WORKING PAPER SERIES

Patent Licensing Revisited: Heterogeneous Firms and Product Differentiation

Rubén Hernández-Murillo
and
Gerard Llobet

_____

Patent Licensing Revisited: Heterogeneous Firms and Product Differentiation[*]

Rubén Hernández-Murillo[†]   Gerard Llobet[‡]
Federal Reserve Bank of St. Louis   CEMFI

September, 2004

Abstract

In this paper we study the optimal licensing agreement between a patentholder of a cost-reducing innovation and firms that have heterogeneous uses for the new technology. We consider the case in which these firms are competitors in a downstream market. We extend the competition environment among the licensees beyond the Cournot/Bertrand models considered by the previous literature to a framework with differentiated products. We also assume that potential licensees have private information about the usefulness of the new technology. We characterize two purposes the optimal licensing contract serves to the patentholder: separation of the licensees and competition softening in the downstream market. We also describe how the optimal contract changes with the degree of product differentiation.

JEL Classification: L13, O32.

Keywords: patent licensing, royalty rate, fixed fees, private information.

1 Introduction

Obtaining a patent for an innovation is a crucial step in many development processes. A patent allows its owner to exclude other firms from producing a similar good or using a related process. Moreover, a patent allows its owner to license the new technology to other firms. Without the protection of a patent, a competing firm could expropriate the knowledge embedded in the innovation.

In many cases, licensing represents an important share of the innovator's return on the investment. Licensing arises in several circumstances. Often, in markets where inventors are not active

competitors, the new technologies have, as side-products, uses that may differ from the original application. In other cases, inventors might be financially constrained and unable to undertake the necessary investment to market the results of their research. In these two situations, selling the right to use the new technology to other firms seems to be the natural alternative. Finally, litigation processes often result in the licensing of the patents under dispute.[1]

In this paper we study the optimal licensing agreement between a patentholder of a cost-reducing innovation and firms that have heterogeneous uses for the new technology. We consider the case in which these firms are competitors in a downstream market. We extend the competition environment among the licensees beyond the Cournot/Bertrand models considered by the previous literature to a framework with differentiated products. We also assume that potential licensees have private information about the usefulness of the new technology. We find that the optimal licensing contract serves two purposes to the patentholder: separation of the licensees and competition softening in the downstream market. We also describe how the optimal contract changes with the degree of product differentiation.

The initial literature surveyed by Kamien (1992) assumes that there is one patentholder, owner of a cost-reducing innovation, and a number of ex ante identical potential licensees of the new technology. These licensees produce a homogeneous good, and, as in our paper, they compete in a downstream market. A feature of this setup is that for a particular firm, the willingness to pay for the new technology depends on how many licenses are allocated. More licenses imply that more firms have lower costs; therefore, competition becomes fiercer and equilibrium prices lower, reducing profits for all firms. Because this reduction in profits is higher for licensees than for non-licensees,

---

[1]This is the focus of Llobet (2003). In that paper, the patentholder decides to license the innovation to a potential infringer to avoid expensive litigation. One of the most interesting results is that the patentholder does not necessarily benefit from having more protection against future infringement. The reason is that more protection deters the arrival of future improvements on the invention and potential licensees.

firms are willing to pay less for the license. This mechanism results in a downward sloping demand for licenses.

Despite the relevance of patent licensing there are limited data on these contracts. A few empirical studies find that, depending on the environment, royalties, flat fees, or combinations of both can be used.[2] Because of this disparity in the use of contracts, there is an extensive literature that provides conditions under which each type of licensing agreement is optimal for the patentholder. The literature reviewed in Kamien (1992) considers two kinds of competition – Bertrand and Cournot – and two kinds of contracts – flat fees and royalties.[3] In those models fees are better than royalties both from the point of view of the patentholder and for society as a whole. In the case of Bertrand competition, in particular, the optimal royalty is offered to all firms and it is set to compensate for the decrease in marginal cost, making firms indifferent between obtaining a license or not.[4] Therefore, the equilibrium price remains unchanged, in spite of all firms having the new technology. When using flat fees the patentholder offers a limited number of licenses. The equilibrium price is lower in the case of both Bertrand and Cournot competition because flat fees do not distort the incentives of firms. In more recent studies, royalties are found to be superior to fees in the case of risk-sharing (Bouquet et al. (1998)) and strategic delegation (Saracho (2001)).

Our paper departs from the traditional literature on licensing in several dimensions. First, we consider firms to be heterogeneous in their uses for the patent and thus in their willingness to pay for a license. Second, we assume that the decrease in marginal cost that the use of the invention generates is private information. Third, potential licensees produce differentiated goods. Finally, the patentholder has access to a broad set of licensing contracts.

---

[2]Kamien(1992) cites Rostoker (1984) where he finds that royalties plus fees are used in 46% of the cases, royalties alone in 36%, and fixed fees in another 13%.

[3]A remarkable exception that considers a combination of a royalty and a fee is Kamien and Tauman (1984).

[4]These results apply when the innovation is not *drastic*—that is, when the monopoly price with the new technology is above the original marginal cost, arguably the most usual situation.

The existence of private information has been studied recently in contexts where there is limited, if any, competition among licensees. For example, Beggs (1992) and Macho-Stadler and Pérez-Castrillo (1996) examine models in which the innovator uses a combination of royalties and fees to separate licensees with different cost reductions. In contrast, Gallini and Wright (1990) examine the case where the patentholder has private information on the quality of the innovation. Royalties in this case are used by the patentholder to signal the value of the innovation. Choi (2001) and Jensen and Thursby (2001) study the moral hazard case.

In our model, the patentholder offers a menu of contracts, which serves two purposes. First, these contracts allow the patentholder to discriminate among the different firms and maximize the surplus extracted from them, along the lines of the price discrimination literature. Second, and more important, the patentholder uses these contracts to soften competition in the downstream market.

Product differentiation has been introduced in Muto (1993), for example, to show that, for the patentholder, royalties might be a superior licensing agreement compared with fees. Faulí-Oller and Sandonis (2002) and Erutku and Richelle (2001) study the optimal combination of royalties and fees in an environment with differentiated products in which the patentholder is a competitor in the downstream market. Saracho (2002) takes a different approach and uses conjectural variations to model competition among the potential licensees.

The framework we propose allows us to parameterize the degree of product differentiation and characterize how the licensing contract changes from the case where firms produce goods that are independent to the case where firms produce homogeneous goods. When goods are poor substitutes, the licensing contract is geared towards discriminating among licensees, while in the case in which goods are close substitutes, the licensing contract is geared towards softening competition in the

downstream market.

In contrast to the traditional Cournot and Bertrand models used in the literature, we place firms in an environment of monopolistic competition. This model is especially well-suited to analyze changes in product differentiation for two reasons. First, this model abstracts from strategic interaction among the licensees, which allows us to examine how the patentholder's choice of contracts affects downstream competition. Second, unlike the Cournot and Bertrand frameworks, where firms are not atomistic, monopolistic competition allows us to characterize menus of contracts that induce self-selection among the licensees.

Among the class of contracts that induce self-selection, we identify the optimal menu of two-part tariffs, in which the patent holder retains a proportion of the licensee's revenue and charges a flat fee to the firm. Firms with a higher valuation for the innovation choose a contract in which they pay a higher flat fee and retain a higher share of revenues. We show that focusing on two-part tariffs is without loss of generality and the optimal allocation can be implemented with a menu of output-based royalties and fees, more traditional in the literature. Our model predicts that firms with a lower valuation for the innovation choose a contract in which they pay a higher royalty and produce less, which is consistent with evidence found in Taylor and Silberston (1973). As a way to build intuition, however, the use of two-part tariffs is preferred because it provides a clear separation of the discrimination and competition softening features mentioned before.

The optimal contract distorts the pricing decisions of the licensees. In particular, by reducing the marginal revenue of sales, the contract induces firms to sell fewer units at a higher price. Because higher prices by competitors increase own profits, firms are willing to pay more for the license, allowing larger licensing payments. These two forces that induce higher prices in the final good market are limited by the standard *double-marginalization* effect: When the patentholder

demands a higher variable payment, the rents extracted from the licensees are obtained at a cost of generating additional distortions and establishing a price above the monopoly price.

We show that the discrimination achieved among licensees is independent of the differentiation in the final goods that firms produce. In contrast, the competition softening induced by the contract in the downstream market is more important as the degree of product differentiation declines and goods become more homogeneous. As a result, in markets with more homogeneous products the patentholder retains a higher share of revenues, even though the double-marginalization distortion becomes more important.

By the same token, the number of licenses granted raises as goods become more homogeneous. The reason is that the profits of a licensee are influenced to a greater extent by the pricing choices of the competitors. Therefore, selling to additional firms also serves the purpose of dampening competition.

Comparing the optimal two-part tariff contract with a simple flat fee contract we find that the former makes the improvement available to more firms. The intuition of this result is similar to the case of a price-discriminating monopolist that chooses to sell to more consumers than in the case in which price discrimination is not possible. With two-part tariffs, however, the associated distortions in the final good market, consequence of the double marginalization, often reduce social welfare compared with cases where only a flat fee is used. This result is the opposite of that obtained in the price discrimination literature, where a monopolist who distinguishes among different consumers with a menu of tariffs can improve social welfare beyond the level attained by using a uniform price.

The structure of the paper is as follows. In section 2 we introduce the model. In section 3 we describe the contracts that are implementable and the one that maximizes profits for the patentholder; we also show the equivalence between the two-part tariff and a menu of royalties and

fees. In section 4 we introduce the optimal flat fee, and section 5 compares this contract with the optimal two-part tariff. Section 6 concludes.

