Optimal Licensing of Non-Drastic and (Super-)Drastic Innovations: The Case of the Inside Patent Holder*

Cuihong Fan†
Shanghai University of Finance and Economics
cuihongf@mail.shufe.edu.cn

Byoung Heon Jun
Korea University, Seoul
bhjun@korea.ac.kr

Elmar G. Wolfstetter
Humboldt-University at Berlin and Korea University, Seoul
wolfstetter@gmail.com

October 31, 2016

Abstract

We reconsider the inside innovators' optimal licensing problem, assuming incomplete information and unit cost profiles that may or may not have the potential to propel a monopoly, taking into account restrictions concerning royalty rates and the use of exclusive licenses implied by antitrust rules. We analyze optimal licensing mechanisms using methods developed in the analysis of license auctions with downstream interaction. The optimal mechanism differs significantly from the mechanisms reported in the literature, which assumed complete information or particular cost profiles or probability distributions.

KEYWORDS: Innovation, licensing, optimal contracts, asymmetric information.
JEL CLASSIFICATIONS: D21, D43, D44, D45.

1 Introduction

The literature on patent licensing in oligopoly markets draws a sharp distinction between licensing by an *outside* innovator and an *inside* patent holder who is a competitor in the product market of potential licensees. Whereas outside innovators are advised to either use fixed fee contracts or auction patent licenses to a limited number of licensees, inside patent holders are advised to employ output based royalty contracts without fixed fees.

The present paper reconsiders the inside innovator's optimal licensing problem, assuming incomplete information and unit cost profiles that may or may not have the potential to propel a monopoly, taking into account restrictions concerning royalty rates and the use of exclusive licenses implied by antitrust rules. Due to the nature of the licensing problem and its constraints, we analyze optimal license contracts using methods developed in the analysis of license auctions

*Research support by Korea University (Grant: K1509111), the *Deutsche Forschungsgemeinschaft (DFG), SFB Transregio 15,* "Governance and Efficiency of Economic Systems", and the National Natural Science Foundation of China (Grant: 71371116) is gratefully acknowledged.

†Corresponding author. School of Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, China.

with downstream interaction. Our analysis indicates that optimal license contracts exhibit a complex pattern of state dependent royalty rates and fixed fees that differ substantially from the mechanisms reported in the literature, which assumed complete information or particular cost profiles or probability distributions.

Our results are broadly in line with the predominance of two-part tariffs with output dependent royalties and fixed fees. In a study of US firms Rostoker (1984, p. 64) reports that: "A down payment with running royalties method was used 46% of the time, while straight royalties and paid-up licenses accounted for 39% and 13%, respectively".[1] The use of output dependent royalties is also mirrored in the fact that license agreements typically "grant inspection rights aimed at controlling the licensees use of the licensed technology" (Brousseau, Coeurderoy, and Chaserant, 2007)

There is an extensive licensing literature. The outside innovators' licensing problem was initiated by Kamien and Tauman (1984) and Kamien and Tauman (1986) and Kamien (1992), who show that fixed-fee licensing and license auctions are more profitable than royalty contracts. However, the subsequent literature showed that royalties may be useful to signal private information of the licensee or the innovator (Gallini and Wright, 1990; Beggs, 1992), to induce the innovator to create and share relevant know-how concerning the use of the innovation (Macho-Stadler, Martínez-Giralt, and Pérez-Castrillo, 1996; Choi, 2001), or to extract additional surplus from the losers of the auction (Giebe and Wolfstetter, 2008; Fan, Jun, and Wolfstetter, 2013; Fan, Jun, and Wolfstetter, 2014).

The inside innovators' licensing problem was initiated by Wang (1998) and Sen and Tauman (2007). They assume that the licensee is not more efficient, conditional on using the innovation and show that royalty contracts with a royalty rate equal to the cost reduction and without fixed fees are optimal. This way, the innovator neutralizes the licensee's cost reduction due to using the innovation and thus protects his turf.

