Patent Royalties When Imitation Is Costly

Fernando Leiva B. and John L. Turner[†]

January 2012

Abstract

We identify welfare-maximizing patent royalties in a model of costly imitation, entry and imperfect competition. When the social planner may impose a compulsory license, optimal royalties either blockade entry, facilitating unregulated monopoly, or yield an aggregate-zero-profit duopoly. When duopoly is optimal, the optimal per-unit royalty pins the equilibrium price at the aggregate average cost and the optimal fixed royalty shifts surplus so the patentee and imitator break even. Interestingly, royalty payments with such royalties may be negative. Because of this, aggregate-zero-profit duopoly may be impossible to achieve if the planner must instead direct the courts to use such royalties.

JEL Code: K2, O3
Keywords: patents, damages, royalties, invention, entry

[†]Leiva: Department of Economics, Arizona State University. Main Campus, PO BOX 873806, Tempe, AZ 85287-3806. Phone: 480-965-2997. E-mail: fernando.leivabertran@asu.edu.
Turner: Department of Economics, University of Georgia, Brooks Hall 5th Floor, Athens, GA 30602-6254. Phone: 706-542-3682. E-mail: jlturner@uga.edu. We thank Matt Mitchell for helpful comments.

1. Introduction

Though highly variable, imitation is typically less costly than invention, even when patents are present. In their survey of firms from chemical, drug, electronics and machinery industries, Mansfield, Schwartz and Wagner (1981, pp. 909-13) find that imitation costs and time are on average about two-thirds of invention costs and time. In about half of cases, imitation costs are either lower than 40% or higher than 90%. The presence of patents raises imitation costs by 11% at the median.[1]

In addition, patents are enforced imperfectly and patentees frequently lose in court. Henry and Turner (2006) estimate that from 1983-2002, about 28% of patents were found "invalid," about 35% were found "not infringed" and about 37% were found "valid and infringed." As a result, rival firms often find imitation and entry prior to a patent's expiration to be profitable. This is true even when such behavior is likely to trigger litigation.[2]

Therefore, the presence of patents often facilitates competition between a small number of firms. Such competition tends to be imperfect, characterized both by equilibrium price markups and potentially wasteful duplication of fixed costs. Optimal policy must weigh such considerations. Surprisingly, reward-theory approaches to optimal patent policy (e.g. Nordhaus 1969; Gilbert and Shapiro 1990; Klemperer 1990) and optimal compulsory licensing (Tandon 1982) largely ignore the possibility that imitators enter and compete imperfectly in the shadow of royalties. On the other hand, papers directly considering how patent royalties and other damages influence competition (e.g. Lanjouw and Schankerman 2001; Anton and Yao 2007; Choi 2009; Henry and Turner 2010) do not consider how the cost of entry by imitators affects optimal damages.

In this paper, we introduce a single-period model of endogenous invention, imitation, entry and imperfect competition where marginal production costs are constant and imitation is costly but cheaper than invention. We study three questions. First, assuming a monopolist may not be regulated, what is the welfare-maximizing, feasible level of entry and output?

[1] In their more comprehensive survey, Levin et al. (1987) find imitation costs to be about 85% for patented products and 65% for unpatented products.

[2] For example, when a generic drug manufacturer files a paragraph IV Abbreviated New Drug Application (ANDA) with the US Food and Drug Administration (FDA), it attempts to directly imitate a patented active ingredient prior to the patent's expiration. Because it is permitted to rely on the results of clinical trials conducted by the original manufacturer, its imitation costs are relatively small. Such filings nearly always trigger litigation.

Second, what royalties would a planner choose for a compulsory license to achieve that level of entry and output? Third, could the planner achieve optimal entry and output by instead directing the court system to use particular damages, then letting firms choose to sue for damages in court?

We show that optimal royalties either blockade entry in the sense of Bain (1956), facilitating unregulated monopoly, or encourage one entrant and achieve an aggregate-zero-profit duopoly.[3] This yields an intuitive sufficient (but not necessary) condition for the optimality of duopoly. If there exist royalties that yield an aggregate-zero-profit duopoly—which we term an "efficient" duopoly—that lowers the average cost of production (including innovation and imitation costs) relative to unregulated monopoly, then efficient duopoly maximizes welfare. Thus, if invention costs are low enough to make monopoly profitable, then efficient duopoly is optimal for imitation costs sufficiently close to zero.

Paradoxically, the royalties that yield efficient duopolies have characteristics that differ markedly from royalty damages assigned by courts in practice. Most notably, efficient royalties depend crucially on invention and imitation costs. Per-unit royalties adjust the effective marginal cost of imitators and are therefore useful for fine-tuning price and output, provided there is at least one imitator. The efficient per-unit royalty sets the equilibrium price at the aggregate average cost of production. Fixed royalties are useful for fine-tuning the level of entry because they alter payoffs without distorting output. The efficient fixed royalty shifts surplus so that the patentee and imitator each break even.

When aggregate invention and imitation costs are lower than the aggregate variable profit (i.e. net of fixed costs) from a no-royalty duopoly, the efficient total royalty payment is positive but the efficient per-unit royalty is *negative*. A negative per-unit royalty subsidizes production by the imitator, increasing total output and lowering aggregate variable profit to the break-even point. Positive fixed royalties shift surplus from the imitator to the inventor so that both firms break even.

When aggregate invention and imitation costs are higher than the variable profit from a no-royalty duopoly, the efficient per-unit royalty is positive. This lowers output and raises aggregate variable profit to the point where both firms can break even. In this case, efficient

[3] In a spatial model with a single potential entrant, Waterson (1990) also notes that monopoly may be optimal.

fixed royalties may be either positive or negative.

We then consider whether optimal entry and output can be achieved through court-imposed royalties. There are two reasons why this may be impossible, covering a large number of cases. First, courts may follow rules such that per-unit royalties and fixed royalties are individually required to be non-negative. If so, then in all cases where variable profit under a no-royalty duopoly exceeds aggregate invention and imitation costs, it is not possible to achieve optimal entry and efficient output because negative per-unit royalties are necessary for efficient duopoly.[4] For higher invention and imitation costs, efficient fixed royalties may be negative, so such cases may also be problematic.

Second, with court-imposed royalties, the patentee must willingly sue for royalty payments to be realized. Because of this, a credibility constraint emerges—the total royalty payment must be non-negative. We show that the total royalty payment with efficient royalties is negative for sufficiently high invention and imitation costs. Hence, the credibility constraint binds. In particular, if such cases coincide with cases where welfare is higher under efficient duopoly than under monopoly (which require sufficiently *low* invention and imitation costs), then optimal entry and output are impossible to achieve with court-imposed royalties.

In a class of demand functions that includes the linear and log-linear specifications, we study the impact of the credibility constraint and show that the curvature of demand is crucial. When curvature is low (the extreme case being linear demand), monopoly is always optimal in cases where efficient duopoly requires a negative total royalty payment. Hence, the credibility constraint is irrelevant. However, as curvature increases, demand elasticity increases for aggregate output above the monopoly level. This makes the effect of duopoly competition on profit more severe for an imitator that pays a positive per-unit royalty, necessitating efficient royalties that yield a higher transfer from patentee to imitator. Higher curvature also increases the welfare gain under efficient duopoly relative to monopoly. When curvature is sufficiently high (the extreme case being log-linear demand), cases emerge where the total royalty payment is negative under efficient duopoly and efficient duopoly yields more welfare than monopoly. Hence, the credibility constraint prevents optimal entry and

[4]For clarity, we refer to welfare-maximizing output conditional on a given level of entry as "efficient output" and welfare-maximizing entry as "optimal entry."

output under court-imposed royalties, and compulsory licensing may be necessary for optimal welfare.