2 The Model

Consider a market with three types of agents: consumers, producers of a variety of final goods, and a patentholder of a cost-reducing innovation. The structure of the final good market corresponds to the *monopolistic competition* framework introduced by Dixit and Stiglitz (1977). There is a continuum of goods along the unit line, indexed by a parameter x, and a representative consumer that chooses how much of each product to consume, according to the utility function

$$u\left(y\right) = \left(\int_0^1 y\left(x\right)^\lambda dx\right)^{\frac{1}{\lambda}},$$

where $y\left(x\right)$ is the quantity purchased of product x and $\lambda \in (0,1)$ measures how differentiated are the final products. In the limit, if $\lambda = 1$ products are perfect substitutes, and if $\lambda = 0$ the utility function becomes Cobb-Douglas, meaning that a fixed proportion of income is devoted to each product.[5] The consumer has total income normalized to 1, so that if $p\left(x\right)$ is the price of good x, the budget constraint can be written as

$$\int_0^1 p\left(x\right) y\left(x\right) dx \leq 1.$$

The utility maximization problem is solved by equating the marginal rate of substitution between any two goods x and x' to the ratio of their prices,

$$\left(\frac{y\left(x\right)}{y\left(x'\right)}\right)^{\lambda-1} = \frac{p\left(x\right)}{p\left(x'\right)}. \tag{1}$$

Substituting quantities into the budget constraint and defining

$$P \equiv \left(\int_0^1 p(x)^{-\frac{\lambda}{1-\lambda}} dx\right)^{-\frac{1-\lambda}{\lambda}}$$

---

[5]The model can accomodate a more general specification that includes an outside good, with similar results but a substantially higher level of complexity.

as the *price index*, the inverse demand function for good x can be obtained as

$$p(x) = \kappa y(x)^{\lambda - 1}, \tag{2}$$

where $\kappa = P^{\lambda}$. As expected, goods are substitutes, and increases in the aggregate price index lead to increases in the consumption of good x.

The patentholder owns a patent for an invention that can be used to reduce the marginal cost of any good x. We assume that the patentholder does not compete in the downstream market,[6] and the production of each good x is carried out by a different monopolist. If this firm produces the downstream good with the initial technology, the marginal cost is constant, independent of x and denoted by c. The patented process allows firms to reduce their marginal cost. However, the magnitude of the reduction is heterogeneous among firms. In particular, the parameter x orders the firms by how much they benefit from the new invention. We assume that x is private information, so that when negotiating a license the patentholder does not observe the valuation of each prospective licensee.[7] A firm producing good x will obtain a reduction in marginal cost of $(1-x)\theta$ so that the new cost is given by

$$c(x) = \begin{cases} c - (1-x)\theta & \text{if a license is purchased, and} \\ c & \text{otherwise,} \end{cases}$$

where θ represents the quality of the invention created by the patentholder. Obviously, it has to be that $\theta \le c$ so that firms operate with non-negative marginal costs. Notice that the firms that benefit the most from the patent are those with the lowest x, while firms with x close to 1 obtain

---

[6]The case where the patentholder is a competitor in the downstream market is addressed by Arora and Fosfuri (1998), Wang (1998) and Yang and Wang (1999) among others. They provide conditions under which the patentholder sells licenses to some of the competitors in the same market.

[7]The definition of x is with a slight abuse of notation as it is used both for the good and the attachment to the invention. However, it is obvious that from the point of view of consumers all goods are symmetric and the decision to consume each of them is independent of whether x is observable or not. Hence, the ordering of x is without loss of generality.

Moreover, because firms are atomistic, the pricing decision only requires firms to know their own type and the distribution of x among the rest of the firms, which coincides with the original distribution. Therefore, whether each licensee can observe the type of the rest of the competitors or not is irrelevant for the results of the paper.

$y(x)$ chosen

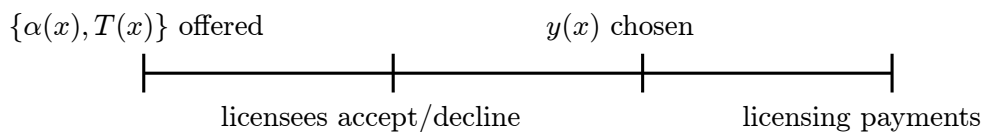licensees accept/decline

licensing payments

Figure 1: Timing of the model

only a small decrease in cost. This heterogeneity in the cost reduction created by the license can be interpreted as a different complementarity between the new patent and the pre-existing technology that each firm possesses. For this reason, the cost reduction is assumed to be independent of the existing degree of differentiation in the market.[8]

The demand for the product of a downstream monopolist is given by equation (2). This demand has the usual properties for $0 < \lambda < 1$. Firms are atomistic and hence, they cannot affect the aggregate price index that determines the value of κ.

In order to license the innovation, the patentholder demands a payment corresponding to a share α of gross revenues of the licensee together with a flat fee T. Moreover, the patentholder has all the bargaining power. This seems a reasonable assumption in this context, since firms are atomistic and compete with each other.

The timing of the game is as follows. First, the patentholder offers of a contract (or a menu of contracts) to all prospective licensees. These firms simultaneously accept or reject the contract. In a third stage, all firms observe the share of firms that purchased the license and choose simultaneously their production. Finally firms collect profits and settle licensing obligations.

The profit function of the firm producing good x under a licensing contract (α, T) corresponds

---

[8]This effect could be relevant if product heterogeneity were assumed to be chosen by firms, affecting for example the marginal cost of each good. Because we consider only the technical implications of the license we abstract from these issues.

to

$$\pi(x) \equiv \max_{y(x)} y(x)((1-\alpha)p(x) - c(x)) - T,$$

where $p(x) = \kappa y(x)^{\lambda-1}$. In other words, the firm maximizes profits $y(x)(p(x) - c(x))$ net of

licensing payments, $\alpha p(x)y(x) + T$. The first order condition of this problem results in a quantity

that corresponds to

$$y(x) = \left(\frac{c(x)}{(1-\alpha)\lambda\kappa}\right)^{\frac{1}{\lambda-1}}, \tag{3}$$

and thus the price of good x becomes

$$p(x) = \frac{c(x)}{(1-\alpha)\lambda}. \tag{4}$$

Notice that when $\alpha = 0$ and $\lambda = 1$, so that products are perfect substitutes, the price equals

marginal cost. The more imperfect substitutes are the products, the higher is the mark-up that

firms will optimally charge. This expression also illustrates the *double-marginalization* that the use

of a variable payment α adds to the typical mark-up over marginal cost. A higher α reduces the

marginal revenue of the producer, cutting down on quantity and raising the price.

Replacing $y(x)$ and $p(x)$ in the profit function of the firm producing good x we obtain

$$\pi(x) = \left(\frac{1-\lambda}{\lambda}\right) c(x)^{-\frac{\lambda}{1-\lambda}} ((1-\alpha)\lambda\kappa)^{\frac{1}{1-\lambda}} - T.$$

It is easy to verify that the profit function, $\pi(x)$, has the usual properties, and in particular, profits

are decreasing in c and α. Moreover, for a given contract (α, T) and because $c(x)$ is increasing in

x, profits are smaller for firms that benefit less from the invention.

3   The Optimal Contract

For illustration purposes in the previous section we have restricted the contracts to a unique two-

part tariff corresponding to a share on revenues and a fee. However, because prospective licensees

are heterogeneous in their valuation for the invention, the patentholder might be interested in offering not only one type of contract but a variety of contracts. In this section we describe the set of contracts that are feasible given the private information concerning this valuation and characterize the optimal menu. We later show that the results would be unaltered if the patentholder had access to a broader set of mechanisms such as a per-unit royalty.

Using the Revelation Principle, we can focus without loss of generality on menus of contracts $\{\alpha(x), T(x)\}$, where each option of the menu is intended to a particular firm type x. That is, denote as $\pi(x, \widehat{x})$ the profits of the firm with product x when it chooses the contract intended for type $\widehat{x}$. From the expressions in the previous section,

$$\pi(x, \widehat{x}) = \left(\frac{1-\lambda}{\lambda}\right)(c - (1-x)\theta)^{-\frac{\lambda}{1-\lambda}}((1-\alpha(\widehat{x}))\lambda\kappa)^{\frac{1}{1-\lambda}} - T(\widehat{x}).$$

As usual, for a menu of contracts to be incentive compatible —or to induce self-selection— declaring its own type has to be a dominant strategy for each firm. If we define the profits of firm x when choosing the type $\widehat{x}$ that maximizes profits π as

$$\Pi(x) = \max_{\widehat{x}} \ \pi(x, \widehat{x}),$$

a menu of contracts $\{\alpha(x), T(x)\}$ is incentive-compatible if for all firms x,

$$\Pi(x) = \pi(x, x). \tag{5}$$

Moreover, firms that do not purchase a license charge a price $p(x) = \frac{c}{\lambda}$ and obtain profits

$$\Pi_0 = (1-\lambda)\left(\frac{c}{\lambda}\right)^{-\frac{\lambda}{1-\lambda}}\kappa^{\frac{1}{1-\lambda}},$$

which are independent of x. Therefore, a firm will purchase a license if $\Pi(x) \geq \Pi_0$.

The next lemma characterizes the set of contracts that allow separation of the different licensees.

11

Lemma 1 *For a menu of contracts $\{\alpha(x), T(x)\}_{x\in[0,1]}$ to be incentive-compatible, the following must be true for all $x \leq \widetilde{x}$:*

1. *$\alpha(x)$ is non-decreasing in x,*

2. *$T(x)$ corresponds to*

$$T(x) = \left(\frac{1-\lambda}{\lambda}\right)(c - (1-x)\theta)^{-\frac{\lambda}{1-\lambda}}\left((1-\alpha(x))\lambda\kappa\right)^{\frac{1}{1-\lambda}} - \int_x^{\widetilde{x}} \theta y(s)\,ds - \Pi_0 - K, \quad (6)$$

where $K(1-\widetilde{x}) = 0$ and $K \geq 0$.

