct a model in which linear sharing rules are optimal due to two-sided moral hazard.

(decreases) the expected value of any concave (convex) function of $\theta$. In parallel with Theorem 1, *ex post* enhanced review is valuable if $r(\theta)$ is sufficiently (relative to $\mu$) concave in $\theta$. Theorem 2 becomes: consumers benefit from enhanced review if $r$ is concave in $\theta$. Theorem 5 directly applies to any "additional" review, not only ideal. If $\theta_V = 1$, Theorem 6 shows that $r$ is concave (not only weakly, because $r(0) = 0$), so enhanced review is *ex post* beneficial if and only if it has positive probability of actually invalidating the patent.[34] If $\theta_V < 1$, $r$ is globally concave if and only if it is concave on $[\theta_V, 1]$. Theorem 8 implies (with due attention to the approximation) that $s(\theta)$ is concave in $\theta$, as in the generalized forms of Theorems 1 and 2.

## E. Patent Validity and Patent Scope

We have cast our analysis so far in terms of uncertainty about patent validity. There is often also (or instead) uncertainty about whether a downstream firm's product actually infringes the patent.[35] The two kinds of uncertainty are equivalent if there is just one licensee. Our analysis extends to cases where patent scope or infringement rather than patent validity is the key issue, if a finding of (non-)infringement against one downstream firm implies that other downstream firms also are (not) infringing and that these firms can stop paying running royalties.

## F. Litigation Costs and Bargaining

Litigation costs make licensing even more attractive relative to litigation than our model suggests. How do they affect the terms on which a probabilistic patent is licensed in the shadow of litigation? If, as above, P makes take-it-or-leave-it offers, of course, it can demand more; if downstream firms had commitment power, they could offer less. Extending the model to include litigation costs would thus seem most natural if we also extended it to more general bargaining, which becomes complex when competing downstream firms bargain with P. If litigation costs are unrelated to $(\theta, v)$ they may dominate the bargaining for small, weak patents, but if

---

[34] The litigation process induces a series of mean-preserving spreads on patent strength, with $\theta = 0$ never arising until final judgment. For patents licensed to downstream rivals, private parties may not pursue litigation to final judgment. After a verbal ruling dismissing Rambus's patent infringement case against Infineon, but before a written opinion that could have set a precedent for other infringement cases brought by Rambus, Rambus and Infineon settled. See Don Clark, "Rambus, Infineon Reach Settlement," *Wall Street Journal,* March 22, 2005.

[35] Michael Waterson (1990) studies how uncertainty about infringement (patent scope) affects rivals' design decisions.

bargaining skill and litigation costs are symmetric, their effect will tend to be neutral, restoring our results.[36] Steven Meurer (1989) considers signaling issues when a vertically integrated holder of a probabilistic patent can litigate with a single downstream rival.

## G. Litigation Credibility

As we noted above, if $N-1$ firms sign lucrative licenses and one infringes, the patent holder might be reluctant to litigate and put its licensing revenues at risk. Jay Pil Choi (1998) models this question of litigation credibility, but does not analyze licensing terms. In his model, the vertically integrated patent holder either excludes rivals completely or they enter with no royalties. Yet clearly P has a strong incentive to ensure that it has a credible threat to sue an infringer. Several mechanisms may help it. First, if infringers divert substantial sales from licensees, as they might (especially if the downstream industry is highly competitive) due to their cost advantage from not paying royalties, litigation may well be credible. Second, reputation effects can make litigation credible. Third, licenses may contractually commit P to sue: for instance, by allowing licensees to stop paying royalties if P fails to challenge an infringer.

If none of these (or other) mechanisms establishes litigation credibility, P may have to adjust its licensing terms to do so. Generalizing our model in this direction exposes the inherent relationship between litigation credibility and relativity. Litigation becomes more credible if licensees must continue paying royalties even if the patent is overturned. *Blonder-Tongue* limits what a license can do in this respect, but licenses may be able to bundle trade secrets (or other patents) with a weak patent. However, the more effectively the license ensures that running royalties continue even if another downstream firm successfully challenges, the greater is the upside to a challenge, and the less the patent holder can exploit relativity. These issues will be a fertile area for future work.

## 8. Conclusion

In fiscal 2006, the U.S. Patent and Trademark Office received 444,000 patent applications and issued 183,000 patents; in the past ten years, it has issued 1.7 million patents.[37] Evidence has

---

[36] Farrell and Robert Merges (2004) explore the role of relativity in determining parties' effort (expenditure) in patent litigation, noting that this makes both litigation costs and the resulting probability $\theta$ endogenous.

mounted that many patents would likely not hold up if tested in court, not surprisingly since on average a patent examiner reportedly spends only about 15-20 hours on a patent application.[38] The Supreme Court recently expressed deep concern that many patents have been improperly issued covering obvious technologies.[39]  Efforts are underway in Congress to reform the patent system, with the information technology sector in particular deeply concerned about the issuance of large numbers of questionable patents.

Since far more patents are licensed than litigated, the economic impact of questionable patents depends largely on how they are licensed.  We modeled how licensing terms vary with patent strength, and found that weak patents licensed to downstream firms that are not rivals (to each other or to the patent holder) command correspondingly low royalties.  In our model, there are no social benefits of examining these patents more closely.  In sharp contrast, weak patents on technology used by downstream firms that are rivals (to each other or to the patent holder) command surprisingly large running royalties, especially if licenses can use unrestricted two-part tariffs.  There are large social benefits, *ex post* and, perhaps more importantly, *ex ante*, of better examining commercially significant patents that will be licensed to downstream rivals.

Closely scrutinizing the hundreds of thousands of patent applications filed each year, many of which end up having no commercial significance, would be very costly.  Our analysis suggests a more targeted approach: re-examination of issued patents covering valuable technology that is useful to multiple downstream firms that compete against each other or against the patent holder. Our analysis thus supports current proposals to expand post-grant review of commercially significant patents, but also identifies downstream competitive conditions as a key indicator of the value of further review.

---

[37] U.S. Patent and Trademark Office (2006), Tables 2 and 6.

[38] Federal Trade Commission (2003),  Chapter 5, p. 5.

[39] Oral Argument in *KSR International, Inc. vs. Teleflex, et. al*, No. 04-1350, November 28, 2006, available at http://www.supremecourtus.gov/oral_arguments/argument_transcripts/04-1350.pdf.