In practice, courts generally treat patent infringement as a tort and base damages on compensation for the injury caused by infringement. Patent infringers are considered "tort-feasors" and courts in England and the US have consistently sought to identify damages equal to the value of property taken (Lipscomb, 1989, pp. 5-27). The one notable exception to this is that the various patent laws have permitted damages to be increased in cases where infringers' conduct is deemed egregious—under current US law, damages may be trebled if infringement is found to be "willful."

Under United States law, a patentee whose patent is found valid and infringed is entitled to "damages adequate to compensate for infringement but in no event less than a reasonable royalty." (US Code Title 35, Part III, Chapter 29, Section 284). Specifically, a patentee is entitled to "lost profits" if it can satisfy the four-factor test described in *Panduit Corp. v. Stahlin Bros. Fibre Works Inc.* [197 USPQ 726 (6 Cir), 1978]. Otherwise it is entitled to royalties. In determining royalty damages, US courts follow the 15-step procedure under *Georgia-Pacific Corp. v. U.S. Plywood-Champion Papers Inc.* [166 USPQ 237-38 (SDNY 1970)]. The 15-step procedure does not explicitly incorporate invention and imitation costs.[5] It also does not yield royalties with negative components.[6]

Our results are at odds with conventional economic wisdom (e.g. Schumpeter 1942), about the necessity and optimality of monopoly to promote R&D. Consider the pharmaceutical industry, which is often the first example offered to defend blockading patent protection as an incentive to invention (Mazzoleni and Nelson 1998). DiMasi, Hansen and Grabowski (2003) estimate average R&D costs of about $800 million for a new drug, while the costs of imitation are a very small fraction of this due to the FDA policy of allowing generic manufacturers to bypass most expenses associated with clinical trials.[7] It is widely argued that monopoly is an absolute necessity for pharmaceuticals precisely because of this high-invention-cost, low-imitation-cost structure,[8] and to this end the FDA *adds* a set of non-patent exclusivities for new chemical entities, orphan drugs, new indications, new for-

[5]See Appendix A for the full list of factors.

[6]Cooley (1993) analyzes damages from 152 cases over 1982-92. Positive fixed and/or royalty damages were paid in 71 of these cases and lost profits were paid in the other cases.

[7]See Voet (2011).

[8]See Boldrin and Levine (2008) for a thorough review of the literature and an extensive counter-argument.

mulations and several other classes of inventions.[9] Yet our results suggest that, in industries with such low imitation costs, duopoly with appropriate royalty payments is quite likely to improve welfare.

2. The Model

Let there be an inventor and an infinitely elastic supply of potential imitators. They play a four-stage game: 1. Invention, 2. Imitation/Entry, 3. Competition, 4. Royalty Payments. The inventor must first pay cost $F > 0$ to invent and patent a new product. Next, any imitator must pay cost θF to imitate, where $\theta \in [0, 1]$. If the inventor does not invent, then there can be no imitation. We say that any firm that either invents or imitates *enters* the market, and henceforth refer to an inventor who enters as the *patentee*. The firms reach a subgame-perfect equilibrium.

Marginal cost is a constant $C > 0$ for all firms. This simplifies the welfare analysis since, conditional on entry, aggregate output alone (and not its distribution among firms) affects aggregate welfare. Inverse demand is given by $P(Q)$, where Q is aggregate output. All imitators must pay to the patentee *fixed* royalty K, plus *per-unit* royalty r for each unit produced. Hence, the total royalties paid by firm j is $R_j = K + rq_j$.

Let N represent the total number of firms that enter. Conditional on N and factoring in (future) royalty payments, firms compete in quantities $\{q_1, ..., q_N\}$. Letting the patentee be indexed by 1, we write profits π_i excluding fixed costs, henceforth *variable profits*, as follows:

$$\begin{aligned} \pi_1(q_1, ..., q_N; N, r, K) &= Max_{\{q_1\}} \quad [P(Q) - C]q_1 + [r(\textstyle\sum_{j \neq 1} q_j) + (N-1)K] \\ \pi_j(q_1, ..., q_N; N, r, K) &= Max_{\{q_j\}} \quad [P(Q) - C]q_j - (rq_j + K), \text{ for } j \neq 1. \end{aligned} \tag{1}$$

The firms reach a Cournot-Nash equilibrium in the competition stage. To guarantee the existence and uniqueness of an equilibrium at the competition stage, let $P(Q)$ be twice continuously differentiable, strictly decreasing in Q as long as P is positive, and log-concave in Q.

Let $\{q_P, q_I\}$ denote the equilibrium output levels for the patentee (P) and the $N - 1$ symmetric imitators (I) and let aggregate equilibrium output be Q_N. Then equilibrium

[9] New chemical entity and orphan drug exclusivities last for seven years. Most others last for three years. See Voet (2011).

total profits, conditional on entry and the royalty rates, satisfy:

$$\begin{aligned}\Pi_P(N, r, K) &= [P(Q_N) - C]q_P + (N-1)(rq_I + K) - F \\ \Pi_I(N, r, K) &= [P(Q_N) - C]q_I - (rq_I + K) - \theta F.\end{aligned} \tag{2}$$

It is useful to define monopoly output as Q_M and monopoly variable profit as π_M. Define also the aggregate variable profit under duopoly, absent royalties, as π_D.[10]

3. The Planner's Problem

The social planner cannot directly regulate price or output but may choose royalties for a compulsory license and costlessly enforce the terms of the license. Hence, the planner cannot regulate a patentee-monopolist but can indirectly affect price and output if there is at least one entrant. In manipulating price and output, the planner may also affect payoffs and the level of entry. The planner's objective is to maximize the sum of firm profits and consumer surplus.

The planner is constrained by feasible levels of entry. For any number of firms N, the aggregate variable profits earned in equilibrium must exceed the total invention and imitation costs. Since aggregate variable profit is bounded above by the monopoly variable profit π_M, free entry results in N firms in equilibrium (regardless of royalties) only if $\pi_M \geq F + (N-1)\theta F$. Because the monopoly profit does not depend on N and the total entry costs increase with N, this bounds N from above. Define $\overline{N}$ to be the maximum possible N satisfying this constraint.

Suppose $\overline{N} = 1$. Although profitable duopoly is not feasible, if imitation costs and royalties are sufficiently low it will be privately profitable for an imitator to enter once invention has occurred. Such entry would result in a negative profit for the inventor. Invention would not obtain in a subgame-perfect equilibrium. Hence, if $\overline{N} = 1$, royalties should be set sufficiently high enough to blockade entry in the sense of Bain (1956) and achieve unregulated monopoly. Any sufficiently high fixed or per-unit royalty will blockade entry. For example, setting a fixed royalty that is higher than monopoly profits or setting a per-unit royalty that is higher than monopoly price will work.

[10] When we use the subscript "D" we refer to a no-royalty duopoly. When we use the subscript "2" we refer to duopoly where royalties may be present.