The previous lemma describes the incentive compatible contracts as a function of $\widetilde{x}$. Only firms with $x \leq \widetilde{x}$ will participate in the mechanism, obtaining profits

$$\Pi(x) = \begin{cases} \Pi_0 + K + \int_x^1 \theta y(s)\,ds & \text{if } \widetilde{x} = 1, \\ \Pi_0 + \int_x^{\widetilde{x}} \theta y(s)\,ds & \text{otherwise.} \end{cases}$$

In particular, if $\widetilde{x} = 1$, all firms decide to obtain a license, and any contract that provides profits of at least Π_0 will be enough to satisfy this constraint. The possible existence of this corner solution is the reason why the parameter K is introduced in the expression (6). If $\widetilde{x} = 1$ then $K = \Pi(1) - \Pi_0 \geq 0$. It is easy to see that $K > 0$ will never be profit maximizing, since the patentholder would be able to increase profits by raising the fee T by the amount K to all licensees. Therefore, in the remaining of the paper we will restrict ourselves to the case where $K = 0$ regardless of whether $\widetilde{x} = 1$ or not, and hence, the relevant participation constraint is equivalent to[9]

$$\Pi(\widetilde{x}) = \Pi_0. \quad (7)$$

---

[9]The solution to this problem might not be *a priori* unique. The reason is that, as we will later emphasize, κ is a function of $\widetilde{x}$ and hence, the previous expression can in principle hold for more than one value of $\widetilde{x}$. In particular, if more licenses implied in equilibrium lower prices and lower willingness to pay for the license; for example, the two following equilibria could arise: one with very few licenses and little competition, in which only the firms with the highest cost reduction would license the patent, and another with many licensees, where market prices would be lower. We will turn to this issue in the next section and show the uniqueness of the solution.

Lemma 1 also shows that a necessary condition for the schedule $\boldsymbol{\alpha} = \{\alpha(x)\}_{x=0}^{\widetilde{x}}$ to be incentive-compatible is that it must be weakly increasing in x. The intuition is that firms with lower x benefit more from the patent, reducing more their marginal cost of production. For this reason, these firms will produce a higher output and raise larger revenues, and as a result, they will be more likely to trade a lower payment on revenues in exchange for a higher flat fee. This result is a parallel to the classic two-part tariff schedules studied in monopoly theory, where a monopolist selling to heterogeneous consumers decides to charge a higher flat fee and a lower per unit price to consumers that are interested in buying more units of the good.

3.1 The Problem of the Patentholder

The objective function of the patentholder can be written as

$$V = \max_{\{\alpha(x)\}_{x=0}^{\widetilde{x}}, \widetilde{x}} \int_0^{\widetilde{x}} (\alpha(x) p(x) y(x) + T(x))\, dx,$$

$$\text{s.t.} \qquad \begin{array}{c} \alpha(x) \geq \alpha(x') \ \text{if} \ x \geq x', \\ (6) \text{ and } K = 0. \end{array}$$

That is, from all firms that buy a license — those with $x \leq \widetilde{x}$ — the patentholder obtains a share $\alpha(x)$ of total sales $p(x) y(x)$ plus the associated flat fee. The constraints imposed in this problem guarantee that the contract satisfies *Lemma 1* and it is therefore incentive-compatible.

We can replace in the previous objective function $p(x) = \frac{c-(1-x)\theta}{(1-\alpha)\lambda}$ and T from (6). Integrating the last term by parts, the problem can be rewritten as

$$V = \max_{\{\alpha(x)\}_{x=0}^{\widetilde{x}}, \widetilde{x}} \int_0^{\widetilde{x}} \left\{ y(x) \left((c - (1-x)\theta) \left(\frac{1}{(1-\alpha(x))\lambda} - 1 \right) - \theta x \right) - \Pi_0 \right\} dx,$$

$$\text{s.t.} \ \alpha(x) \geq \alpha(x') \ \text{if} \ x \geq x'. \tag{8}$$

The patentholder chooses the optimal royalty profile and the number of licenses sold, denoted by $\boldsymbol{\alpha}^* = \{\alpha^*(x)\}_{x=0}^{\widetilde{x}^*}$, and $\widetilde{x}^*$. These two choices are enough to characterize the profile of optimal fees $\{T^*(x)\}_{x=0}^{\widetilde{x}^*}$. Although the previous expression is similar to the typical mechanism design

problem (see for example Maskin and Riley (1984)), one of its features does not allow us to solve it using the standard technique of maximizing for each of the types separately and later verifying that the optimal contract satisfies the incentive-compatibility constraint. Instead, competition among licensees means that in our model, each particular $\alpha(x)$, affects the price that all firms charge in equilibrium through κ. From the point of view of a particular firm, this effect is negligible since firms are atomistic. However, the monopolist is aware that when considering a subinterval of $[0, \tilde{x}]$, monotonic increases in the schedule $\boldsymbol{\alpha}$ raise the price index. Because a higher price index means a higher valuation for the license, the patentholder takes this effect into account in the optimization problem.

3.2 The Optimal Menu of Contracts

The next proposition shows that in fact the optimal mechanism implies that firms that make a better use of the patent will be offered a contract that allows them to keep a strictly larger share of revenues, but in exchange they pay a higher flat fee.

Proposition 2 *For a given marginal firm, $\tilde{x}$, the optimal schedule $\boldsymbol{\alpha}^* = \{\alpha^*(x)\}_{x=0}^{\tilde{x}}$ is implicitly defined by*

$$\alpha^*(x) = [1 - s(\boldsymbol{\alpha}^*, \lambda)]\gamma(x) + s(\boldsymbol{\alpha}^*, \lambda), \tag{9}$$

where $\gamma(x) \equiv \frac{\theta x}{c - (1 - 2x)\theta}$ and

1. *$s(\boldsymbol{\alpha}^*, \lambda)$ is common to all x's,*

2. *$0 \leq s(\boldsymbol{\alpha}^*, \lambda) \leq 1$,*

3. *$s(\boldsymbol{\alpha}^*, \lambda) \leq V$ with equality when evaluated at the optimal $\tilde{x}$,*

4. *$\alpha^*(x)$ is increasing in x.*

An important property of this solution needs to be emphasized. The definition of $\alpha^*(x)$ implicitly involves the entire profile of contracts offered to other firms. That is, the solution to the patent-holder's optimization problem is a fixed-point in a functional equation problem. However in the appendix, by transforming the fixed point in the profile $\boldsymbol{\alpha}^*$ into a fixed point in the parameter s, we show that a solution to this fixed-point problem exists and it is unique.

The optimal profile $\boldsymbol{\alpha}^*$ that a patentholder will offer to a set of firms is increasing in x, satisfying the condition for incentive compatibility. Moreover the previous proposition can be interpreted as a convex combination of two effects that we denote *price discrimination* and *competition softening*.

To interpret the first effect, in the next proposition we study a similar problem where the patentholder is facing a single potential licensee, who is a monopolist in the final good market, with a cost function $c(x) = c - (1-x)\theta$, where x is private information. The proposition shows that the contract in that case will include a share on revenues of $\gamma(x)$.

Proposition 3 *Suppose that a patentholder is willing to license a patent to a single firm with cost function $c(x) = c - (1-x)\theta$ when using the innovation. The downstream firm is of type x, which is uniformly distributed between 0 and 1, and is private information. This firm is a monopolist facing an inverse demand function $p = \kappa y^{\lambda-1}$ for a fixed κ. Then, the optimal contract corresponds to*

$$\widetilde{\alpha}(x) = \frac{\theta x}{c - (1-2x)\theta} = \gamma(x),$$

$$\widetilde{T}(x) = \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{c - (1-x)\theta}{(c - (1-2x)\theta)^{\frac{1}{\lambda-1}}} - c^{-\frac{\lambda}{1-\lambda}}\right)(\lambda\kappa)^{\frac{1}{1-\lambda}} + \int_x^{\widetilde{x}} \theta y(s)\,ds.$$

Hence, in the model with multiple downstream firms, the first effect, with weight $1 - s(\boldsymbol{\alpha}^*, \lambda)$, can be interpreted as the optimal share on revenues that the patentholder would retain if the effect of α on the aggregate price index P were not taken into account. Notice that $\gamma(x)$ is independent of

the demand elasticity $\frac{1}{1-\lambda}$. In other words, $\gamma(x)$ would also correspond to the optimal contract to be offered to a set of monopolists in independent markets, for any positive value of κ. In this case, a higher share α allows to extract more profits from the licensee. The limit to this effect is what we have previously denominated, borrowing the term from the vertical relationships literature, as *double-marginalization.* That is, a higher share implies a larger distortion in the decision of the licensee, reducing profits gross of licensing payments, and therefore the willingness to pay for a license. The optimal share $\gamma(x)$ takes into account the trade-off between rent-extraction from the heterogeneous licensees – achieved through the flat fee $T(x)$, which does depend on the demand elasticity – and the problem of double-marginalization that it generates. The relevance of the rent extraction component increases as x raises – or in other words $\gamma'(x) > 0$ – since the firm produces a lower quantity and the distortion due to double-marginalization becomes less important.

The second effect, with weight $s(\boldsymbol{\alpha}^*, \lambda)$, corresponds to the interest of the patentholder in softening competition among the licensees. The profit that each licensee obtains increases with the price that all the other firms charge through the aggregate price index P. Therefore, the optimal policy fully internalizes this effect. As an illustration, if for example all firms were identical, and hence no rents of private information ought to be guaranteed, the optimal share on profits would correspond to $\alpha(x) = 1$ to all firms, increasing $p(x)$, so that the patentholder would appropriate all the consumer income and at the same time minimize the units sold $y(x)$ and the cost of those sales. This intuition carries over to the general case, where the monopolist, by increasing the profile $\alpha(x)$ to all licensees, sets the price index P sufficiently high to guarantee some market power to each licensee. The patentholder will then partially appropriate the surplus through the fee $T(x)$.

The function $s(\boldsymbol{\alpha}^*, \lambda)$ that determines the weights of these two effects is a complicated expression of the parameters of the model and the entire royalty profile. Given our normalization of the consumer's income, $s(\boldsymbol{\alpha}^*, \lambda)$ can range between 0 and 1. This function counterbalances the

double-marginalization cost with the two effects: competition softening and price discrimination. Notice that $s(\boldsymbol{\alpha}^*, \lambda)$ can also be interpreted as the optimal royalty for the firm with $x = 0$. That is, $\alpha^*(0) = s(\boldsymbol{\alpha}^*, \lambda)$. Numerical results show that $s(\boldsymbol{\alpha}^*, \lambda)$ is increasing in λ, and the next lemma characterizes the limiting behavior.