# References

Anand, Bharat and Tarun Khanna, (2000), "The Structure of Licensing Contracts," *Journal of Industrial Economics*, 48:103-135.

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

Aoki, Reiko, and Jin-Li Hu (1999) "Licensing vs. Litigation: The Effect of the Legal System on Incentives to Innovate," *Journal of Economics & Management Strategy,* 8:133-160.

Bhattacharyya, Sugato, and Francine Lafontaine (1995), "Double-Sided Moral Hazard and the Nature of Share Contracts," *Rand Journal of Economics,* 26:761-781.

Bulow, Jeremy (2004), "The Gaming of Pharmaceutical Patents," *Innovation Policy and the Economy*, Adam Jaffe, Scott Stern, and Josh Lerner, eds., 4:145-187 MIT Press.

Chiao, Benjamin, Josh Lerner, and Jean Tirole (2006), "The Rules of Standard Setting Organizations: An Empirical Analysis," *Rand Journal of Economics*, forthcoming.

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

Choi, Jay Pil (2002, revised 2006), "Patent Pools and Cross Licensing in the Shadow of Patent Litigation," available at http://www.msu.edu/~choijay.

Choi, Jay Pil (2005), "Live and Let Live: A Tale of Weak Patents," *Journal of the European Economic Association*, 724-733.

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,* available at http://www.ftc.gov/os/2002/07/genericdrugstudy.pdf.

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

Federal Trade Commission (2005), "Agreements Filed with the Federal Trade Commission under the Medicare Prescription Drug, Improvement, and Modernization Act of 2003," available at http://www.ftc.gov/os/2006/04/fy2005drugsettlementsrpt.pdf.

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

Jaffe, Adam and Josh Lerner (2004), *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It*, Princeton, N.J.: Princeton University Press.

Kamien, Morton (1992), "Patent Licensing," in *Handbook of Game Theory and Economic Applications*, North-Holland, Chapter 12, 331-354.

Katz, Michael and Carl Shapiro (1986), "How to License Intangible Property," *Quarterly Journal of Economics*, 101: 567-590.

Kesan, Jay and Gwendolyn Ball (2006), "How Are Patent Cases Resolved?  An Empirical Analysis of the Adjudication and Settlement of Patent Disputes," University of Illinois.

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.

Liao, Chun-Hsiung, and Debapriya Sen (2005), "Subsidy in Licensing: Optimality and Welfare Implications," *The Manchester School* 73:3 (June) 281-299.

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

Marco, Alan (2006), "Learning by Suing: Structural Estimates of Court Errors in Patent Litigation," working paper, Vassar College, October.

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

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

Moore, Kimberly, (2000), "Judges, Juries, and Patent Cases: An Empirical Peek inside the Black Box," *Michigan Law Review,* 99:365-409.

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

Rey, Patrick and Thibaud Vergé (2004), "Bilateral Control with Vertical Contracts," *Rand Journal of Economics,* 35:728-746.

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

Sen, Dibapriya and Yair Tauman (forthcoming), "General Licensing Schemes for a Cost-Reducing Innovation," *Games and Economic Behavior*.

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

Shapiro, Carl (2006), "Injunctions, Hold-Up, and Patent Royalties," available at http://faculty.haas.berkeley.edu/shapiro/royalties.pdf.

U.S. Patent and Trademark Office, (2006), *Performance and Accountability Report, Fiscal Year 2006,* at http://www.uspto.gov/web/offices/com/annual/2006/2006annualreport.pdf.