The more interesting case is where $\overline{N} \geq 2$, where some entry is feasible. Conditional on N, equilibrium satisfies the Cournot-Nash equilibrium conditions, where output satisfies

$$\begin{aligned} P(Q_N) + P'(Q_N)\, q_P &= C \\ P(Q_N) + P'(Q_N)\, q_I &= C + r. \end{aligned} \tag{3}$$

Factoring in the entry decision, profits must also satisfy

$$\begin{aligned} &\Pi_P(N, r, K) \geq 0 \\ &\Pi_I(N, r, K) \geq 0 \\ &\Pi_I(N+1, r, K) < 0. \end{aligned} \tag{4}$$

For any N satisfying $2 \leq N \leq \overline{N}$, define $\overline{Q}_N(F, \theta)$ to be the highest Q such that

$$[P(Q) - C]Q = F + (N-1)\theta F.$$

Intuitively, $\overline{Q}_N(F, \theta)$ is the highest level of aggregate output such that aggregate variable profits are sufficient to cover aggregate entry costs. At this output, the non-negative profit constraints in (4) can both hold only if they bind. Price equals the aggregate average cost of production:

$$P(\overline{Q}_N(F, \theta)) = C + \frac{F + (N-1)\theta F}{\overline{Q}_N(F, \theta)}.$$

Thus, conditional on N, $\overline{Q}_N(F, \theta)$ yields the highest feasible consumer surplus and welfare.

The following result shows that the planner can achieve $\overline{Q}_N(F, \theta)$ through the effect of royalties on competition.[11]

Lemma 1. *Conditional on N firms competing, where $2 \leq N \leq \overline{N}$, there exist royalties (r, K) such that $\overline{Q}_N(F, \theta)$ is the equilibrium aggregate output.*

Intuitively, the per-unit royalty determines the effective marginal cost of imitators. By setting $r > 0$, the planner effectively taxes output by imitators and lowers total output. By setting $r < 0$, the planner effectively subsidizes output by imitators and raises total output. Total output varies continuously with r. Hence, there exists an r such that $Q_N = \overline{Q}_N(F, \theta)$.

Conditional on N, fixed royalties do not affect downstream competition, but do affect total profit and entry. By setting $K > 0$, the planner lowers the imitators' profits and raises

[11] All proofs are in Appendix B.

the patentee's profit. Profits vary continuously with K. Hence, if output $\overline{Q}_N(F,\theta)$ is chosen, the planner can adjust profits for all firms to zero, guaranteeing invention and the desired level of entry.

Now consider the problem of choosing the optimal level of entry. Since aggregate variable profit is decreasing in aggregate output when $Q_N > Q_M$, it is clear that, for $\theta > 0$, $\overline{Q}_N(F,\theta)$ is strictly lower when N is higher. Thus, as N increases, the maximum possible consumer surplus falls while entry costs rise. Welfare falls.

Hence, it is never optimal to induce entry by more than one imitator. As we will frequently compare welfare under duopoly with output $\overline{Q}_2(F,\theta)$ to welfare under monopoly, the following definition is convenient.

Definition 1. *We define* (r_2^*, K_2^*) *as* **efficient duopoly royalties** *if they induce invention, entry by one imitator and equilibrium output* $\overline{Q}_2(F,\theta)$*.*

Welfare under monopoly is higher than under efficient duopoly when the extra output under efficient duopoly generates an increase in consumer surplus, CS, that is less than the unregulated monopoly profit. We have the following result.

Proposition 1. *Let* $\overline{N} \geq 2$*. If* $CS(\overline{Q}_2(F,\theta)) - CS_M > \pi_M - F$*, then efficient duopoly is optimal. Otherwise, monopoly is optimal.*

Denote the aggregate average costs for monopoly and duopoly as AC_M and AC_2, respectively. If the imitation cost is near zero, then AC_2 nearly equals AC_M and the planner can lower average cost and significantly expand output by using royalties to achieve efficient duopoly. For a significant imitation cost, however, AC_2 is significantly higher than AC_M and monopoly may be optimal.

Figure 1 illustrates a case where $\overline{N} = 2$ but monopoly is optimal. The additional consumer surplus from efficient duopoly is the sum of areas A and C, while the reduction in aggregate profit from efficient duopoly is the sum of areas A and B. As shown in the diagram, C is smaller than B, so monopoly is optimal.

In general, C will be larger than B when the imitation cost is sufficiently small.

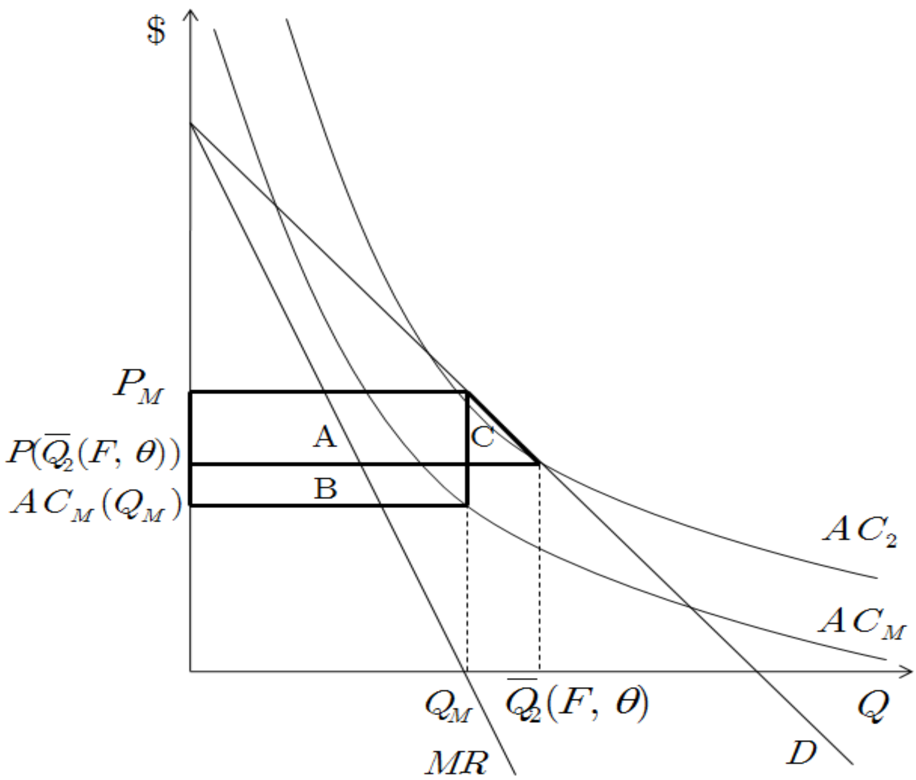

Figure 1: *Monopoly vs. Efficient Duopoly*

Corollary 1. *Suppose $\pi_M > F$. Then efficient duopoly is optimal for a sufficiently small positive imitation cost.*

Intuitively, if $\theta = 0$, then the efficient duopoly price is below the equilibrium average cost under monopoly (there is no area B) so efficient duopoly is optimal. A small θ will not alter this result.

3.1. Characterization

Efficient duopoly royalties are summarized in Table 1. The royalty payment is a combination of positive and/or negative fixed and per-unit royalties. The intuition is simplest for the sign of the efficient per-unit royalty.

Proposition 2. *The efficient per-unit royalty $r_2^* < 0$ if and only if the aggregate variable profit under duopoly absent royalties exceeds the sum of the cost of invention and the cost of imitation by a single entrant, $\pi_D > F(1 + \theta)$.*

Primitives	Efficient Per-Unit Royalty r_2^*	Efficient Fixed Royalty K_2^*
$\pi_D > F(1+\theta)$	-	+
$\pi_D = F(1+\theta)$	0	+
$\pi_D < F(1+\theta)$	+	- or +

Table 1: *Sign of Efficient Duopoly Royalties,* $\theta < 1$

Start by imagining that $r = 0$ but K may be non-zero. Then aggregate variable duopoly profit is that of a standard symmetric Cournot duopoly. If this exceeds $F(1+\theta)$, then since the fixed royalty merely transfers surplus between the firms, there is *some* K such that the inventor earns back more than its invention cost and one imitator earns back more than its imitation cost. Hence, when $\pi_D > F(1+\theta)$, the planner can increase welfare by lowering $r < 0$ to increase output to $\overline{Q}_2(F, \theta)$. This lowers aggregate variable profits to $F(1+\theta)$. Since a negative per-unit royalty transfers surplus from the patentee to the imitator and the imitator has an entry cost advantage, the fixed royalty must be positive for both firms to earn zero profit.