Lemma 4 $lim_{\lambda \to 0} s(\boldsymbol{\alpha}^*, \lambda) = 0$ and $lim_{\lambda \to 1} s(\boldsymbol{\alpha}^*, \lambda) \le \frac{\theta}{c}$.

When goods are not very sustitutable (λ is low), the price discrimination effect dominates, and the patentholder sets a share on profits as if firms were not competing with each other, along the lines of *Proposition 3*. As goods become more homogeneous (λ is high), the firms' price approaches marginal cost c and as a result the payoff from inducing higher markups increases with respect to the cost of the distortion due to double-marginalization.

Another surprising interpretation of $s(\boldsymbol{\alpha}^*, \lambda)$ in *Proposition 2* is that it corresponds to the profits that the patentholder achieves when choosing the optimal profile of α and T. That is, evaluated at the optimal $\widetilde{x}$, $s(\boldsymbol{\alpha}^*, \lambda) = V$. If consumer income, denoted as W, were not normalized to 1, then $s(\boldsymbol{\alpha}^*, \lambda) = \frac{V}{W}$. Moreover, because $s(\boldsymbol{\alpha}^*, \lambda)$ tends to 0 when λ goes to 0, profits for the patentholder approach 0 when final goods are not substitutes of each other. The reason is that the invention does not lead to a competitive advantage nor an increase in the production of any firm. Therefore, firms are in the limit willing to pay 0 to obtain it. Profits increase with λ because of the competitive advantage that it generates to licensees and because the willingness to pay increases accordingly.

3.3 Per-Unit Royalties

So far, we have restricted the analysis to a general family of two-part tariff contracts that a patentholder can offer to a set of potential licensees. This menu of contracts has been interpreted as a

fixed fee and a share on the value of sales. Such a contract, although observed in practice, does not correspond to the usual per-unit royalty studied in the literature. In this section, we will show that under the assumptions of this paper, both mechanisms implement the same optimal allocation.

Consider a per unit royalty r and a fee F as an alternative licensing scheme. In this case, profits of a firm x correspond to

$$\pi(x) = \max_{y(x)} \, y(x) \left(\kappa y(x)^{\lambda-1} - c(x) - r \right) - F \tag{10}$$

where we have already replaced the price with $p(x) = \kappa y(x)^{\lambda-1}$. The quantity chosen will be $y(x) = \left(\frac{c(x)+r}{\kappa\lambda} \right)^{\frac{1}{\lambda-1}}$, which is decreasing in r. In a similar way as in the previous section, the patentholder can set up a menu of contracts $\{r(x), F(x)\}$ in order to separate among the different values of x. The next lemma, parallel to *Lemma 1*, states the conditions that this menu must satisfy in order to be incentive compatible.

Lemma 5 *For a menu of contracts $\{r(x), F(x)\}_{x \in [0,1]}$ to be incentive-compatible, the following must be true for all $x \leq \tilde{x}$:*

1. *$r(x)$ is non-decreasing in x,*

2. *profits of licensee x correspond to*

$$\Pi(x) = \begin{cases} \Pi_0 + K + \int_x^1 \theta y(s)\, ds & \text{if } \tilde{x} = 1, \\ \Pi_0 + \int_x^{\tilde{x}} \theta y(s)\, ds & \text{otherwise.} \end{cases} \tag{11}$$

where $K \geq 0$, $\tilde{x} \leq 1$ and $K(1 - \tilde{x}) = 0$.

Therefore, for a contract to be incentive compatible, firms that obtain a bigger reduction in costs should pay a smaller per-unit royalty, exactly in the same way as with a proportional payment on the value of sales.[10] Moreover, notice that if firms produce the same quantity as in the two-part

---

[10]This result is consistent with the evidence obtained in Taylor and Silberston (1973).

tariff studied earlier, for a given value of $\tilde{x}$, profits for the licensee decrease with x at an identical rate. This observation suggests that if the menu $\{r(x), F(x)\}_{x \in [0, \tilde{x}]}$ can induce firms to produce the same quantities as in the case with $\{\alpha(x), T(x)\}_{x \in [0, \tilde{x}]}$, both mechanisms should be equivalent. Indeed, this intuition is correct, as stated in the next result.

Proposition 6 *Any allocation implemented using an incentive-compatible menu of two-part tariffs* $\{\alpha(x), T(x)\}_{x \in [0, \tilde{x}]}$ *can be implemented with a menu of per-unit royalties and fees* $\{r(x), F(x)\}_{x \in [0, \tilde{x}]}$, *where*

$$r(x) = c(x) \frac{\alpha(x)}{1 - \alpha(x)} \tag{12}$$

is strictly increasing in x.

It can also be shown that no other menu of per-unit royalties and fees can improve upon the one corresponding to $r^*(x) = c(x) \frac{\alpha^*(x)}{1 - \alpha^*(x)}$ making the previous results general. A remarkable difference, however, is that $F(x)$ is larger than $T(x)$. The reason for this difference in the fees is that, despite raising identical profits, under each of the mechanisms the per-unit margin is different. In particular, notice that under the previous specification of the per-unit royalty,

$$p(x) - c(x) - r(x) = p(x) - \frac{c(x)}{1 - \alpha(x)} \leq (1 - \alpha(x))p(x) - c(x) \tag{13}$$

and therefore $F(x) \geq T(x)$, with strict inequality for $\alpha(x) > 0$. Replacing in (9) the royalty can be rewritten as

$$r(x) = x\theta + c(x) \frac{s(\boldsymbol{\alpha}^*, \lambda)}{1 - s(\boldsymbol{\alpha}^*, \lambda)}. \tag{14}$$

Two final comments are in order. First, the equivalence stated in the previous proposition between both mechanisms does not carry through to the case of a unique (α, T) and (r, T). It is easy to observe from the proposition that under a constant α, a per-unit royalty that implements

the same quantity for each value of x ought to be increasing in x. This observation suggests that, restricted to a uniform contract, the screening possibilities of each scheme will be different.

Second, the previous results suggest that there is a family of mechanisms that achieves the same allocation. Although it is not the goal of this paper to characterize this family, it is worth mentioning one mechanism that is not in this family: a contract with a fee and a share of the net revenue of the licensee. It is easy to observe that in this case, the quantity $y(x)$ does not depend on the contract and therefore, the patentholder will not be able the affect the pricing decision of the licensees.

3.4 The Optimal Proportion of Licensees

The results in the previous section point to an intuition that has been rarely mentioned in the literature: Licensing proceeds should be higher in markets where firms produce a more homogeneous good because an innovation that reduces costs allows licensed firms to steal a larger market share from non-licensed competitors. To this analysis we add two important features. The first is that these larger proceeds are obtained by setting a higher variable payment and the second is that, as we will argue next, the proportion of licensees is in general higher when the good is more homogeneous.

In order to characterize the optimal profile $\boldsymbol{\alpha}^*$ set by the patentholder we have kept fixed the threshold on the proportion of firms that decide to obtain a license, $\widetilde{x}$. Notice, however, that the value of the weight that the contract assigns to the competition softening effect, $s\left(\boldsymbol{\alpha}^*, \lambda\right)$, depends on $\widetilde{x}$. From the maximization problem in (8), we can obtain the optimal threshold value $\widetilde{x}^*$, with the corresponding first-order condition

$$y\left(\widetilde{x}\right)\left[\left(c-\left(1-\widetilde{x}^*\right)\theta\right)\left(\frac{1}{\left(1-\alpha\left(\widetilde{x}^*\right)\right)\lambda}-1\right)-\theta\widetilde{x}^*\right]-\Pi_0+\frac{\partial\kappa}{\partial\widetilde{x}}\frac{s}{\kappa(1-\lambda)}\geq 0. \quad (=0 \text{ if } \widetilde{x}^*<1) \ (15)$$

This condition defines the optimal proportion of licenses offered. The first term evaluates the

marginal increase in licensing revenue that selling to the marginal firm $\widetilde{x}$ generates, while the second is the minimum amount of profits that the licensee can guarantee for itself by not buying the license. These two terms are standard in the mechanism design literature. More licenses should be offered as long as the additional payments that new licensees provide are higher than the increase in rents that infra-marginal licensees must obtain to induce self-selection.

The last term, which is not standard in the mechanism design literature, can be interpreted as the effect of selling one more license on the equilibrium prices that firms charge, and consequently on κ, which reflects the effect on the price index. This effect can be computed as,

$$\frac{\partial \kappa}{\partial \widetilde{x}} = - (1 - \lambda) \, \kappa^{\frac{2-\lambda}{1-\lambda}} \left(p \left(\widetilde{x}^* \right)^{-\frac{\lambda}{1-\lambda}} - \left(\frac{c}{\lambda} \right)^{-\frac{\lambda}{1-\lambda}} \right) \lessgtr 0$$

and so, the sign of $\frac{\partial \kappa}{\partial \widetilde{x}}$ is in principle inconclusive. The existing literature advocates that firms are willing to pay more for a license when fewer licenses are allocated, since in those models equilibrium prices turn out to be higher. Allocating more licenses reduce the return from obtaining them, which results in a downward sloping demand for licenses. In our environment, this conduct would be consistent with $\frac{\partial \kappa}{\partial \widetilde{x}} < 0$, and it would make the patentholder decrease the proportion of licenses compared to the case where κ is a fixed parameter along the lines of *Proposition 3*. In other words, the existing literature is consistent with $p \left(\widetilde{x} \right) < \frac{c}{\lambda}$ (that is, if obtaining a license implies a decrease in the price that the firm will charge). In this case, competition among downstream firms is increased.