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.
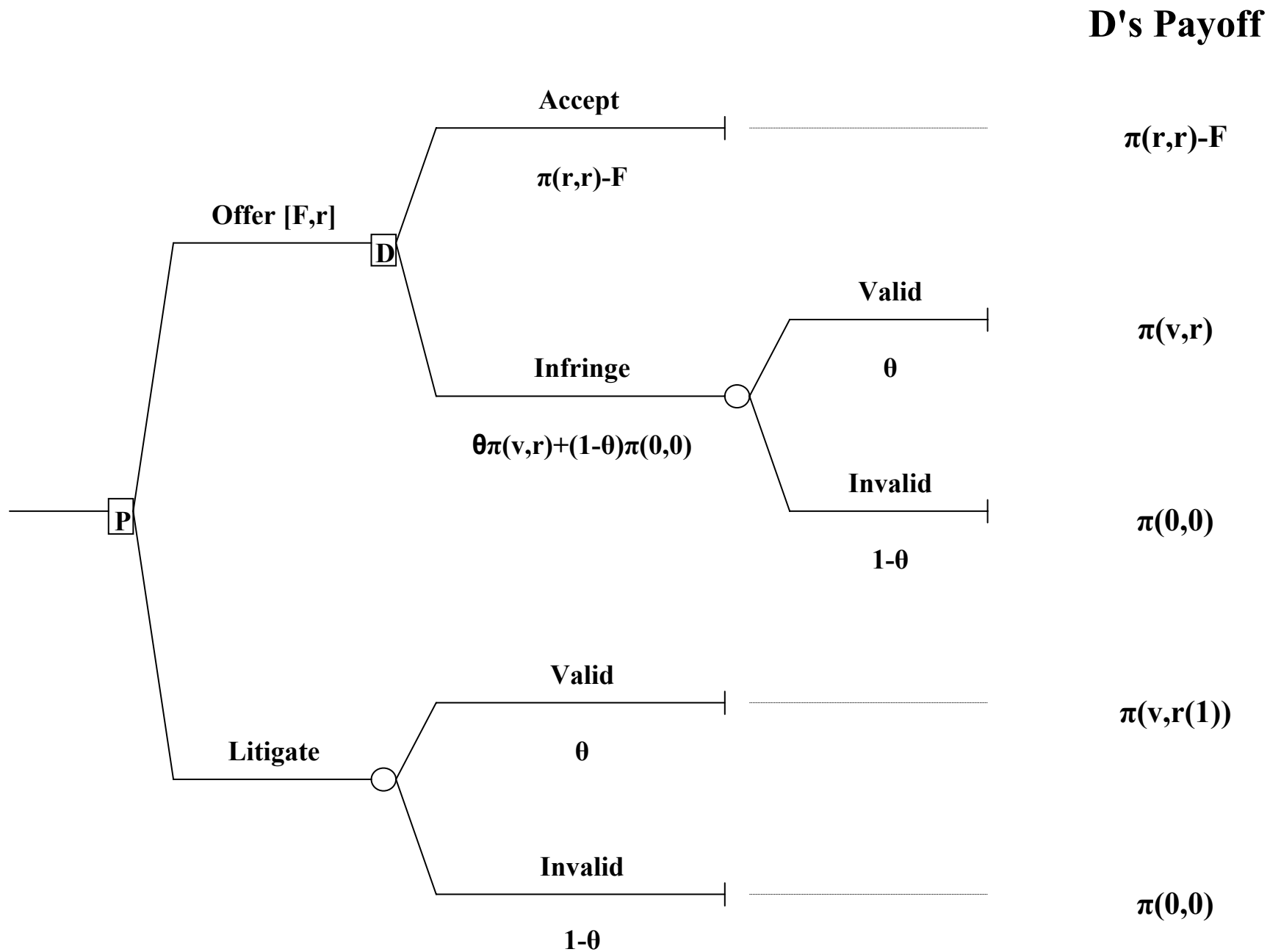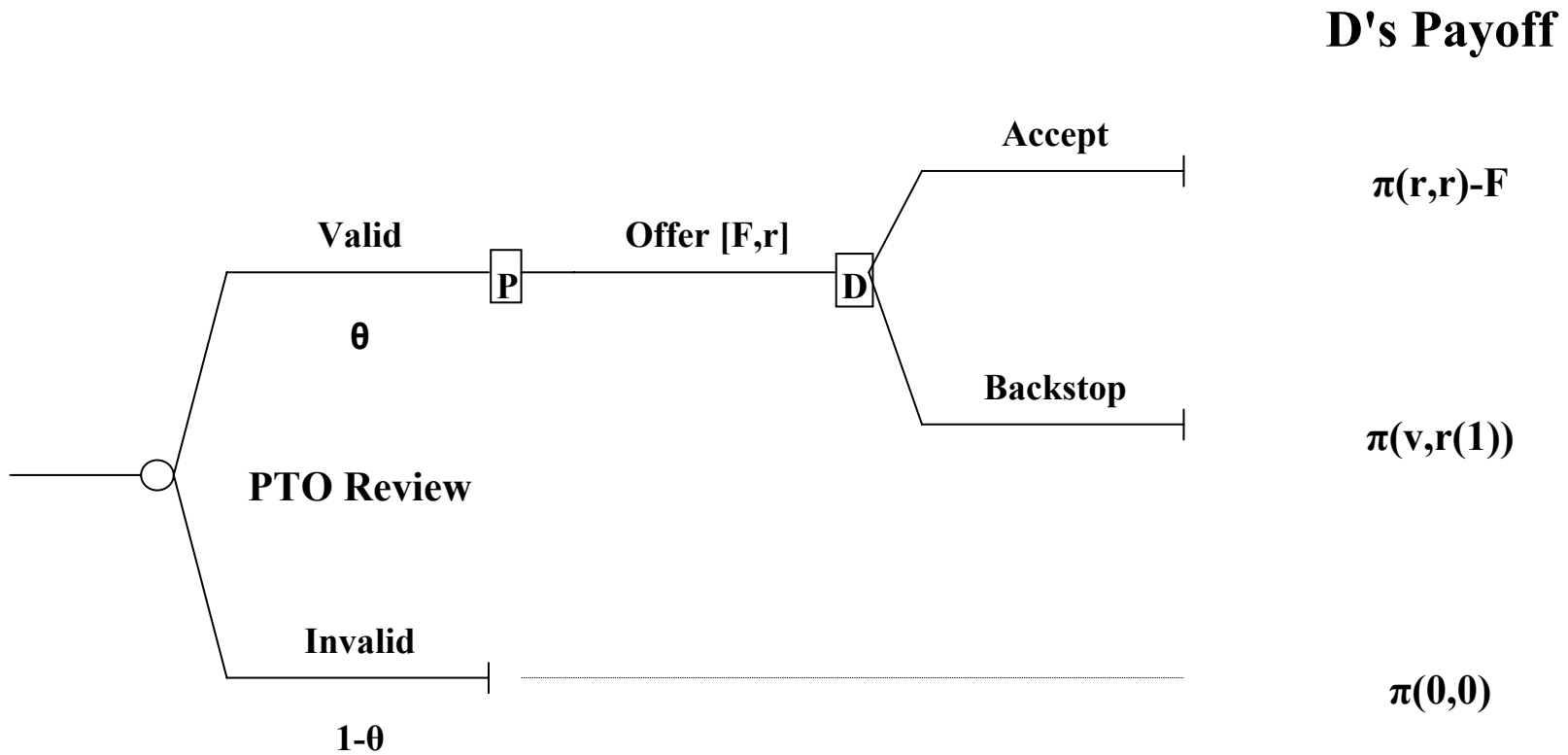
# Figure 1: Licensing Game