Corollary 2. *If* $\pi_D > F(1+\theta)$, *then* $K_2^* > 0$.

On the other hand, suppose $\pi_D = F(1+\theta)$. If $\theta = 1$, then $r_2^* = 0$ and $K_2^* = 0$. If $\theta < 1$, then $K_2^* > 0$.

Finally, if $\pi_D < F(1+\theta)$ and $r = 0$, then there is no K such that two firms earn back their fixed costs. To achieve entry by an imitator, it is necessary to reduce output toward the monopoly level. A positive per-unit royalty accomplishes this, enhancing the patentee's payoff and lowering the imitator's payoff. When $\theta < 1$ and the per-unit royalty necessary to achieve $\overline{Q}_2(F, \theta)$ is small, the decrease in the imitator's payoff is less than its entry cost advantage, so the efficient fixed royalty $K_2^* > 0$. For higher F, however, it is possible to have $K_2^* < 0$. We illustrate these phenomena with the example that follows.

3.2. An Example

Let inverse demand be given by $P(Q) = \left(1 + \frac{1-Q}{t}\right)^t$, where $t > 0$. This specification includes as special cases linear demand ($t = 1$) and log-linear demand ($t = \infty$), and has the useful property that the monopoly price and output $P_M = 1$ and $Q_M = 1$ are independent of t. Let each firm produce at zero marginal cost.

Consider first the case $t = 1$, so that $P = 2 - Q$. Then $\pi_M = 1$, $\pi_D = \frac{8}{9}$ and the Cournot-Nash equilibrium for $N = 2$ satisfies $q_P = \frac{2+r}{3}$ and $q_I = \frac{2(1-r)}{3}$, so that aggregate output satisfies $Q_2 = \frac{4-r}{3}$.

The level of output such that $N = 2$, and both firms exactly break even, satisfies $\left[2 - \overline{Q}_2(F,\theta)\right] \overline{Q}_2(F,\theta) = F(1+\theta)$, so that

$$\overline{Q}_2(F,\theta) = 1 + \sqrt{1 - F(1+\theta)}.$$

Output $\overline{Q}_2(F,\theta)$ obtains and the firms earn zero profits when

$$\begin{aligned} r_2^* &= 1 - 3\sqrt{1 - F(1+\theta)} \\ K_2^* &= 4 - F(4+5\theta). \end{aligned}$$

The efficient per-unit royalty is negative and the efficient fixed royalty is positive when

$$F < \frac{8}{9(1+\theta)}. \tag{5}$$

To induce output $\overline{Q}_2(F,\theta)$, the planner sets a negative per-unit royalty to subsidize production by the imitator. The positive expected fixed royalty K_2^* shifts profit from the imitator to the patentee so that $\pi_P = F$ and $\pi_I = \theta F$.

Both types of efficient royalties are positive when

$$\frac{8}{9(1+\theta)} < F < \frac{4}{4+5\theta}. \tag{6}$$

The intuition is simplest for the case $\theta = 0$, where (6) becomes $F \in (\frac{8}{9}, 1) \equiv (\pi_D, \pi_M)$. Duopoly is not feasible with fixed royalties only, because the aggregate variable profit is too small. A positive r_2^* increases aggregate variable profit to F. Because the second firm pays no actual fixed cost, however, to move its profit to zero requires a positive fixed royalty.

The efficient per-unit royalty is positive while the efficient fixed royalty is negative if

$$\frac{4}{4+5\theta} < F < \frac{1}{1+\theta}. \tag{7}$$

Intuitively, as $F(1+\theta)$ gets close to $\pi_M = 1$, per-unit royalties must be set to achieve nearly the monopoly profit. In doing this, q_I must shrink to zero. As this occurs, the imitator's revenue shrinks to zero, so a negative fixed royalty is necessary to induce entry.

Consider now the question of whether efficient duopoly or unregulated monopoly is optimal. If $F \in \left(\frac{1}{1+\theta}, 1\right]$, then $\overline{N} = 1$ and monopoly is optimal. If $F = \frac{1}{1+\theta}$, then efficient duopoly is feasible but monopoly is optimal. Intuitively, the monopoly price obtains in either case, but entry costs are higher under duopoly.

If $F < \frac{1}{1+\theta}$, either duopoly or monopoly may be optimal. Welfare under efficient duopoly equals consumer surplus exactly (since profits are zero), so we have $W_2 = \frac{\overline{Q}_2(F,\theta)^2}{2}$. Welfare under monopoly equals $W_M = \frac{3}{2} - F$. Collecting terms and doing a bit of algebra, we see that welfare under monopoly is higher if

$$F > \frac{\sqrt{(1+3\theta)^2 + 3(1-\theta)^2} - (1+3\theta)}{(1-\theta)^2}. \tag{8}$$

The right hand side is strictly decreasing in θ. Consistent with Corollary 1, for very low θ, duopoly is preferred for all but the highest possible values of F.[12]

Since the right-hand side of (8) is higher than the right-hand side of (5) and the left-hand side (7) for low θ, each of the three combinations of efficient royalties $\{(r_2^* < 0, K_2^* > 0), (r_2^* > 0, K_2^* > 0), (r_2^* > 0, K_2^* < 0)\}$ may hold, depending on the size of F. For higher θ, monopoly is preferred for a wider range of fixed costs. Since the right-hand side of (8) is lower than the right-hand side of (5) and the left-hand side of (7) for sufficiently high θ, all such cases of efficient duopoly require $r_2^* < 0$ and $K_2^* > 0$.

4. Optimal Royalties and the Courts

Now consider the case where the planner cannot impose a compulsory license,[13] but can direct the courts to choose royalties to influence welfare. Under such an arrangement, the planner faces additional constraints. First, the nature of the law may constrain the royalties

[12] As θ approaches 0, the cutoff value of F approaches the monopoly profit, 1.

[13] Compulsory licenses were quite common in the United States during the 1950s, 1960s and early 1970s. In 1981, the Reagan Administration directed the Department of Justice not to use antitrust laws against companies refusing to license their patents and there have been no notable compulsory licenses in the US since then (Scherer 2009).

that may be chosen. Under a body of law based on torts, negative per-unit and fixed royalties may be prohibited. If the courts are restricted this way, then if $\pi_D > F(1+\theta)$, it is not possible to achieve output level $\overline{Q}_2$ because a negative per-unit royalty is necessary.

Second, even if the courts may use negative royalty components, a *credibility constraint* emerges. The patentee will credibly sue for damages only if it finds it profitable to do so. Under costless litigation, the patentee sues only if the total royalty payment is non-negative. Because efficient duopoly royalties may involve a negative per-unit or fixed component, the total royalty payment may be negative. If so, the patentee's credibility constraint prevents efficient duopoly output.

To illustrate the effect of the credibility constraint, we add a litigation stage to the model, which alters the timing of the game: 1. Invention, 2. Imitation/Entry, 3. Competition, 4. Litigation, 5. Royalty Payments. Assume further that litigation is costless,[14] that patents are enforced perfectly and that the credibility constraint must be met.[15] The following definition proves useful.

Definition 2. *We define* (r_2, K_2) *as* **credible royalties** *if they induce a non-negative royalty payment.*

We will also refer to an efficient duopoly as *credible* if the efficient duopoly royalties are credible.