The previous expression does not rule out multiplicity of values for which the first order condition is guaranteed. However, the next lemma shows that in our case the problem is well-behaved and the optimal menu of contracts is unique.

Lemma 7 *There is at most one value of $\widetilde{x}$ that satisfies the first order condition. This first order condition is also sufficient.*

When the first order condition is not satisfied for any interior value of $\widetilde{x}$ the patentholder decides to license the invention to all the competitors in the market. Numerical results suggest that such a strategy might be optimal only if λ is large.

We finally need to point out an aspect that we are not considering in this paper. Antitrust laws preclude licensors from charging a per unit royalty that more than compensates the decrease in cost that the invention generates on the licensee. Such a constraint would correspond to $r(x) \leq (1-x)\theta$ or using *Proposition 6*, to $p(x) < \frac{c}{\lambda}$. Obviously, in our model higher prices are possible as long as fees are negative. Moreover, notice that the expression for $p(x)$ depends crucially on the variable payment $\alpha(x)$ that the patentholder demands for the license:

$$p(x) = \frac{c - (1-x)\theta}{(1 - \alpha(x))\lambda}.$$

As it is obvious from this expression, the price must be increasing in x. In particular, if the patentholder sells enough licenses and x approaches 1, the price will be above the original price and therefore $\frac{\partial \kappa}{\partial \widetilde{x}} > 0$. While the constraint will not bind for low values of λ, for very homogeneous goods the patentholder might be interested in selling to all firms as a way to artificially raise the price that firms charge. Imposing the condition in our environment will keep most of the features of the model unchanged. The optimal contract will still involve a menu of contracts to different licensors, with the difference that values of x above a certain threshold might be bunched into the same contract.

4   Flat Fees

In the previous sections we have characterized the optimal share on profits (or the optimal per-unit royalty) that the patentholder will demand. Another way to provide intuition about the bearings of the model is by comparing this optimal contract with an alternative mechanism that relies only

on flat fees. This contract is also interesting on its own because of its prevalence in practice and because it has been widely studied in the literature. In order to ease the exposition we refer the reader to the appendix for a more detailed analysis.

In the absence of other contractual arrangements, this contract specifies a fee T that firms have to pay in order to obtain a license. It is important to notice that such a contract prevents the possibility of separating different users of the invention and of softening competition among licensees.

For a given contract, firms with a lower x obtain higher profits; therefore, if a firm located at x decides to purchase the license, all others with $x' < x$ will purchase it as well. However, not all firms will buy a license, since the firm located at $x = 1$ obtains no improvement in cost by doing so. As before, a firm will be interested in purchasing a license if profits are larger with the license, as opposed to the case where the firm uses the initial technology. In other words, a firm will buy the license if $\overline{\Pi}(x) \geq \overline{\Pi}_0$, where $\overline{\Pi}(x)$ and $\overline{\Pi}_0$ represent, respectively, the profits of licensees and non-licensees. This choice determines a threshold $\overline{x}$ such that all firms with an $x > \overline{x}$ decide to use the initial technology; hence $\overline{x}$ also represents the proportion of firms that buy the license. This condition derives a downward slopping demand for licenses. In the next section we compare both contracts.

5   Comparing Contracts

We use numerical methods to establish the characteristics of both types of contracts, their effect on utility and private profits. The numerical characterizations that follow correspond to the parametrization $c = 1$ and $\theta = \frac{4}{5}$; that is, the firm at $x = 0$, obtains a reduction of 80 percent in marginal cost. Since the marginal cost, c, and the degree of attachment to the new technology
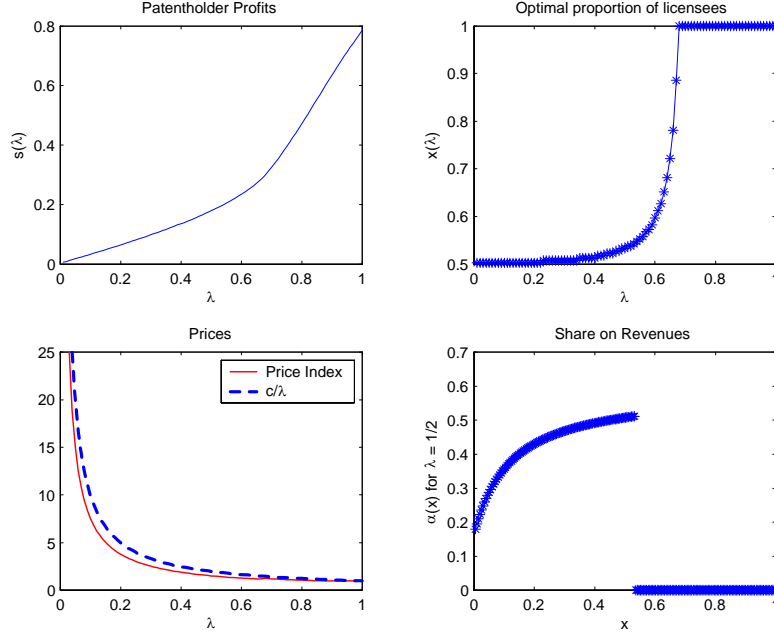
Figure 2: Royalty Rates. Parametrization: $c = 1, \theta = \frac{4}{5}$

by the firms, θ, are the only non-free parameters, the numerical exercise with our parametrization is very general. Because the objective is to analyze the environment as the degree of product differentiation varies, we allow λ to change.

5.1 Two-Part Tariffs

Figure 2 illustrates the features of the optimal two-part tariff. In particular, it shows that the equilibrium variable $s(\boldsymbol{\alpha}^*, \lambda)$ is increasing with the degree of product differentiation (as λ increases products become more homogeneous). The number of licenses, $\tilde{x}$, in the upper right panel increases with λ. The share on revenues retained by the patentholder, $\alpha(x)$, given $\lambda = \frac{1}{2}$, is increasing in x. Although not shown, the entire profile $\boldsymbol{\alpha}^*$ is weakly increasing in λ since $\frac{\partial \alpha^*(x)}{\partial \lambda} = (1 - \gamma(x)) \frac{\partial s}{\partial \lambda}$.

5.2 Licensed Firms and Prices

The share of firms that purchase a license is consistently higher with two-part tariffs than with only flat fees. *Figure 3* illustrates this property. As a result, more firms gain access to the new technology when two-part tariffs are used. This effect is due to two reasons. First, as it is common in a monopolistic setup, having the possibility of price-discriminating entices the patentholder to increase the quantity of licenses offered. Second, by adding an additional licensee in the case of two part-tariffs, the price is modified from $p(x) = \frac{c}{\lambda}$ to $p(x) = \frac{c-(1-2x)\theta}{\lambda(1-s)}$, while in the case of a flat fee the price becomes $p(x) = \frac{c-(1-x)\theta}{\lambda}$, which is strictly smaller. Hence, one more licensee lowers the price by more in the case of a flat fee, reducing even more the quantity sold.

Furthermore, as λ increases (the degree product differentiation declines), the number of licenses granted with two-part tariffs increases, while the number of licenses with flat fees declines. This is so because more competition lowers the profits of the producers when more firms have the technology. Eventually all firms gain access to the new technology with two-part tariffs; in contrast, the number declines toward zero when flat fees are used.

In both cases, the price index declines as λ increases, but in the case of two-part tariffs since $s(\alpha^*, \lambda)$ is increasing in λ, the wedge increases. *Figure 4* illustrates this behavior.

5.3 Welfare

We can also estimate the effect on consumer utility and firms' profits of each contract. The indirect utility of the consumer given prices $\{p(x)\}_{x=0}^1$ is defined in terms of the schedule of prices as

$$v\left(\{p(x)\}_{x=0}^1\right) = \left(\int_0^1 y(x)^\lambda \, dx\right)^{\frac{1}{\lambda}},$$

where $y(x)$ are the optimal consumption bundles. In this case, the indirect utility function is homogeneous of degree 1 in income, which implies that for any level of income W, the corresponding
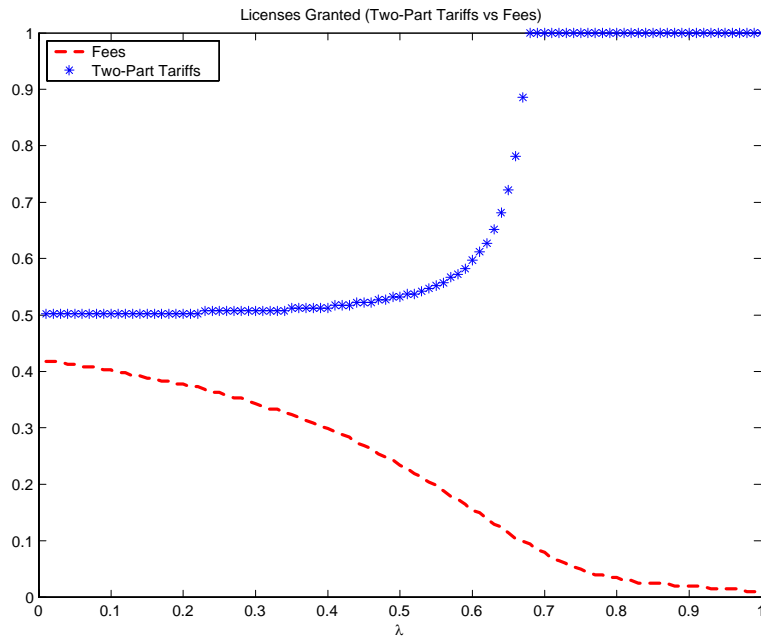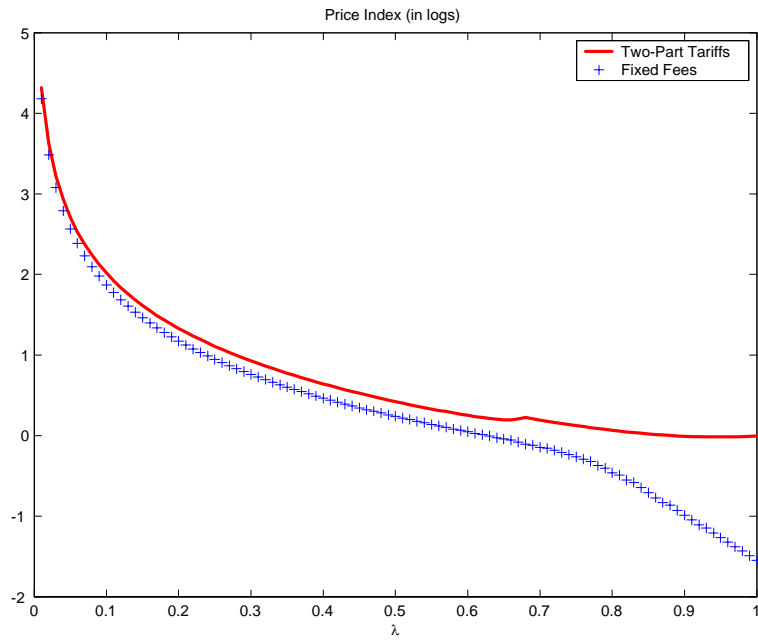
25