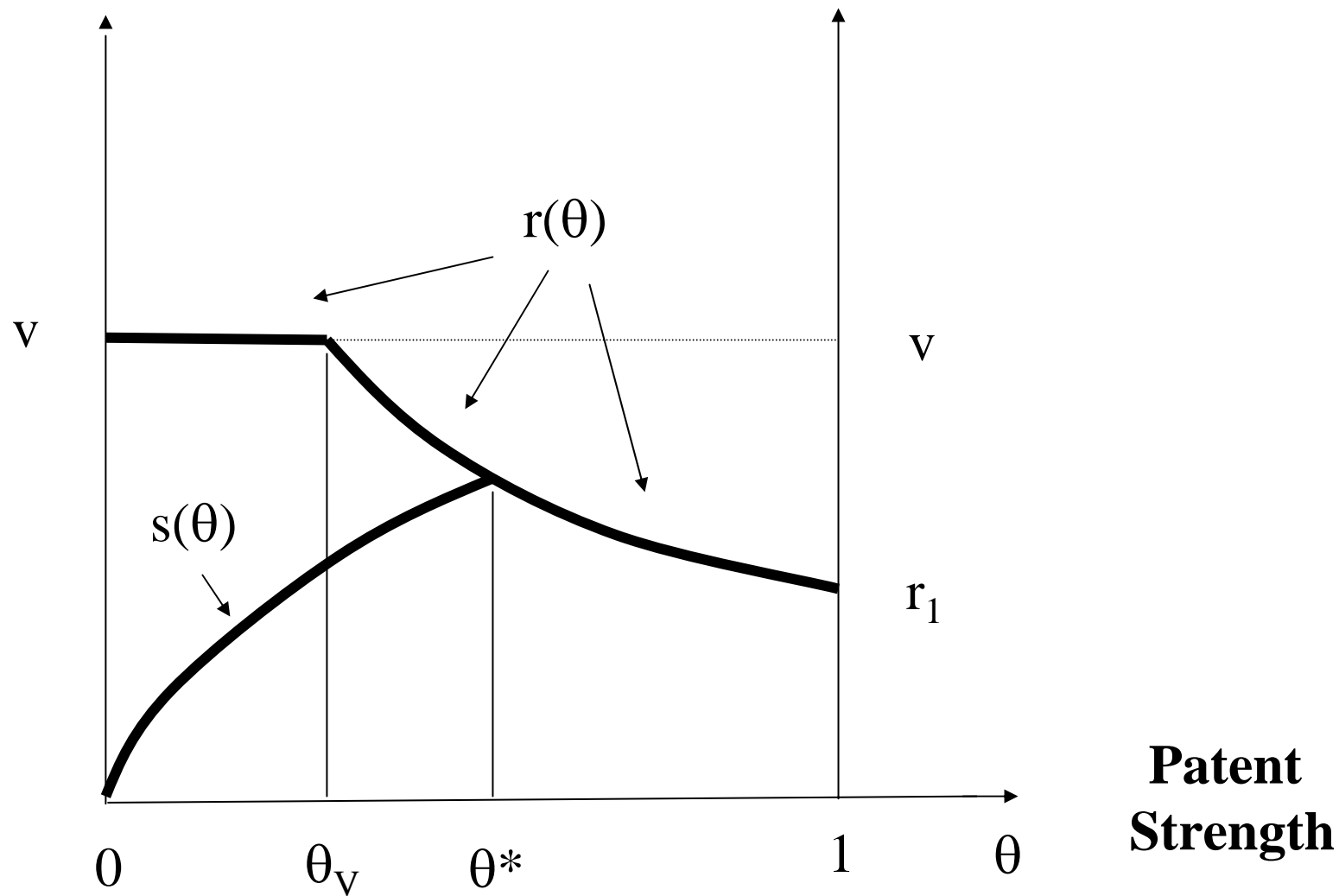**D's Payoff**

Accept

$\pi(r,r)-F$

$\pi(r,r)-F$

Offer [F,r]

D

Valid

$\pi(v,r)$

$\theta$

Infringe

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

Invalid

$\pi(0,0)$

$1-\theta$

P

Valid

$\pi(v,r(1))$

$\theta$

Litigate

Invalid

$\pi(0,0)$

$1-\theta$

# Figure 2: Licensing Game with Ideal PTO Review

**D's Payoff**

**Accept**

$\pi(r,r)-F$

**Valid**          **Offer [F,r]**

P          D

$\theta$

**Backstop**

$\pi(v,r(1))$

**PTO Review**

**Invalid**

$\pi(0,0)$

$1-\theta$

# Figure 3: Equilibrium Royalty Rates

**Royalty Rate**

$r(\theta)$

v                                v

$s(\theta)$

$r_1$

0         $\theta_V$       $\theta^*$              1     $\theta$

**Patent Strength**

# Appendix

**Proof of Theorem 1:** $B(\theta) \geq [r(\theta) - \mu\theta r(1)]\lambda_{\min}$ .

Applying the intermediate value theorem to $w(r)$, and using $0 \leq r(\theta), r(1) \leq v$, there exist $t_1 \in [0, v]$ and $t_2 \in [0, v]$ such that $w(r(1)) = w(0) + r(1)w'(t_1)$ and $w(r(\theta)) = w(0) + r(\theta)w'(t_2)$. Substituting into $B(\theta) = [\theta w(r(1)) + (1-\theta)w(0)] - w(r(\theta))$ and simplifying gives $B(\theta) = \theta r(1)w'(t_1) - r(\theta)w'(t_2)$. Since $w'(t_1) < 0$ and $w'(t_2) < 0$, we have $B(\theta) = r(\theta)|w'(t_2)| - \theta r(1)|w'(t_1)| \geq r(\theta)\lambda_{\min} - \theta r(1)\lambda_{\max} = [r(\theta) - \mu\theta r(1)]\lambda_{\min}$ .

We next discuss upper bounds on $\mu$, precisely in the Cournot case and heuristically more generally. We have $w'(r) = [p(r) - c]\dfrac{d}{dr} Nx(r, r)$, where $c$ is the marginal social cost of production using the patented technology. Differentiating again, and using $Nx(r, r) \equiv X(p(r))$, yields $w''(r) = [p'(r)]^2 [X'(p) + (p-c)X''(p)] + p''(r)(p-c)X'(p)$. Assuming $p''(r) = 0$, this implies that $w''(r)$ has the sign of $X'(p) + (p-c)X''(p)$, which is negative if demand is linear (or concave) in the range $p(0) \leq p \leq p(v)$, whatever the oligopoly behavior. That implies that $\lambda_{\min}$ occurs at $r = 0$, $\lambda_{\max}$ occurs at $r = v$, $\mu = \dfrac{|w'(v)|}{|w'(0)|}$, and $B(\theta) \geq r(\theta)|w'(0)| - \theta r(1)|w'(v)|$.

For Cournot oligopoly, our formula for $w'(r)$ yields $|w'(r)| = [p'(r)]\left[\varepsilon N \dfrac{p(r) - c}{p(r)}\right][x(r, r)]$ where $\varepsilon$ is the absolute value of the elasticity of demand. Since $\mu$ is a ratio of $|w'(t)|$'s, it is the product of the ratios of the three factors in brackets making up $|w'(r)|$. If $p'(r)$ is a constant, the first ratio will be unity. In Cournot oligopoly, $\dfrac{p(r) - r - c}{p(r)} = \dfrac{1}{\varepsilon N}$, so $N\varepsilon \dfrac{p(r) - c}{p(r)} = \dfrac{p(r) - c}{p(r) - r - c}$, which equals unity at $r = 0$ and is increasing in $r$. Therefore, the second ratio is bounded above by $\dfrac{p(v) - c}{p(v) - v - c} = 1 + \dfrac{v}{p(v) - v - c}$. The third ratio is bounded above by $\dfrac{x(0, 0)}{x(v, v)}$, which reflects the proportionate increase in output resulting from the innovation, if it is available royalty-free. Therefore, in Cournot oligopoly $\mu \leq \dfrac{x(0, 0)}{x(v, v)}\left[1 + \dfrac{v}{p(v) - c - v}\right]$. Below, we will be comparing $\mu$ to numbers typically above two.