Efficient per-unit royalties are negative when $\pi_D > F(1+\theta)$. However, the following result shows that the total royalty payment is always non-negative in such cases.

Proposition 3. *If* $\pi_D \geq F(1+\theta)$, *then efficient duopoly is credible.*

Intuitively, if $r_2^* < 0$, the imitator produces more output and earns more revenue than the patentee, net of royalty payments. Since the imitator also pays a lower fixed cost of entry, the imitator's profit, net of royalty payments, is higher than the patentee's profit. To make both profits equal zero, the total royalty payment must be positive. Hence, it is possible to

[14]Alternatively, assume a model where litigation is always settled out of court via Nash bargaining. If bargaining shares are equal and the gains to settling equal foregone symmetric litigation costs, then each firm earns its' threat-point payoff, i.e. what it would earn under costless litigation.

[15]With risk-neutral parties, probabilistic patents would introduce merely cosmetic differences to the analysis. Change "royalty" to "expected royalty" and everything else stays the same.

induce efficient duopoly by directing the courts to choose efficient royalties.

On the other hand, consider the case where $\pi_D < (1+\theta)F$ and start by assuming $\theta = 1$. While a no-royalty duopoly is not feasible, duopoly with compulsory royalties is feasible as long as $\pi_M > 2F$. Specifically, efficient royalties would set $r_2^* > 0$. However, the total royalty payment is negative, following similar logic as above. Now, the imitator produces less output and, with $\theta = 1$, does not enjoy an entry-cost advantage. To make both profits equal zero, the total royalty payment must be negative and litigation is not credible. We have the following result.

Proposition 4. *If $\pi_D < (1+\theta)F$ and $\theta = 1$, then efficient duopoly is not credible.*

Finally, for the case where $\pi_D < F(1+\theta)$ and $\theta < 1$, efficient duopoly may be credible provided $F(1+\theta)$ is not too close to π_M.

For the credibility constraint to harm welfare it must be the case that efficient duopoly, while achieving higher welfare than monopoly, is not credible. The former requires invention and imitation costs to be sufficiently low while the latter requires them to be sufficiently high. So it may be possible that the credibility constraint plays no role in obtaining optimal royalties. This would occur if for all instances where efficient duopoly was not credible, monopoly achieved higher welfare anyway. The curvature of the demand function has direct influence on whether the credibility constraint is relevant or not. To gain some insight, we return to the example of section 3.

In the linear demand example ($t = 1$), efficient duopoly royalties are credible provided

$$F \leq \frac{1}{1+\theta+\frac{\theta^2}{4}}. \tag{9}$$

So, for example, if $F \leq 4/9$ then efficient duopoly royalties are credible regardless of θ.

Two conditions must be met for the credibility constraint to harm welfare: First, F must be lower than the right-hand side of (8) so that efficient duopoly is optimal. Second, F must be greater than the right-hand side of (9) so that efficient duopoly royalties are non-credible. For the case of linear demand the right-hand side of (8) is always lower than the right-hand side of (9) and no such cases exist.

However, when t is higher, such cases may exist. Intuitively, at output exceeding the monopoly level $Q = 1$, demand becomes more elastic as t increases. Because of this, the effects of duopoly competition are more severe for an imitator who must pay a positive per-unit royalty. The wedge between the imitator's and the inventor's equilibrium outputs is wider under more elastic demand, and imitator revenue net of fixed royalties is lower. Hence, the fixed royalty must provide an even higher transfer from patentee to imitator.

In addition, the welfare gain from having an efficient duopoly is higher under more elastic demand. In Figure 1, area C grows relative to area B as t increases. Hence, a more elastic demand allows for higher values of F and θ to result in efficient duopoly being optimal.

Each of these two effects increases the likelihood that efficient duopoly yields higher welfare than monopoly but is not credible under court-imposed royalties. In our example, both effects are most pronounced for $t \to \infty$, where inverse demand is log-linear, that is $P(Q) = e^{1-Q}$. Further, let $F = \frac{9}{10}e^{-1/2}$ and $\theta = \frac{6}{9}$. Then since $Q_M = P_M = 1$, it follows that

$$W_M = e - \frac{9}{10}e^{-1/2} - 1 > 0.$$

We also see that $\overline{Q}_2 = \frac{3}{2}$ and $P(\overline{Q}_2) = e^{-1/2}$, so that

$$W_2 = e - \frac{9}{10}e^{-1/2} - \frac{16}{10}e^{-1/2}.$$

Since $\frac{16}{10}e^{-1/2} < 1$, efficient duopoly achieves higher welfare than monopoly.

To achieve aggregate-average cost pricing under duopoly, $r_2^* = \frac{e^{-1/2}}{2}$ must hold. For both firms to earn zero profits, $K_2^* = -\frac{7}{20}e^{-1/2}$. With these royalties, equilibrium output satisfies $q_P = 1, q_I = \frac{1}{2}$. Hence total royalty payment $r_2^* q_I + K_2^* = -\frac{e^{-1/2}}{10} < 0$, so efficient duopoly is not credible. It is therefore necessary to use compulsory licensing to achieve efficient duopoly.

This result is not an artifact of log-linear inverse demand.[16] Whether efficient duopoly royalties can be guaranteed to be credible—for all θ and F such that efficient duopoly maximizes welfare—depends crucially on the curvature of the demand function.

A common measure of curvature is given by the elasticity of the *slope* of demand with respect to price, which we denote ε_{sd}. The interpretation is simple: A one percent increase in price leads to an ε_{sd} percent decrease in the slope of the demand function. In our example, $\varepsilon_{sd} = (t-1)/t$,[17] which gives us an intuitive interpretation of the curvature: With $t = 1$ the

[16] which, naturally, is the most convex inverse demand that still remains weakly log-concave.

[17] because

$$\varepsilon_{sd} = \frac{-P''(Q)P(Q)}{[P'(Q)]^2}. \tag{10}$$

slope of the demand function does not change with an increase in price, so naturally $\varepsilon_{sd} = 0$. As $t \to \infty$, the result of a percentage increase in price approaches (monotonically) an equal percentage decrease in the slope of the demand function, that is $\varepsilon_{sd} \to 1$.[18]

We can now answer the question of exactly how convex demand can be while still guaranteeing positive payments by the imitator (within the limitations of our specification of demand).

Proposition 5. *For all $\varepsilon_{sd} < 0.593$, if efficient duopoly maximizes welfare, then efficient duopoly royalties are credible.*

That is, the amount of curvature that is required to guarantee positive royalty payments is such that a one percent increase in price leads to a less than 0.593 percent decrease in the slope of the demand function. If this is so, then litigation will invariably occur in equilibrium whenever duopoly is optimal.

5. Conclusion

Under the reward theory, optimal royalties limit entry to either zero or one entrant. When invention and imitation costs are very high, optimal royalties blockade entry and achieve unregulated monopoly. On the other hand, when imitation costs are near zero (the classic case motivating the need for patent-based monopoly), entry by one imitator is optimal and the efficient per-unit royalty may be negative. In all cases, optimal royalties depend crucially on fixed invention and imitation costs. Hence, they differ significantly from royalties observed in damages awards from patent infringement suits.