Figure 3: Proportion of Licenses

Figure 4: Price Indexes

indirect utility function is given by

$$v\left(\{p(x)\}_{x=0}^1, W\right) = W v\left(\{p(x)\}_{x=0}^1\right).$$

We can then perform the following experiment: starting from a situation where the patentholder is constrained to using flat fees we compute the increase in income that would make the consumer indifferent with a situation where the patentholder uses two-part tariffs. The negative of this change in income corresponds to the usual compensating variation:

$$CV = 1 - \frac{v\left(\{p^{FF}(x)\}_{x=0}^1\right)}{v\left(\{p^{2PT}(x)\}_{x=0}^1\right)},$$

where $p^{FF}(x)$ and $p^{2PT}(x)$ denote the prevailing prices in the cases of flat fees and two-part tariffs, respectively.

Given optimal consumption bundles in the cases of flat fees and two-part tariffs, $y^{FF}(x)$ and $y^{2PT}(x)$, respectively, we can define producer surplus as the sum of the profits of the final good producers and the patentholder. In the case of two-part tariffs, profits are, respectively,

$$FS = \int_0^{\widetilde{x}} \left[((1 - \alpha(x)) p^{2PT}(x) - c(x)) y^{2PT}(x) - T(x)\right] + \int_{\widetilde{x}}^1 y^{2PT}(x) \left[p^{2PT}(x) - c(x)\right] dx,$$

$$PatS = \int_0^{\widetilde{x}} \left[\alpha(x) p^{2PT}(x) y^{2PT}(x) + T(x)\right] dx.$$

In the case of flat fees, the corresponding expressions are

$$FS = \int_0^1 \left[(p^{FF}(x) - c(x)) y^{FF}(x)\right] dx - T\overline{x},$$

$$PatS = T\overline{x}.$$

The results of our simulations in *Figure 5* show the following: As expected, from the producers standpoint, a two-part tariff is always preferred because, by inducing a higher price for the final good (and therefore a higher price index) and by making the technology available to more firms, competition among final good producers is reduced. Because of the increase in prices, the distortion
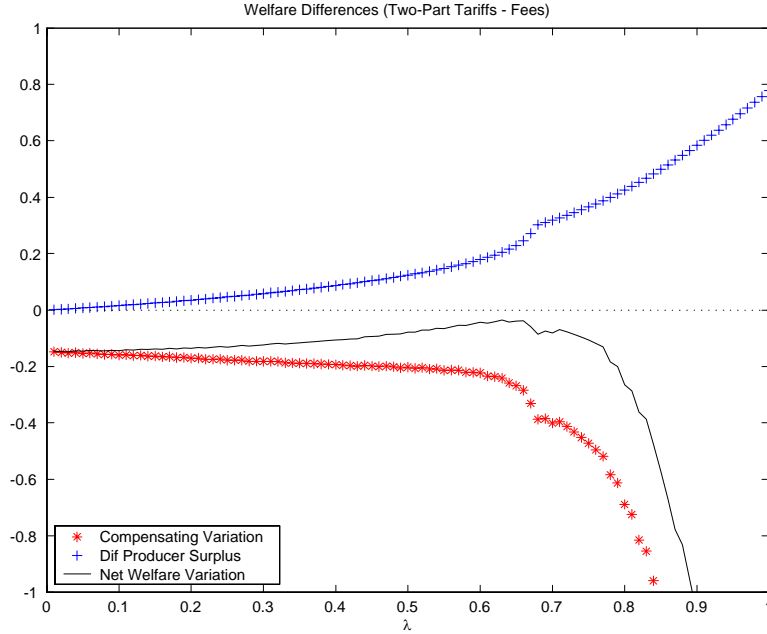
Figure 5: Welfare Comparison

on the consumer side is larger, leading to a lower consumer utility, or a negative compensating variation.

We can define a variation in net welfare as the sum of the compensating variation and the difference in producer surplus. We can see that from *Figure 5* that, since consumers are always worse-off under two-part tariffs, even though more firms have access to the new technology, the variation in net welfare is negative.[11]

6 Concluding Remarks

This paper presents a model of patent licensing that integrates two important features largely unexplored in the literature: private information and a varying degree of product differentiation in the downstream market. Both of these features turn out to have a systematic relationship with the

---

[11]We also verified numerically that, compared with the case in which all firms operate with the inital technology, licensing by means of two-part tariffs increases both consumer utility and firms' profits, and therefore social welfare improves.

optimal contract that the patentholder chooses. Heterogeneity among prospective licensees gives the patentholder the opportunity to offer a menu of two-part tariffs that extracts a smaller share of revenues from those firms that make a better use of the technology. By doing so, the patentholder's profits increase with respect to the usual case of a single flat fee studied in the literature.

We examine a monopolistic competition framework, which presents a natural way of dealing with product differentiation in the downstream market. Moreover, by simplifying the strategic interaction among downstream competitors, we can focus the analysis on the optimization problem of the patentholder. Our model illustrates the different trade-offs in the choice of licensing contracts that the patentholder faces, and how they are affected by changes in the degree of differentiation in the downstream market. We obtain an intuitive characterization of the optimal contract within the class of two-part tariffs. Our results are general because this contract can be rewritten in terms of per-unit royalties and fees, which are more common in the literature. In this contract, royalties are also smaller for those firms that benefit more from the invention. These results also provide useful insights for other types of vertical relationship problems, such as franchising.

The varying degree of product differentiation has also important implications. In particular, when goods are more homogeneous and firms enjoy small markups over marginal cost in their prices, the profit-maximizing contract calls for higher variable payments and lower fees. The reason is that the marginal revenue of each firm is reduced, causing a decrease in the quantity they sell and this reduces effective competition. In the same direction, a second remarkable effect arises. The number of licenses is higher when goods are more homogeneous, as a way to reduce the total quantity that firms produce, which benefits the patentholder. We have interpreted both effects as the ability of inducing higher markups among the licensees, or of softening competition. Numerical results suggests, however, that welfare might be lower when two-part tariffs are used instead of flat fees because of this softening of competition, in spite of more firms having access to the new technology.

The setup we have introduced would allow us to ask more general questions. For example, we could easily accommodate a second patentholder producing an alternative technology that could benefit more firms with higher x and examine the choice of contracts offered by patentholders of alternative technologies.

The availability of data on royalty rates and licensing contracts in general is very limited. There are few instances where there is information on the details of licensing contracts—for example, in the case of franchise contracts and in the case of licensing contracts for international transfer of technology via foreign direct investment. The use of this type of data to test the model is left for future research.

A Proofs

Proof of Lemma 1

Using the envelope theorem, we observe that profits are decreasing in x,

$$\frac{\partial \pi\left(x, x\right)}{\partial x} = -\theta\left[c - \left(1 - x\right)\theta\right]^{-\frac{1}{1-\lambda}}\left[\left(1 - \alpha\left(x\right)\right)\lambda\kappa\right]^{\frac{1}{1-\lambda}} = -\theta y\left(x\right) < 0, \tag{16}$$

and the typical sorting condition can be computed as

$$\frac{\partial^2 \pi\left(x, x\right)}{\partial x \partial \alpha} = -\theta\frac{\partial y}{\partial \alpha} = \theta\frac{y\left(x\right)}{\left(1 - \lambda\right)\left(1 - \alpha\right)},$$

which is positive.[12]

Integrating equation (16) into the profit function we have

$$\Pi\left(\widetilde{x}\right) - \Pi\left(x\right) = \int_x^{\widetilde{x}} \frac{\partial \pi\left(x, x\right)}{\partial x}dx.$$

When $\widetilde{x} < 1$, $\Pi\left(\widetilde{x}\right) = \Pi_0$. Solving for $T\left(x\right)$ we obtain the desired result. When $\widetilde{x} = 1$ we know that $\Pi\left(1\right) \geq \Pi_0$ and it follows that $K \geq 0$.

Proof of Proposition 2

---

[12]In fact, if firms were not atomistic the cross-derivative would be

$$\frac{\partial^2 \Upsilon}{\partial x \partial \alpha} = -\theta\frac{\partial y}{\partial \alpha} = \theta\frac{y\left(x\right)}{\left(1 - \lambda\right)\left(1 - \alpha\right)}\left[1 - \lambda\left(\frac{P}{p\left(x\right)}\right)^{\frac{\lambda}{1-\lambda}}\right].$$

This last effect originates from the competition among final good producers. That is, each firm could take into account that an increase in the $\alpha\left(x\right)$ not only affects the quantity $y\left(x\right)$ and the corresponding price $p\left(x\right)$ but also the aggregate price index, P. Since firms are atomistic, we follow the standard simplification in the literature (see Dixit and Stiglitz, 1993) and we do not include this effect.