**Proof of Theorem 2:** If $r(\theta) \geq \theta r(1)$ **then consumers benefit from ideal PTO review.**

We prove this for weakly concave $p$, $p''(r) \leq 0$, not just for linear $p(r)$. With $p''(r) \leq 0$, for any $t_1$ and $t_2$ and $\lambda \in [0,1]$, $p(\lambda t_1 + (1-\lambda)t_2) \geq \lambda p(t_1) + (1-\lambda)p(t_2)$. Write $CS(p)$ for consumer surplus as a function of the representative downstream price. Since $CS(p)$ is decreasing in price, this in turn implies that $CS(p(\lambda t_1 + (1-\lambda)t_2)) \leq CS(\lambda p(t_1) + (1-\lambda)p(t_2))$. Since $CS(p)$ is convex in price, $CS(\lambda p(t_1) + (1-\lambda)p(t_2)) \leq \lambda CS(p(t_1)) + (1-\lambda)CS(p(t_2))$. Combining these two inequalities gives $CS(p(\lambda t_1 + (1-\lambda)t_2)) \leq \lambda CS(p(t_1)) + (1-\lambda)CS(p(t_2))$, so the composite function $S(t) \equiv CS(p(t))$ is convex in $t$. This proves that consumers are risk loving in the royalty rate (a fact useful beyond this Theorem), so they benefit from mean-preserving spreads in $r$. Since they also prefer lower royalty rates, they welcome ideal PTO review, which is a combination of a mean preserving spread and a possible reduction in the expected royalty rate when $r(\theta) \geq \theta r(1)$.

**Proof of Theorem 3:** If $r(\theta) > \theta v$ **then** $P(\theta) > K(\theta)$**.**

Each downstream firm's reservation (litigation) payoff is $\theta \pi(v,r) + (1-\theta)\pi(0,0)$. Without the patent holder's activities, there is a probability $\theta$ that the patented technology would be unavailable, resulting in per-firm profits of $\pi(v,v)$; with probability $1-\theta$, the technology would be available without royalties, resulting in per-firm profits of $\pi(0,0)$. The difference between equilibrium payoff and expected but-for payoff is thus $\theta[\pi(v,r) - \pi(v,v)]$, which is (weakly) negative when $r \leq v$ since profits are increasing in rivals' cost level. When downstream firms actually compete ($\pi_2 > 0$) and $r < v$, downstream firms are strictly hurt.

Without the patent holder's activities, with probability $\theta$ the patented technology would be unavailable, which for consumers is the same as there being a royalty rate of $v$. With probability $1-\theta$ the patented technology would be available without royalties. The proof of Theorem 2 showed that consumers are weakly risk-loving in the royalty rate. Since they also prefer lower royalties, if $r(\theta) \geq \theta v$, they are harmed by the patent holder's activities, and strictly so if $r > \theta v$.

Since the patent holder's activities harm both downstream firms and consumers, the patent holder's profits must exceed its social contribution; the proof shows that if $\theta < 1$ and $\pi_2 > 0$, at least one group is strictly harmed, so the comparison is strict. Like Theorem 2, Theorem 3 holds for $p''(r) \leq 0$, not just for linear $p(r)$.

**Proof of Theorem 4:** If $B(\theta) > 0$**, then** $P(\theta)/P(1) > K(\theta)/K(1)$**.**

By definition, $B(\theta) \equiv [\theta W(1) + (1-\theta)W(0)] - W(\theta)$ and $K(\theta) \equiv W(\theta) - [(1-\theta)W(0) + \theta \overline{W}]$.

Adding these together gives $B(\theta) + K(\theta) = \theta[W(1) - \overline{W}]$. Since $K(\theta)$ is the patent holder's

contribution if the patent is not examined more carefully, and $B(\theta)$ is the additional benefit arising from ideal PTO review, their sum is the patent holder's expected contribution under ideal PTO review, which is precisely the contribution from an ironclad patent times the probability that the patent will indeed be found valid under ideal PTO review.

Since $K(1) = W(1) - \overline{W}$, this implies that $B(\theta) + K(\theta) = \theta K(1)$. Writing this as $K(\theta) = \theta K(1) - B(\theta)$, when $B(\theta) > 0$ we must have $K(\theta) < \theta K(1)$. Since $P(\theta) \geq \theta P(1)$,

$\dfrac{P(\theta)}{K(\theta)} \geq \dfrac{\theta P(1)}{K(\theta)}$. Using $K(\theta) < \theta K(1)$, we get $\dfrac{P(\theta)}{K(\theta)} \geq \dfrac{\theta P(1)}{K(\theta)} > \dfrac{\theta P(1)}{\theta K(1)} = \dfrac{P(1)}{K(1)}$.

## Proof of Theorem 6

**The patent holder strictly prefers licensing to litigation.**

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

For $\theta \leq \theta_V$, $r(\theta) = v$ and $H(\theta) = T(v) - [\theta\pi(v,v) + (1-\theta)\pi(0,0)]$. Since $\lim_{\theta \to 0} H(\theta) = T(v) - \pi(0,0) = T(v) - T(0) > 0$ (recall that $T(r)$ increases with $r$ for $0 \leq r \leq v$), for sufficiently weak patents we must have $H(\theta) > \theta H(1)$.

If $r(1) = v$ then $r(\theta) = v$ for all $\theta > 0$, making $H$ linear in $\theta$. 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 strictly greater than the payoff from litigation for all $\theta < 1$.