It may also be infeasible to achieve optimal entry and output via the courts. If either the efficient per-unit royalty or the efficient fixed royalty is negative, courts that follow a law based on torts cannot choose optimal damages. Courts not required to follow such a law may still not be able to achieve efficient duopoly because the implied total royalty payment may be negative. When demand has high curvature, negative total royalties may obtain in

[18] Alternatively we can use the elasticity of the slope of *inverse* demand with respect to output as a measure of curvature. In this case the curvature would not be independent of Q, but evaluated at monopoly output $Q = 1$ would also equal $(t-1)/t$. For more on the importance of the elasticity of the slope with Cournot Competition see Vives (1999, pp. 104-5).

cases where efficient duopoly yields more welfare than monopoly. Economic arguments for compulsory licensing are strongest in these cases. Finally, in the one case where efficient per-unit and fixed royalties are both positive, efficient royalties favor imitation and entry over unregulated monopoly. In all cases, the intuition behind these royalties differs significantly from a damages theory based on torts.

Nearly thirty years have passed since Mansfield et al. (1981) established the stylized fact that imitation costs are positive, yet non-prohibitive. Unfortunately, their work has not led to more theoretical work studying implications of this stylized fact. We hope that our paper sparks renewed interest in filling this long-neglected gap in the literature.[19] Our study lends itself to interesting direct extensions.

One could alter the cost structure of firms. With increasing marginal costs, optimal entry could involve more than two firms. With symmetric entry costs, if firms can produce at a globally minimum efficient scale, then a first-best competitive equilibrium will routinely include more than two firms. Whenever the optimal level of entry is feasible with zero royalties, however, the efficient per-unit royalty will be negative because the planner can expand output without reducing entry.

One could also consider alternative models of imperfect competition. Another interesting extension would consider royalty damages in a model of spatial competition. As Henry and Turner (2010) show, per-unit royalty damages affect neither static deadweight losses nor imitator profit in a two-firm Hotelling framework. This suggests that such damages, while useful for elevating the patentee's profit, may be ineffective at curtailing excessive entry. It would be interesting to consider this problem in a more general model with entry, such as the circle model of Salop (1979) or the spokes model of Chen and Riordan (2007).

[19] In their conclusion (p. 917), Mansfield et al. made a specific plea for such work:

> *"Our findings should help to promote the search for more realistic and useful models of the innovative process. In recent years, there has been a tendency for such models to assume that the innovator receives all of the benefits from an invention and that imitation can be ignored...We hope that the excellent theorists who are working in this area will soon be able to relax these assumptions."*

Appendix A

In determining royalty damages, US courts follow the 15-step procedure under *Georgia-Pacific Corp. v. U.S. Plywood-Champion Papers Inc.* [166 USPQ 237-38 (SDNY 1970)]. The factors are:

1. The royalties received by the patentee for the licensing of the patent in suit, proving or tending to prove an established royalty.

2. The rates paid by the licensee for the use of other patents comparable to the patent in suit.

3. The nature and scope of the license, as exclusive or non-exclusive; or as restricted or non-restricted in terms of territory or with respect to whom the manufactured product may be sold.

4. The licensor's established policy and marketing program to maintain his patent monopoly by not licensing others to use the invention or by granting licenses under special conditions designed to preserve that monopoly.

5. The commercial relationship between the licensor and licensee, such as, whether they are competitors in the same territory in the same line of business; or whether they are inventor and promotor.

6. The effect of selling the patented specialty in promoting sales of other products of the licensee; the existing value of the invention to the licensor as a generator of sales of his non-patented items; and the extent of such derivative or convoyed sales.

7. The duration of the patent and the term of the license.

8. The established profitability of the product made under the patent; its commercial success; and its current popularity.

9. The utility and advantages of the patent property over the old modes or devices, if any, that had been used for working out similar results.

10. The nature of the patented invention; the character of the commercial embodiment of it as owned and produced by the licensor; and the benefits to those who have used the invention.

11. The extent to which the infringer has made use of the invention; and any evidence probative of the value of that use.

12. The portion of the profit or of the selling price that may be customary in the particular business or in comparable businesses to allow for the use of the invention or analogous inventions.

13. The portion of the realizable profit that should be credited to the invention as distinguished from non-patented elements, the manufacturing process, business risks, or significant features or improvements added by the infringer.

14. The opinion testimony of qualified experts.

15. The amount that a licensor (such as the patentee) and a licensee (such as the infringer) would have agreed upon (at the time the infringement began) if both had been reasonably and voluntarily trying to reach an agreement; that is, the amount which a prudent licensee – who desired, as a business proposition, to obtain a license to manufacture

and sell a particular article embodying the patented invention – would have been willing to pay as a royalty and yet be able to make a reasonable profit and which amount would have been acceptable by a prudent patentee who was willing to grant a license.

Appendix B

Proof of Lemma 1. For notational simplicity, define $\overline{Q}_N(F,\theta)$ as $\overline{Q}_N$. Also, define royalties that yield $\overline{Q}_N$ as (r_N^*, K_N^*). We require the following four conditions to hold:

$$\begin{array}{lll} (i) & [P(\overline{Q}_N) - C] + P'(\overline{Q}_N)q_P & = 0 \\ (ii) & [P(\overline{Q}_N) - C - r_N^*] + P'(\overline{Q}_N)q_I & = 0 \\ (iii) & [P(\overline{Q}_N) - C]q_P + (r_N^* q_I + K_N^*)(N-1) - F & = 0 \\ (iv) & [P(\overline{Q}_N) - C]q_I - (r_N^* q_I + K_N^*) - \theta F & = 0, \end{array}$$

where $q_I = (\overline{Q}_N - q_P)/(N-1)$.

Adding (iii) and (iv) for all firms, we have

$$\pi\left(Q_N\right) - F - (N-1)\,\theta F = 0$$

This condition determines $\overline{Q}_N$. Such a level of output exists because, due to continuity and log-concavity of inverse demand, $\pi(Q)$ is continuous and strictly decreasing in Q for $Q > Q_M$ and can be as low as needed, for Q high enough.

Adding (i) and (ii) for all firms and rearranging we have:

$$P\left(\overline{Q}_N\right) - C + \pi'\left(\overline{Q}_N\right)/\left(N-1\right) = r_N^*. \tag{11}$$

This condition determines r_N^*. This royalty level exists because $0 < P\left(\overline{Q}_N\right) - C < P\left(Q_M\right) - C$, and $-\infty < \pi'\left(\overline{Q}_N\right) < 0$.

Rearranging (iv) and substituting in using (ii) and (11), we have:

$$\left[\pi'\left(\overline{Q}_N\right)/\left(N-1\right)\right]^2 / \left[-P'\left(\overline{Q}_N\right)\right] - \theta F = K_N^*$$

This condition determines K_N^*. This royalty level exists due to continuity of $P'\left(Q\right)$ and because $-\infty < P'\left(\overline{Q}_N\right) < 0$.

Finally, condition (i) determines $q_P \in (0, \overline{Q}_N)$. Such a level of patentee output exists because for q_P close enough to $\overline{Q}_N$, we have $P\left(\overline{Q}_N\right) - C + P'\left(\overline{Q}_N\right) q_P < 0$ and for $q_P = 0$ we have $P\left(\overline{Q}_N\right) - C > 0$. Thus, by continuity, there exists a level of $q_P \in (0, \overline{Q}_N)$ that satisfies the first condition. **QED**

Proof of Proposition 1. A corollary of Lemma 1 is that if we want to induce any entry at all, we would like it to be minimal. Patentee profit is:

$$\Pi_P = [P(Q) - C]\, Q - F - (N-1)\,\theta F.$$

Keeping it constant, we obtain the following relationship between output and entry:

$$\frac{dN}{dQ} = \frac{\pi'(Q)}{\theta F} < 0,$$

since aggregate output is at higher than monopoly levels.