As usual in mechanism design, we assume that the incentive compatibility constraint is satisfied and we later verify. From the problem in (8) we can obtain the first-order condition with respect to $\alpha(x)$ as

$$\int_0^{\widetilde{x}} \left\{ \frac{\partial y(x')}{\partial \kappa} \frac{\partial \kappa}{\partial \alpha^*(x)} \left[(c - (1-x')\theta) \left(\frac{1}{(1-\alpha^*(x'))\lambda} - 1 \right) - \theta x' \right] \right.$$
$$\left. \frac{\Pi}{(1-\lambda)k} \frac{\partial \kappa}{\partial \alpha^*(x)} \right\} dx - \frac{y(x)}{(1-\alpha^*(x))(1-\lambda)} \left[(c - (1-x)\theta) \frac{\alpha^*(x)}{1-\alpha^*(x)} - \theta x \right] = 0$$

Define s as follows (to save notation, we omit the dependence on $\boldsymbol{\alpha}^*$ and λ)

$$s = \int_0^{\widetilde{x}} y(x') \left[(c - (1-x')\theta) \left(\frac{1}{(1-\alpha^*(x'))\lambda} - 1 \right) - \theta x' \right] dx' - \Pi_0 \widetilde{x}.$$

Given that $\frac{\partial y(x)}{\partial \kappa} = \frac{y(x)}{(1-\lambda)\kappa}$ we can rewrite the first-order condition as

$$\frac{\partial \kappa}{\partial \alpha^*(x)} \frac{sP^{-\lambda}}{1-\lambda} - \frac{y(x)}{(1-\alpha^*(x))(1-\lambda)} \left[(c - (1-x)\theta) \frac{\alpha^*(x)}{1-\alpha^*(x)} - \theta x \right] = 0,$$

and we can obtain the expression for $\frac{\partial \kappa}{\partial \alpha(x)}$ as

$$\frac{\partial \kappa}{\partial \alpha^*(x)} = P^\lambda \lambda \left[\int_0^1 p(x)^{-\frac{\lambda}{1-\lambda}} dx \right]^{-1} \frac{p(x)^{\frac{-\lambda}{1-\lambda}}}{1-\alpha(x)} = \frac{\kappa^2 \lambda y(x)^\lambda}{1-\alpha}.$$

Replacing in the previous equation we have

$$\lambda y(x)^\lambda \frac{\kappa}{1-\alpha^*(x)} \frac{s}{1-\lambda} - \frac{y(x)}{(1-\alpha^*(x))(1-\lambda)} \left[(c - (1-x)\theta) \frac{\alpha^*(x)}{1-\alpha^*(x)} - \theta x \right] = 0;$$

using the expression $y(x) = \left(\frac{P^\lambda}{p(x)} \right)^{\frac{1}{1-\lambda}}$, we can solve for $\alpha^*(x)$ as

$$\alpha^*(x) = \frac{s(c - (1-x)\theta) + \theta x}{c - (1-2x)\theta}. \tag{17}$$

Notice that from the budget constraint of the consumer we obtain that $V \le 1$. Moreover, it can be verified that

$$s = \int_0^{\widetilde{x}} y(x') \left[(c - (1-x')\theta) \left(\frac{1}{(1-\alpha^*(x'))\lambda} - 1 \right) - \theta x' \right] dx' - \Pi_0 \widetilde{x}$$
$$\le \max_{\{\alpha(x)\}_0^{\widetilde{x}}, \widetilde{x}} \int_0^{\widetilde{x}} y(x') \left[(c - (1-x')\theta) \left(\frac{1}{(1-\alpha(x'))\lambda} - 1 \right) - \theta x' \right] dx' - \Pi_0 \widetilde{x} = V$$

with equality when evaluated at the optimal $\widetilde{x}$. Since offering a contract $\alpha(x) = T(x) = 0$ is always feasible and leads to $s = 0$ we can conclude that $s \in [0, 1]$.

To verify that this expression characterizes a maximum, we can rewrite the first order condition as

$$\frac{y(x)^\lambda}{(1-\alpha(x))(1-\lambda)} \left[\lambda s \kappa - y(x)^{1-\lambda} \left(c(x) \frac{\alpha(x)}{1-\alpha(x)} - \theta x \right) \right].$$

The second derivative evaluated at $\alpha^*(x)$ corresponds to

$$\frac{y(x)^\lambda \kappa}{(1-\alpha^*(x))(1-\lambda)} \left[\lambda \left. \frac{\partial s}{\partial \alpha(x)} \right|_{\alpha^*(x)} + \frac{\lambda}{c(x)} \left(c(x) \frac{\alpha^*(x)}{1-\alpha^*(x)} - \theta x \right) - \frac{y(x)^{1-\lambda}}{\kappa} c(x) \frac{1}{(1-\alpha^*(x))^2} \right].$$

By the Envelope condition, $\frac{\partial s}{\partial \alpha(x)} = 0$ when evaluated at $\alpha^*(x)$ and therefore the expression can be rewritten as

$$\frac{\kappa y(x)^\lambda \lambda}{(1-\alpha^*(x))(1-\lambda)} \left[-\frac{\theta x}{c(x)} - 1 \right] < 0.$$

Existence and uniqueness of the equilibrium of the fixed point defined in (17) can be immediately verified through the following lemma:

31

Lemma 8 *Consider the function $\alpha^*(x)$ defined as $\alpha^*(x) = \gamma(x)(1-s) + s$, where*

$$s = F(s) \equiv \int_0^{\widetilde{x}} f(\alpha^*(x), x)\, dx \tag{18}$$

with f continuously differentiable, $\left.\frac{\partial F}{\partial \alpha(x)}\right|_{\alpha^*} = 0$ and $F(s) \in [0,1]$ for $s \in [0,1]$. Then, $\alpha^*(x)$ is uniquely defined.

Proof. Given that $\int_0^{\widetilde{x}} f(\alpha(x), x)\, dx$ is continuous in s in the interval $[0,1]$ for $s \in [0,1]$, existence of a fixed point is guaranteed using Brouwer's Fixed Point Theorem. Hence, there exists at least one value of s that satisfies (18). To show uniqueness of s, notice that $\frac{\partial F}{\partial s} = 0$ by assumption. As a result, in all points where $s = F(s)$ the derivative is 0. However, if there were more than one solution, for at least one of them $F'(s) > 1$ which is a contradiction. If s is unique, then α^* must be also unique. ∎

Proof of Proposition 3

The problem that the patentholder solves corresponds to

$$\max_{\alpha(x), \widetilde{x}} \int_0^{\widetilde{x}} \left\{ y(x)\left[(c - (1-x)\theta)\left(\frac{1}{(1-\alpha(x))\lambda} - 1\right) - \theta x\right] - \Pi \right\} dx,$$

under the same incentive compatibility constraint described in Lemma 1. The first-order condition becomes in this case

$$\frac{y(x)}{(1-\alpha)(1-\lambda)}\left[(c - (1-x)\theta)\frac{\alpha}{1-\alpha} - \theta x\right] = 0.$$

After solving for α we obtain the desired result.

Proof of Lemma 4

For the first notice that

$$
\begin{aligned}
s &= \int_0^{\widetilde{x}} y(x')\left[(c - (1-x')\theta)\left(\frac{1}{(1-\alpha^*(x'))\lambda} - 1\right) - \theta x'\right] dx' - \Pi_0 \widetilde{x} \\
&\leq \widetilde{x}\left(y(0)\left((c-\theta)\left(\frac{1}{(1-\alpha^*(0))\lambda} - 1\right)\right) - \Pi_0\right) \\
&= \widetilde{x}\kappa^{\frac{1}{1-\lambda}}\left(\left(\frac{(1-s)\lambda}{c-\theta}\right)^{\frac{\lambda}{1-\lambda}} - (c-\theta)\left(\frac{(1-s)\lambda}{c-\theta}\right)^{\frac{1}{1-\lambda}} - (1-\lambda)\lambda^{\frac{\lambda}{1-\lambda}}c^{\frac{-\lambda}{1-\lambda}}\right)
\end{aligned}
$$

where the limit of both terms is

$$\lim_{\lambda \to 0} \widetilde{x}\kappa^{\frac{1}{1-\lambda}} \leq 1$$

$$\lim_{\lambda \to 0} \left(\frac{(1-s)\lambda}{c-\theta}\right)^{\frac{\lambda}{1-\lambda}} - (c-\theta)\left(\frac{(1-s)\lambda}{c-\theta}\right)^{\frac{1}{1-\lambda}} - (1-\lambda)\lambda^{\frac{\lambda}{1-\lambda}}c^{\frac{-\lambda}{1-\lambda}} = 0$$

which means that the previous expression converges to 0. Since the patentholder can always guarantee profits of 0, $s = 0$.

With respect to the second expression, notice that when $\lambda = 1$ the price that a non-licensee charges is $p = c$. In this case, the highest price that a license can charge is $p(x) \leq c$. As a result, the sum of profits of the patentholder and all licensees, defined as

$$\int_0^{\widetilde{x}} (p(x) - c(x))y(x)\, dx \leq 1 - \int_0^{\widetilde{x}} c(x)y(x)\, dx \leq 1 - (c-\theta)\int_0^{\widetilde{x}} y(x)\, dx \leq \frac{\theta}{c}$$

where the last inequality originates from the fact that using the consumer budget constraint $\int_0^1 y(x)\, dx \leq \frac{1}{c}$. As a result, $s \leq \frac{\theta}{c}$.

Proof of Lemma 5

In a similar way as in *Lemma 1* we can define the profits of a licensee of type x that declares to be of type $\hat{x}$ as,

$$\pi(x, \hat{x}) = (1 - \lambda)\kappa \left(\frac{c(x) + r(\hat{x})}{\kappa\lambda} \right)^{\frac{1}{\lambda-1}} - F(\hat{x})$$

and it is easy to verify that using the envelope theorem,

$$\frac{\partial \pi(x, x)}{\partial x} = -\theta y(x) < 0$$

$$\frac{\partial \pi(x, x)}{\partial x \partial r} = -\theta \frac{\partial y}{\partial r} > 0$$

which proves the first point. The second result can be obtained by integrating the first of the previous expressions.