Alternatively, if $r(1) < v$ then in the range $0 < \theta \leq \theta_V$, $H(\theta)$ is as just discussed. For $\theta \geq \theta_V$, $r$ varies, so $H(\theta)$ is a convex function of $\theta$ 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 $\theta$. 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 $H'(1) \leq H(1)$. Now $H'(1) = \pi(0,0) - \pi(v, r(1))$ and $H(1) = T(r(1)) - \pi(v, r(1))$. So $H'(1) \leq H(1)$ if and only if $\pi(0,0) = T(0) \leq T(r(1))$, a condition that must be satisfied since $T(r)$ is increasing in $r$ for $r \leq v$.

**If the downstream oligopoly game involves strategic substitutes, then $r(1) < v$.**

For $\theta = 1$, the patent holder maximizes $G(r;1) = rx(r,r) + \pi(r,r) - \pi(v,r)$. Since $G(r;1)$ is single-peaked, $r(1) < v$ 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, $\pi_1$. We can decompose $\pi_1$ 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. With strategic substitutes, including Cournot oligopoly, the indirect effect is negative: when rivals learn a firm 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 $r(1) < v$.

## Proof of Theorem 7

**For $\theta \leq \theta_V$ and $\theta < \dfrac{w(0) - w(v)}{w(0) - w(v) + vx(v,v)}$, $K(\theta) < 0$. For $\theta \leq \theta_V$, if $K(\theta) > 0$, then $P(\theta)/K(\theta)$ strictly decreases with $\theta$.**

When $r = v$, $P(\theta) = vx(v,v) - (1-\theta)[\pi(0,0) - \pi(v,v)]$ and $K(\theta) = \theta vx(v,v) - (1-\theta)[w(0) - w(v)]$. Rearranging this last equation shows that $K(\theta) < 0$ for $\theta < \dfrac{w(0) - w(v)}{w(0) - w(v) + vx(v,v)}$.

Both $P(\theta)$ and $K(\theta)$ are increasing and linear in $\theta$. Their difference is $P(\theta) - K(\theta) = (1-\theta)[vx(v,v) + (w(0) - w(v)) - (\pi(0,0) - \pi(v,v))]$, which is zero at $\theta = 1$. (These linear functions do not apply for $\theta > \theta_V$; we are using this fact only to demonstrate the properties of $P(\theta)/K(\theta)$ in the range $\theta \leq \theta_V$.) Therefore, showing that $P(0) > K(0)$ is sufficient to conclude that $P(\theta)/K(\theta)$ strictly decreases with $\theta$. $P(0) > K(0)$ if $vx(v,v) + (w(0) - w(v)) - (\pi(0,0) - \pi(v,v)) > 0$. Writing $w(r) = T(r) + S(r)$, where $S(r)$ is the consumer surplus when royalties are $r$, this expression is equivalent to $vx(v,v) + T(0) + S(0) - T(v) - S(v) - \pi(0,0) + \pi(v,v) > 0$. Simplifying, this becomes $S(0) > S(v)$, which holds.

## Licensing vs. Litigation without Negative Fixed Fees

The patent holder strictly prefers licensing to litigation if $s(\theta)x(s(\theta),s(\theta)) > \theta P(1)$. If $s(\theta) = k\theta v$, this becomes $kvx(s,s) > rx(r,r) + [\pi(r,r) - \pi(v,r)]$, where $r = r(1)$. Since

$vx(s,s) \geq vx(v,v) \geq rx(r,r)$, this condition is satisfied if $(k-1)vx(s,s) > \pi(r,r) - \pi(v,r)$. If $r(1) = v$ then this becomes $(k-1)vx(s,s) > 0$, which is satisfied for all $k > 1$.

If $r(1) < v$ we can use the intermediate value theorem to write $\pi(r,r) - \pi(v,r) = (r-v)\pi_1(t,r) = (v-r)|\pi_1(t,r)|$ for some $t \in [r(1), v]$. Substituting, the sufficient condition becomes $(k-1)vx(s,s) > (v-r)|\pi_1(t,r)|$.

Now $\pi_1(t,r) = -x(t,r) + IE$, where *IE* is the indirect effect of the higher costs on the firm's profits that arises because the firm's rivals adjust their behavior. With strategic complements, including Bertrand oligopoly, the indirect effect is positive, so $|\pi_1(t,r)| < x(t,r)$. In this case, the sufficient condition for the patent holder to prefer licensing is satisfied if $(k-1)vx(s,s) > (v-r)x(t,r)$. If $s \leq r(1)$, then $x(s,s) > x(r,r) \geq x(t,r)$, so the patent holder prefers licensing to litigation for $k \geq 2$. This condition is sufficient but far from necessary.

With strategic substitutes, including Cournot oligopoly, the indirect effect is negative. With linear demand and constant marginal costs, $|\pi_1(t,r)| = x(t,r)\dfrac{2N}{N+1} < 2x(t,r)$, so we get the sufficient condition $(k-1)vx(s,s) > 2(v-r)x(t,r)$. If $s \leq r(1)$, this condition is satisfied for $k \geq 3$. Again, this condition is sufficient but far from necessary.

**Proof of Theorem 8:** **For small *v*,** $s(\theta) \approx \theta v \dfrac{\rho}{1 + (\rho-1)\theta}$.

Recall that $s(\theta)$ satisfies $\pi(s,s) = \theta\pi(v,s) + (1-\theta)\pi(0,0)$. By the intermediate value theorem, there exist $a, b \in (0,1)$ such that the left hand side is equal to $\pi(0,0) + s[\pi_1(as,as) + \pi_2(as,as)]$ and the right hand side is equal to $\pi(0,0) + v\pi_1(bv,bs) + s\pi_2(bv,bs)$, where the subscripts denote partial derivatives. Therefore, $s[\pi_1(as,as) + \pi_2(as,as) - \theta\pi_2(bv,bs)] = \theta v\pi_1(bv,bs)$, or

$$\frac{s}{\theta v} = \frac{\pi_1(bv,bs)}{\pi_1(as,as) + \pi_2(as,as) - \theta\pi_2(bv,bs)}.$$

Since $0 \leq s \leq v$, for small *v*, one can approximate *s* by $\dfrac{s}{\theta v} \approx \dfrac{\pi_1(0,0)}{\pi_1(0,0) + (1-\theta)\pi_2(0,0)}$. Using the

definition of $\rho$, this is equivalent to $s(\theta) \approx \theta v \dfrac{\rho}{1 + (\rho-1)\theta}$.