This negative relationship between output and entry means that reducing entry as much as possible not only reduces the fixed entry costs but also increases aggregate output and, with it, welfare. Because $\overline{N} \geq 2$ then, conditional on entry, the optimal number of firms is $N^* = 2$.

The only instance where we would rather have a monopoly would be where welfare under monopoly exceeded welfare with a duopoly. That is:

$$\pi_M + CS_M - F > \pi_P(\overline{Q}_2) + \pi_I(\overline{Q}_2) + CS(\overline{Q}_2) - F - \theta F = CS(\overline{Q}_2).$$

QED

Proof of Corollary 1. Suppose $\pi_M > F$. Then there exists some $\overline{\theta} > 0$ such that $\pi_M > F + \overline{\theta} F$. Thus, $\pi(\overline{Q}_2(F,\overline{\theta})) < \pi_M$, which implies $\overline{Q}_2(F,\overline{\theta}) > Q_M$. Also, because $\pi'(Q) < 0$ for all $Q > Q_M$, then for all $\theta \leq \overline{\theta}$, $\overline{Q}_2(F,\theta)$ is decreasing in θ.

From Proposition 1, we want to find values of θ such that:

$$\theta F < \pi(\overline{Q}_2(F,\theta)) + CS(\overline{Q}_2(F,\theta)) - \pi_M - CS_M.$$

The left-hand side is increasing in θ, ranging from 0 to F. Also, because $\pi(Q) + CS(Q)$ is increasing in Q and $\overline{Q}_2(F,\theta)$ is decreasing in θ (for $\theta \leq \overline{\theta}$) then the right-hand side is always positive and decreasing in θ for all $\theta \leq \overline{\theta}$. Thus, there exists $\underline{\theta} \leq \overline{\theta}$ such that for all $\theta \leq \underline{\theta}$, efficient duopoly is optimal. **QED**

Proof of Proposition 2. If duopoly is optimal, then the optimal output is $\overline{Q}_2(F,\theta)$. For notational simplicity, let that output be $\overline{Q}_2$. Let aggregate output under a no-royalty duopoly be Q_D, so that $\pi_D = \pi(Q_D)$.

Suppose $\pi_D < F(1+\theta)$. We show the following:

$$\pi\left(\overline{Q}_2\right) = (1+\theta)\, F > \pi\,(Q_D) = -\pi'\,(Q_D)\, Q_D > -\pi'\left(\overline{Q}_2\right)\overline{Q}_2.$$

The first equality holds by construction of the efficient royalty and the first inequality by assumption.

The second equality holds by first order conditions in a symmetric (no royalty) Cournot duopoly. Since $[P\,(Q_D) - C] + P'\,(Q_D)\, Q_D/2 = 0$, we have $P\,(Q_D) - C = -\,[P\,(Q_D) - C] - P'\,(Q_D)\, Q_D$, or $P(Q_D) - C = -\pi^{'}(Q_D)$. Multiplying both sides by Q_D yields $\pi\,(Q_D) = -\pi'\,(Q_D)\, Q_D$.

The last inequality is due to strict concavity of aggregate variable profits at all $Q \in (Q_M, Q_D]$, as we show below:

For any Q, if $P''(Q) \leq 0$ then $\pi''(Q) < 0$. If, on the other hand, $P''(Q) > 0$ let $s \in [0, P(Q_M) - C]$ and let Q be such that:

$$Q = \frac{2[P(Q) - C] - s}{-P'(Q)}.$$

Notice that Q ranges continuously from Q_D at $s = 0$ to Q_M as $s = P(Q_M) - C$. So, for all $Q \in (Q_M, Q_D]$, there exists $s \in [0, P(Q_M) - C]$ such that

$$\pi''(Q) = 2P'(Q) + P''(Q)\frac{2[P(Q) - C] - s}{-P'(Q)} = \left[P''(Q)\left[P(Q) - C - \frac{s}{2}\right] - P'(Q)^2\right]\frac{2}{-P'(Q)} < 0,$$

where the inequality results from log-concavity of inverse demand. Thus, aggregate variable profits are strictly concave for all $Q \in (Q_M, Q_D]$.

Finally, the above implies

$$\pi_M > \pi\left(\overline{Q}_2\right) \geq \pi(Q_D) \Leftrightarrow Q_D \geq \overline{Q}_2 > Q_M \Leftrightarrow -\pi'(Q_D) \geq -\pi'\left(\overline{Q}_2\right) > -\pi'(Q_M) = 0,$$

so that $-\pi'(Q_D)\,Q_D > -\pi'\left(\overline{Q}_2\right)\overline{Q}_2$.

To sum up, we know that

$$\left[P\left(\overline{Q}_2\right) - C\right]\overline{Q}_2 = \pi\left(\overline{Q}_2\right) > -\pi'\left(\overline{Q}_2\right)\overline{Q}_2,$$

so we conclude that

$$r_2^* = \left[P\left(\overline{Q}_2\right) - C\right] + \pi'\left(\overline{Q}_2\right) > 0.$$

Next, suppose $\pi_D = F(1+\theta)$. Then $\overline{Q}_2 = Q_D$ by construction and $r_2^* = [P(Q_D) - C] + \pi'(Q_D) = 0$.

Finally, suppose $\pi_D > F(1+\theta)$. Then $\overline{Q}_2 > Q_D$, so that

$$\pi\left(\overline{Q}_2\right) = (1+\theta)F < \pi(Q_D) = -\pi'(Q_D)\,Q_D < -\pi'\left(\overline{Q}_2\right)\overline{Q}_2.$$

Then $r_2^* = \left[P\left(\overline{Q}_2\right) - C\right] + \pi'\left(\overline{Q}_2\right) < 0$. **QED**

Proof of Corollary 2. Since variable profits are sufficient for a no-royalty duopoly then from Proposition 2 we know that $r_2^* < 0$. Thus, at the Nash Equilibrium we have:

$$q_P = \frac{P\left(\overline{Q}_2\right) - C}{-P'\left(\overline{Q}_2\right)} < \frac{P\left(\overline{Q}_2\right) - C - r_2^*}{-P'\left(\overline{Q}_2\right)} = q_I.$$

Finally, from the free entry condition we know that

$$K_2^* = \left[P\left(\overline{Q}_2\right) - C - r_2^*\right]q_I - \theta F > \left[P\left(\overline{Q}_2\right) - C\right]q_P - F = 0.$$

QED

Proof of Proposition 3. The total royalty payment received by the patentee is $(rq_I + K)$. Consider the case where $r_2^* < 0$ and $K_2^* > 0$. With efficient royalties, the following four conditions must hold:

$$\begin{aligned}
&(i) && [P(\overline{Q}_2) - C] + P'(\overline{Q}_2)q_P && = 0\\
&(ii) && [P(\overline{Q}_2) - C - r_2^*] + P'(\overline{Q}_2)q_I && = 0\\
&(iii) && [P(\overline{Q}_2) - C]q_P + (r_2^* q_I + K_2^*) - F && = 0\\
&(iv) && [P(\overline{Q}_2) - C]q_I - (r_2^* q_I + K_2^*) - \theta F && = 0,
\end{aligned}$$

where $q_I = \overline{Q}_2 - q_P$.