Proof of Proposition 6

Take an incentive compatible menu $\{\alpha(x), T(x)\}_{x \in [0, \tilde{x}]}$. From the argument in the text, if all firms produce the same quantities under both mechanisms profits for the licensees should be identical. This is immediate, since

$$y(x) = \left(\frac{c(x) + r}{\kappa\lambda} \right)^{\frac{1}{\lambda-1}} = \left(\frac{c(x) + c(x)\frac{\alpha(x)}{1-\alpha(x)}}{\kappa\lambda} \right)^{\frac{1}{\lambda-1}} = \left(\frac{c(x)}{\kappa\lambda(1 - \alpha(x))} \right)^{\frac{1}{\lambda-1}}.$$

Hence, if production is identical and prices do not change, consumers should obtain the same utility and the patentholder the same level of profits.

Proof of Lemma 7

Because the objective function in (8) is twice-differentiable, and the feasible set has the usual properties, showing that in any interior value of $\tilde{x}$ that satisfies the first order condition the objective function is locally concave is also enough to guarantee uniqueness. After some algebra, the first derivative can be rewritten as

$$\kappa^{\frac{1}{1-\lambda}} \left(p(\tilde{x})^{-\frac{\lambda}{1-\lambda}} (1 - (1 - s)\lambda - s) + \left(\frac{c}{\lambda} \right)^{-\frac{\lambda}{1-\lambda}} (s - (1 - \lambda)) \right)$$

and the second derivative evaluated at $\tilde{x}^*$ is

$$-\frac{\lambda(1 - (1 - s)\lambda - s)}{1 - \lambda} \kappa^{\frac{1}{1-\lambda}} p(\tilde{x}^*)^{-\frac{1}{1-\lambda}} \frac{\partial p}{\partial \tilde{x}}$$

which is negative, since $\frac{\partial p}{\partial \tilde{x}} > 0$. Notice that for this result we have used the envelope condition on s, so that $\frac{\partial s}{\partial \tilde{x}} = 0$ at that point.

B First order condition to the Flat Fees problem.

Given that firms with $x \leq \overline{x}$ purchase a license, the corresponding price index can be obtained as

$$P(\overline{x}) = \left[\int_0^{\overline{x}} \left(\frac{c - (1 - x)\theta}{\lambda} \right)^{-\frac{\lambda}{1-\lambda}} dx + \int_{\overline{x}}^1 \left(\frac{c}{\lambda} \right)^{-\frac{\lambda}{1-\lambda}} dx \right]^{-\frac{1-\lambda}{\lambda}},$$

where the individual prices are replaced by their expressions given by $p(x) = \frac{c(x)}{\lambda}$. In particular, firms with $x > \overline{x}$ have a marginal cost of c. The expression for κ is now

$$\kappa(\overline{x}) = P(\overline{x})^{\lambda},$$

while profits when a license is purchased are

$$\overline{\Pi}(x) = y(x)(p(x) - c(x)) = (c - (1 - x)\theta)^{-\frac{\lambda}{1-\lambda}} \lambda^{\frac{\lambda}{1-\lambda}} P(\overline{x})^{\frac{\lambda}{1-\lambda}} (1 - \lambda) - T.$$

If the firm does not obtain the license, profits are given by

$$\overline{\Pi}_0 = (1-\lambda)\left(\frac{c}{\lambda}\right)^{-\frac{\lambda}{1-\lambda}} P\left(\overline{x}\right)^{\frac{\lambda}{1-\lambda}}.$$

It is easy to verify that the increase in profits from obtaining a license,

$$\Delta\pi\left(x\right) - T = \overline{\Pi}\left(x\right) - \overline{\Pi}_0,$$

is decreasing in x. Because a firm x is willing to pay for the license up to the total increase in profits that it generates, the threshold will be $\overline{x} = 0$ if $\Delta\pi\left(0\right) \leq T$ and $\overline{x} = 1$ if $\Delta\pi\left(1\right) \geq T$; while it will correspond to $\Delta\pi\left(\overline{x}\right) = T$ otherwise. In the interior solution, choosing $\overline{x}$ is equivalent to choosing T. Hence, the problem that the patentholder solves is

$$\overline{V} = \max_{\overline{x}} \ \Delta\pi\left(\overline{x}\right)\overline{x}.$$

and the optimal number of licenses will be therefore

$$\overline{x} = \operatorname*{argmax}_{\hat{x}} \ \left(c\left(\hat{x}\right)^{-\frac{\lambda}{1-\lambda}} - c^{-\frac{\lambda}{1-\lambda}}\right) P\left(\hat{x}\right)^{\frac{\lambda}{1-\lambda}} \hat{x}.$$

And the first order condition for this problem can be written as

$$-\overline{x}^* \left[\left(c(\overline{x}^*)^{-\frac{\lambda}{1-\lambda}} - c^{-\frac{\lambda}{1-\lambda}}\right) P\left(\overline{x}^*\right)^{\frac{2\lambda}{1-\lambda}} \left(p\left(\overline{x}^*\right)^{\frac{-\lambda}{1-\lambda}} - \left(\frac{c}{\lambda}\right)^{\frac{-\lambda}{1-\lambda}}\right) + \right. \tag{19}$$

$$\left. +\frac{\lambda\theta}{1-\lambda}c(\overline{x}^*)^{\frac{-1}{1-\lambda}} P\left(\overline{x}^*\right)^{\frac{\lambda}{1-\lambda}}\right] + \left(c(\overline{x}^*)^{-\frac{\lambda}{1-\lambda}} - c^{-\frac{\lambda}{1-\lambda}}\right) P\left(\overline{x}^*\right)^{\frac{\lambda}{1-\lambda}} = 0,$$

where $p\left(\overline{x}^*\right) = \frac{c(\overline{x}^*)}{\lambda}$. Moreover, notice that $p\left(\overline{x}^*\right)^{\frac{-\lambda}{1-\lambda}} - \left(\frac{c}{\lambda}\right)^{\frac{-\lambda}{1-\lambda}} = \lambda^{\frac{\lambda}{1-\lambda}}\left(c(\overline{x}^*)^{-\frac{\lambda}{1-\lambda}} - c^{-\frac{\lambda}{1-\lambda}}\right)$. Since $P\left(\overline{x}^*\right)^{\frac{\lambda}{1-\lambda}} > 0$, this condition can be further simplified into

$$-\overline{x}^*\left[\frac{\lambda\theta}{1-\lambda}c(\overline{x}^*)^{\frac{-1}{1-\lambda}} + \lambda^{\frac{\lambda}{1-\lambda}}\left(c(\overline{x}^*)^{-\frac{\lambda}{1-\lambda}} - c^{-\frac{\lambda}{1-\lambda}}\right)^2 P\left(\overline{x}^*\right)^{\frac{\lambda}{1-\lambda}}\right] + \left(c(\overline{x}^*)^{-\frac{\lambda}{1-\lambda}} - c^{-\frac{\lambda}{1-\lambda}}\right) = 0.$$

References

Arora, A. and A. Fosfuri (1998) "Licensing in the presence of competing technologies," *mimeo.*

Beggs, A. W. (1992) "The Licensing of Patents under Asymmetric Information," *International Journal of Industrial Organization*, **10**: 171-191.

Bouquet, A., Cremer, H., Ivaldi, M. and M. Wolkowicz (1998) "Risk Sharing in Licensing," *International Journal of Industrial Organization*, **16**: 535-554.

Choi, J. P. (2001) "Technology Transfer with Moral Hazard," *International Journal of Industrial Organization,* **19**: 249-266.

Dixit, A. and J. E. Stiglitz (1977) "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, **67**: 297-308.

Dixit, A. and J. E. Stiglitz (1993) "Monopolistic Competition and Optimum Product Diversity: Reply" *American Economic Review*, **83**: 302-304.

Erutku C. and Y. Richelle (2001) "Licensing a New Product with Non-Linear Contracts," mimeo.

Faulí-Oller, R. and J. Sandonis (2002) "Welfare Reducing Licensing," *Games and Economic Behavior*, **41**: 192-205.

Gallini, N. T. and B. D. Wright (1990) "Technology Transfer under Asymmetric Information," *RAND Journal of Economics*, **21**: 147-160.

Jensen, R. and M. Thursby (2001) "Proofs and Prototypes for Sale: The Licensing of University Inventions," *American Economic Review*, **91**: 240-259.

Kamien, M. I. (1992) "Patent Licensing" in the *Handbook of Game Theory* Vol 1, Elsevier Science Publishers.

Kamien, M. I. and Y. Tauman (1984) "The Private Value of a Patent: A Game Theoretic Analysis," *Journal of Economics*, **4**: 93-118.

Llobet, G. (2003) "Patent Litigation when Innovation is Cumulative," *International Journal of Industrial Organization*, **21**: 1135-1158.

Macho-Stadler, I. and D. Pérez-Castrillo (1996) "An Introduction to the Economics of Information: Incentives and Contracts," *Oxford University Press.*

Maskin, E. and J. Riley (1984) "Monopoly with Incomplete Information," *RAND Journal of Economics*, **15**: 171-196.

Muto, S. (1993) "On Licensing Policies in Bertrand Competition," Games and Economic Behavior, **5**: 257-267.

Rostoker (1984) "A survey of corporate licensing", *IDEA*, **24**: 59-92.

Saracho, A. (2001) "Patent Licensing under Strategic Delegation," *Journal of Economics and Management Strategy*, **11**:225-251.

Saracho, A. (2002) "The Relationship Between Patent Licensing and Competitive Behavior," *mimeo*.

Taylor, C. T. and Z. A. Silberston (1973) "The Economic Impact of the Patent System: A Study of the British Experience," Cambridge University Press.

Wang, X. H. (1998) "Fee versus royalty licensing in a Cournot duopoly model," *Economic Letters*, **60**: 55-62.

Wang, X. H. and B. Yang (1999) "On Licensing under Bertrand Competition," *Australian Economic Papers*, **38**:106-119.