The result is approximate because we substituted (0,0) for the varying arguments in the partial derivatives of $\pi$. Because we are concerned with a ratio, we need to bound the *proportional* error introduced by that substitution. Technically this requires that $\pi_1(0,0)$ and $\pi_1(0,0) + \pi_2(0,0)$ are nonzero (otherwise, continuity would bound only the *absolute* error in

numerator or denominator, leaving open the possibility of large proportional errors). Since we have assumed that (see Section 2B), the Theorem indeed holds as a limiting statement for small enough $v$. But how is it likely to fare for moderate but not infinitesimal $v$? In the course of calculating $\rho$ in Cournot oligopoly next, we show that the partial derivatives of $\pi$ vary with output. At least in simple cases, this implies that the proportional error introduced by substituting for the varying arguments is bounded by the proportional difference in output as $(a,b)$ varies over $[0,v]\times[0,v]$. In those cases, and (we suggest) plausibly in general in moderately competitive markets with moderate $v$, that error factor will not be large compared to the ratio by which the approximation exceeds the benchmark $\theta v$.

**Relativity Ratio in Cournot Oligopoly: Comparison of $\rho$ and $N$**

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 $\lambda_i \equiv \dfrac{-p'(X)-x_i p''(X)}{-p'(X)}$, so with $dc_i = dr_i$ we have $dx_i = -\lambda_i dX + \dfrac{dr_i}{p'(X)}$.

Writing $\Lambda \equiv \sum_i \lambda_i$ and adding up across all firms gives $\dfrac{dX}{dr_1} = \dfrac{1}{[1+\Lambda]p'(X)}$. Substituting for $dX$ using this expression, we get $\dfrac{dx_1}{dr_1} = \dfrac{1+\Lambda-\lambda_1}{[1+\Lambda]p'(X)}$ and $\dfrac{dx_j}{dr_1} = \dfrac{-\lambda_j}{[1+\Lambda]p'(X)}, j \neq 1$.

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 firm $j$'s equilibrium output $x_j$ times the change in price resulting from the equilibrium change in output by all *other* firms, $dX - dx_j$. This price change is given by $p'(X)[dX - dx_j]$, which equals $\dfrac{1+\lambda_j}{1+\Lambda}dr_1$. Since this expression does not contain any parameters specific to firm 1, the effect on firm $j$'s profits of a small increase $dr$ in *all other firms'* running royalties is given by $(N-1)x_j \dfrac{1+\lambda_j}{1+\Lambda}dr$. Returning to our main notation, we therefore have $\pi_2 = (N-1)x_j \dfrac{1+\lambda_j}{1+\Lambda}$.

Similarly the effect on firm 1's profits of a small increase $dr_1$ in its own running royalty is equal to the direct cost effect, $-x_1 dr_1$, plus the effect of the price change caused by other firms' output changes, $x_1 p'(X)[dX - dx_1] = -\dfrac{\Lambda-\lambda_1}{1+\Lambda}dr_1$. Therefore $|\pi_1| = x_1 \dfrac{1+2\Lambda-\lambda_1}{1+\Lambda}$.

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 $\lambda_i \equiv \dfrac{-p'(X) - Xp''(X)/N}{-p'(X)} = 1 + Xp''(X)/Np'(X)$.

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 $E = N(1-\lambda)$. Hence, we obtain $\dfrac{|\pi_1|}{|\pi_1 + \pi_2|} = N\dfrac{2}{2-E}$, 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 $\dfrac{|\pi_1|}{|\pi_1 + \pi_2|} \geq N$. For linear

demand, $E = 0$, so $\dfrac{|\pi_1|}{|\pi_1 + \pi_2|} = N$. When demand has constant elasticity $\varepsilon > 1$ ($\varepsilon > 1$ is the

regularity condition for $\pi_1 + \pi_2 < 0$), we have $E = 1 + \dfrac{1}{\varepsilon}$, so $\dfrac{|\pi_1|}{|\pi_1 + \pi_2|} > N$.

## Vertically Integrated Patent Holder

Define $\pi^I(a,b)$ as the profits of a downstream firm with cost $a$, given that the other downstream firms have cost $b$ and the patent holder competes using the patented technology. This downstream firm's output is $x^I(a,b)$. We assume that $\pi^I(a,b)$ satisfies the three assumptions that Section 2B assumed for $\pi(a,b)$.

Write $\phi(r)$ for P's product market profits if the rivals all pay royalty $r$. We make the very mild assumption that $\phi'(r) \geq 0$; P earns no less profits from its downstream operations, the higher are the royalties paid by other downstream firms.

Define $T_I(r) = \phi(r) + Nrx^I(r,r) + N\pi^I(r,r)$ as the joint profits of P and the downstream firms if all downstream firms pay royalties $r$. We assume that $T_I(r)$ is increasing with $r$ in the range $0 \leq r \leq v$, now even if $N = 1$. With these definitions, the analysis proceeds just as in the non-integrated case, using $T_I(r)$ rather than $T(r)$ and $\pi^I(r,r)$ rather than $\pi(r,r)$. So $r_I(\theta)$ maximizes $G^I(r,\theta) = T_I(r) - \theta N\pi^I(v,r) - (1-\theta)N\pi^I(0,0)$ subject to $r \leq v$.

If $N = 1$ then $G^I(r,\theta) = T_I(r) - \theta\pi^I(v) - (1-\theta)\pi^I(0)$, which increases with $r$ in $r \leq v$, so $r_I(\theta) = v$ for all $\theta$. A similar logic applies if downstream firms using the backstop technology

cannot profitably compete against the patent holder: $\pi^I(v,v) = 0$. This condition implies that $\pi^I(v,r) = 0$ for all $r \le v$, so $G^I(r,\theta) = T_I(r) - (1-\theta)N\pi^I(0,0)$, and thus $r_I(\theta) = v$ for all $\theta$.

More generally, we have $r(\theta) = v$ for all $\theta \le \theta_{VI}$ where $\theta_{VI} \equiv \dfrac{T_I'(v)}{N\pi_2^I(v,v)}$. As in the non-integrated case, $B(\theta) > 0$ for all $\theta$, and $K(\theta) < 0$ for weak patents.