By conditions (i) and (ii) and log-concavity of demand, $q_I > q_P$ when $r < 0$ (see proof of Corollary 2). Given this, then

$$[P(\overline{Q}_2) - C]q_I - \theta F > [P(\overline{Q}_2) - C]q_P - F. \tag{12}$$

Using conditions (iii) and (iv), we see that $[P(\overline{Q}_2) - C]q_P - F = -\left\{[P(\overline{Q}_2) - C]q_I - \theta F\right\}$. Combining (12) with condition (iv), we see that the total royalty payment $(r_2^* q_I + K_2^*) = [P(\overline{Q}_2) - C]q_I - \theta F > 0$. **QED**

Proof of Proposition 4. Let $\theta = 1$. With efficient royalties, the following four conditions must hold:

$$\begin{aligned}
&(i) && [P(\overline{Q}_2) - C] + P'(\overline{Q}_2)q_P && = 0\\
&(ii) && [P(\overline{Q}_2) - C - r_2^*] + P'(\overline{Q}_2)q_I && = 0\\
&(iii) && [P(\overline{Q}_2) - C]q_P + (r_2^* q_I + K_2^*) - F && = 0\\
&(iv) && [P(\overline{Q}_2) - C]q_I - (r_2^* q_I + K_2^*) - F && = 0,
\end{aligned}$$

where $q_I = \overline{Q}_2 - q_P$.

From Proposition 2, $r_2^* > 0$. Then by conditions (i) and (ii) and log-concavity of inverse demand, $q_I < q_P$. Given this then

$$[P(\overline{Q}_2) - C]q_I - F < [P(\overline{Q}_2) - C]q_P - F. \tag{13}$$

Using conditions (iii) and (iv), we see that $[P(\overline{Q}_2) - C]q_I - F = -\left\{[P(\overline{Q}_2) - C]q_P - F\right\}$. Combining this with (13), the total royalty payment $(r_2^* q_I + K_2^*) = [P(\overline{Q}_2) - C]q_I - F < 0$. **QED**

Proof of Proposition 5. With marginal cost $C = 0$ and inverse demand given by:

$$P(Q) = \left(1 + \frac{1 - Q}{t}\right)^t,$$

we can first solve for the non-royalty duopoly:

$$Q_D = 2\frac{1 + t}{2 + t},\ P_D = \left(\frac{1 + t}{2 + t}\right)^t.$$

Then, solving for the efficient royalty duopoly we have:

$$\begin{array}{ll}(i) & R = F - \overline{P}_2\, q_P \\ (ii) & 0 = \overline{P}_2 + \overline{P}_2'\, q_P \\ (iii) & 1 = P_M = Q_M \\ (iv) & D = \frac{t}{t+1}\left(1 - \overline{P}_2\,\overline{Q}_2\right) \\ (v) & F > \overline{P}_2\, Q_M - D \\ (vi) & F + \theta F = \overline{P}_2\,\overline{Q}_2 > P_D\, Q_D,\end{array}$$

where R is the total royalty payment, $\overline{P}_2 = P(\overline{Q}_2)$, $\overline{P}_2' = P'(\overline{Q}_2)$, and D is defined as the deadweight loss that results from having a monopoly instead of an efficient royalty duopoly (or area C in figure 1). After replacing and rearranging we have:

$$\begin{aligned} R &= F - \overline{P}_2\, q_P \\ &= F - \overline{P}_2^{\frac{t+1}{t}} \\ &> \overline{P}_2 - \frac{t}{t+1}\left(1 - \overline{P}_2\,\overline{Q}_2\right) \\ &= (1+t)\,\overline{P}_2 - \frac{1+t+t^2}{1+t}\overline{P}_2^{\frac{t+1}{t}} - \frac{t}{1+t}, \end{aligned}$$

where the second line results from solving for q_P as a function of $\overline{P}_2$ in (ii), the third line from using (v) and then replacing for D with (iv) and for Q_M with (iii), and the last line from solving for $Q\,(P)$ using the inverse demand function, replacing and rearranging.

Since we are assuming that $F + \theta F > P_D\, Q_D$ from (vi), then this implies $1 > \overline{P}_2 > P_D$. Knowing this, the expression on the right hand side of the last line can be shown to be either greater than zero for all $\overline{P}_2$ or minimized at $\overline{P}_2 = P_D$ (because this expression is strictly concave in $\overline{P}_2$, achieves a maximum at $\overline{P}^* \in (P_D, 1)$ and is zero at $\overline{P}_2 = 1$). Thus, after replacing for $\overline{P}_2 = P_D = [(1+t)\,/\,(2+t)]^t$ and rearranging we arrive at $t < 2.457$ or, equivalently, $\varepsilon_{sd} < 0.593$. **QED**

References

Anton, J, and D. Yao. 2007. "Finding 'Lost' Profits: An Equilibrium Analysis of Patent Infringement Damages." *Journal of Law, Economics and Organization* 27, 186-207.

Ayres, I. and P. Klemperer. 1999. "Limiting Patentee's Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies," *Michigan Law Review* 97, 985-1033.

Bain, J. 1956. *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries,* Harvard University Press: Cambridge, MA.

Boldrin, M. and D. Levine. 2008. *Against Intellectual Monopoly.* Cambridge, UK: Cambridge University Press.

Chen, Y., Riordan, M. 2007. "Price and Variety in the Spokes Model," *Economic Journal* 117, 897-921.

Cooley, R. "Overview and Statistical Study of the Law on Patent Damages," *Journal of the Patent and Trademark Office Society* 75, 515-37.

Choi, J.P. 2009. "Alternative Damage Rules and Probabilistic Intellectual Property Rights: Unjust Enrichment, Lost Profits and Reasonable Royalty Remedies," *Information Economics and Policy* 21, 145-157.

DiMasi, J.; Hansen, R.; Grabowski, H. 2003. "The Price of Innovation: New Estimates of Drug Development Costs," *Journal of Health Economics* 22, 151 - 85

Gilbert, R. and C. Shapiro. 1990. "Optimal Patent Length and Breadth," *RAND Journal of Economics* 21, 106-12.

Henry, M. and J. Turner. 2006. "The Court of Appeals for the Federal Circuit's Impact on Patent Litigation." *Journal of Legal Studies* 35, 85-117.

Henry, M. and J. Turner. 2010. "Patent Damages and Spatial Competition," *Journal of Industrial Economics* 58, 279-305.

Klemperer, P. 1990. How Broad Should the Scope of Patent Protection Be?, *RAND Journal of Economics* 21, 113-30.

Levin, R; Klevorick, A; Nelson, R.; Winter, S. 1987. "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Industrial Activity* 3, 1987, 783-831.

Lipscomb, E. 1989. *Lipscomb's Walker on Patents*, Bancroft-Whitney Co.: Rochester, NY.

Mansfield, E.; Schwartz, M.; Wagner, S. 1981. "Imitation Costs and Patents: An Empirical Study," *Economic Journal* 91, 907-18.

Mazzoleni, R. and R Nelson. 1998. "The benefits and costs of strong patent protection: a contribution to the current debate," *Research Policy,* 27(3), pp. 273-284.

Nordhaus, W. 1969. *Invention Growth and Welfare,* Cambridge, Mass. : MIT Press.

Salop, S. 1979. "Monopolistic Competition with Outside Goods," *RAND Journal of Economics* 10, 141-56.

Scherer, F.M. 2009. "The Political Economy of Patent Policy Reform in the United States," *Journal of Telecommunications and High-Technology Law,* 2009, 101-48.

Schumpeter, J. 1942. *Capitalism, Socialism, and Democracy,* Harper Perennial (1962).

Schankerman, M. and S. Scotchmer. "Damages and Injunctions in Protecting Intellectual Property." *RAND Journal of Economics* 32 (2001): 199-220.

Vives, X. 1999. *Oligopoly Pricing: Old Ideas and New Tools,* Cambridge, Mass. : MIT Press.

Voet, M. 2011. *The Generic Challenge: Understanding Patents, FDA and Pharmaceutical Life-Cycle Management (third edition),* Brown Walker Press.