The analysis without negative fixed fees also closely parallels the case of the non-integrated patent holder. P still sets the highest acceptable royalty rate, all the more so if $\phi'(r) > 0$. The same acceptance condition applies, using $\pi^I(a,b)$ instead of $\pi(a,b)$. Therefore, for small values of $\theta$, we get $s(\theta) \approx \rho^I \theta v$ where $\rho^I = \dfrac{\left|\pi^I_1(0,0)\right|}{\left|\pi^I_1(0,0) + \pi^I_2(0,0)\right|}$. With a symmetric Cournot oligopoly, the Appendix shows that $\rho^I = \dfrac{N+1}{2}$. For small $v$ we again have $B(\theta) > 0$ for all $\theta$.

## Short-Term vs. Long-Term Licenses

In equilibrium, there is no litigation, so the only impact of using short-term rather than long-term licenses is on the payoff to a downstream firm that infringes rather than accepts a license. In the analysis above with unrestricted two-part tariffs, for a given $\theta$, if the equilibrium running royalty rate is $r(\theta)$ with long-term licenses, this reservation payoff was $\theta\pi(v,r(\theta)) + (1-\theta)\pi(0,0)$. With short-term licenses, this reservation payoff becomes $\theta\pi(v,r(1)) + (1-\theta)\pi(0,0)$. The patent holder has an incentive to use short-term licenses if and only if $r(1) < r(\theta)$. If $r(1) = v$, then $r(\theta) = v$ for all $\theta$ and the patent holder is indifferent between using short-term and long-term licenses. However, if $r(1) < v$, then $r(1) < r(\theta)$ for all $\theta < 1$ and the patent holder strictly prefers to use short-term licenses. Define $\theta_{ST}*$ such that $\pi(v,v) = \theta\pi(v,r(1)) + (1-\theta)\pi(0,0)$. For $\theta > \theta_{ST}*$, the downstream firm's threat point is to use the backstop technology rather than infringe and engage in litigation. For all $\theta$, the downstream firm's reservation payoff is independent of $r$, so the patent holder has no incentive to reduce $r$ below $v$. Therefore, $r = v$ for all patent strengths. Negative fixed fees are used for all $\theta < \theta_{ST}*$; no fixed fee is used for $\theta \ge \theta_{ST}*$. The *ex post* welfare analysis is exactly the same as the case already studied in which $\theta_V = 1$. *Ex ante,* for all $\theta < 1$, $P(\theta)$ is higher than we had earlier and $K(\theta)$ is unchanged or lower, so the bias resulting from $P(\theta)/K(\theta) > P(1)/K(1)$ is stronger than we had earlier.

If negative fixed fees are not feasible, with long-term licenses the downstream firm's payoff from infringing was $\theta\pi(v,s(\theta)) + (1-\theta)\pi(0,0)$. With short-term licensees, for $\theta < \theta_{ST}*$ this payoff becomes $\theta\pi(v,r(1)) + (1-\theta)\pi(0,0)$, so the patent holder has an incentive to use short-term licenses if and only if $s(\theta) > r(1)$. With $r(1) > 0$, this condition will not be met for relatively weak patents, so owners of those patents will use long-term licenses. For stronger patents, the patent holder has an incentive to use short-term licenses. For $\theta \ge \theta_{ST}*$, the

downstream firm's threat point is to use the backstop technology, so $s(\theta) = v$. Since the ability strategically to use short-term licenses raises $s(\theta)$, our welfare results are strengthened.

## Consumer and Downstream Firm Risk Preferences on Linear Royalties

Each downstream firm's payoff from $s$ is $\pi(s) = [p(x(s)) - s]x(s) = \max_x[[p(x) - s]x]$. For each $x$, $[p(x) - s]x$ is linear and decreasing in $s$, so the upper envelope $\pi(s)$ is convex and decreasing in $s$. Thus (as is well known), the downstream firm prefers lower $s$ but is risk-loving in $s$. Since $\pi'(s) = -x(s)$ and $\pi''(s) = -x'(s)$, the downstream firm's risk preference in $s$ is measured by the coefficient of "absolute risk aversion" $\dfrac{\pi''(s)}{-\pi'(s)} = \dfrac{-x'(s)}{x(s)}$. (Because $\pi(s)$ is decreasing, this standard "risk aversion coefficient" is mathematically positive even though the downstream firm is risk-preferring.) Turning to consumers, write $V(s)$ for consumer surplus, and $p(s)$ for downstream price, as functions of $s$. Then $V'(s) = -p'(s)x(s)$. If pass-through is linear, $p''(s) = 0$, then $p'(s)$ is a positive constant, so $V'(s)$ is a preference-preserving transformation of $\pi'(s)$, so consumers have the same risk attitudes as the downstream firm. In more detail, we have

$$\frac{V''(s)}{-V'(s)} = \frac{-x'(s)}{x(s)} - \frac{p''(s)}{p'(s)},$$

so, when $p''(s) = 0$, consumers' risk preference coefficient equals $\dfrac{V''(s)}{-V'(s)} = \dfrac{-x'(s)}{x(s)}$, which (we just saw) is also the risk preference coefficient for the downstream firm. If price is convex in $s$ then consumers are less risk-loving in $s$ than the downstream firm; if price is concave, they are more risk-loving.

## Benefits of Ideal PTO Review with Linear Licenses

By Theorem 1, $B(\theta) \geq [r(\theta) - \mu\theta r(1)]\lambda_{\min}$. With linear licenses, $r(1) = v$ so $B(\theta) \geq \theta v\left[\dfrac{r(\theta)}{\theta v} - \mu\right]\lambda_{\min}$ and $K(1) = vx(v,v)$. In general $B(\theta) + K(\theta) = \theta K(1)$; with $r(1) = v$ this gives $B(\theta) + K(\theta) = \theta vx(v,v)$. Therefore, $\dfrac{B(\theta)}{B(\theta) + K(\theta)} \geq \left[\dfrac{r(\theta)}{\theta v} - \mu\right]\left[\dfrac{\lambda_{\min}}{x(v,v)}\right]$. As shown in the proof of Theorem 1, with Cournot oligopoly $|w'(r)| = p'(r)\left[\dfrac{p(r) - c}{p(r) - r - c}\right]x(r,r)$. Therefore

$\lambda_{\min} \geq p'(r)x(v,v)$, and $\dfrac{B(\theta)}{B(\theta) + K(\theta)} \geq \left[\dfrac{r(\theta)}{\theta v} - \mu\right]p'(r)$.