The recent literature on the inside innovators' licensing problem made inroads towards covering more general cost profiles and introducing incomplete information. Poddar and Sinha (2010) consider an example in which the licensee is more efficient, assuming a common value innovation under complete information.[2] There, the greater efficiency of the licensee after licensing occurs because prior to the innovation the licensee has lower cost and using the innovation reduces their cost by the same amount.

Heywood, Li, and Ye (2014) made a first contribution to incorporate incomplete information, assuming a binary model of private value cost reductions. They confirm the optimality of pure royalty contracts.[3] However, like the bulk of the literature, they consider non-drastic innovations with a particular profile of cost reductions, where the innovator is at least as efficient as the licensee with probability one.

In the present paper we solve the licensing problem between one inside innovator and one potential licensee, assuming that the licensee has private information about the cost induced by using the innovation and the two firms play a Cournot market game. The innovator designs and proposes the licensing mechanism. We assume that the innovation can be transferred and fully disclosed at zero

[1] See also Calvert (1964), Taylor and Silberston (1973), Macho-Stadler, Martínez-Giralt, and Pérez-Castrillo (1996), and Vishwasrao (2007) on foreign technology licensing to Indian firms.

[2] An extension of this analysis to three firms is in Wang, Liang, and Chou (2013).

[3] Heywood, Li, and Ye (2014) also consider *ad valorem* royalties and show that for most parameter values output based royalties are superior and fixed fees are never employed.

cost, without causing moral hazard problems that may distort incentives for sharing the know-how related to using the innovation.

In a nutshell, our results are as follows: if the innovation is non-drastic, the optimal licensing mechanism generally exhibits four regimes: a two-part tariff regime with decreasing royalty rates and positive constant fixed fees, a two-part tariff regime with increasing royalty rates and decreasing fixed fees, a pure royalty regime with decreasing royalty rates and zero fixed fee, and an exclusion regime where no license is awarded. Royalties are always used; yet, pure royalty contracts are optimal in the special case when the innovator is always more efficient than the licensee.

In the case of drastic innovations we identify cases in which the innovator either issues no license, and makes himself a monopoly, or issues an exclusive license and makes the licensee a monopoly. In the latter case, the optimal mechanism prescribes state dependent two-part tariffs. In the presence of incomplete information output based royalties are essential because linking the license fee to a variable that is correlated with the licensee's private information tends to lower information rents.

The plan of the paper is as follows: In Section 2 we state the model. In Section 3 we spell out basic distinctions between exclusive and non-exclusive licenses and between non-drastic and (super-)drastic innovations. In Section 4 we state the mechanism design problem for non-drastic innovations and characterize the optimal mechanism for different probability distributions of cost profiles induced by the innovation. In Section 5 we extend the analysis to (super-)drastic innovations and characterize the optimal patent sale mechanism that makes either the licensee or the innovator a monopoly. We close with a brief discussion.

2 The model

Consider a dynamic licensing game played between an inside innovator of a process innovation (firm 1) and a potential licensee (firm 2). In the first stage the innovator offers a menu of license contracts from which the licensee either accepts one or rejects. In the second stage firms play a Cournot game.

Prior to the innovation firms' unit cost profile is $(c_1, c_2) = (c, c)$.[4] Using the innovation reduces the unit cost of firm 1 from c to d, and licensing reduces the unit cost of firm 2 to x, which is that firm's private information.

A license is either exclusive, in which case only the licensee can use the innovation, or non-exclusive, in which case the innovator can also use the innovation. Non-exclusive licensing changes the profile of unit costs from $(c_1, c_2) = (d, c)$ to (d, x), whereas exclusive licensing changes the cost profile to (c, x).

From the perspective of the innovator x is a random variable drawn from the continuous c.d.f. G with support $[\underline{x}, \bar{x}]$, $0 \le \underline{x} < \bar{x} \le c$. Generally, we assume G to be log-concave, unless we explicitly remove that restriction.

The payoff functions of the duopoly games are: $\pi_1(q_1, q_2; c_1) = (P(Q) - c_1)q_1$, $\pi_2(q_1, q_2; c_2) = (P(Q) - c_2)q_2$, $Q := q_1 + q_2$. For simplicity and in order to obtain closed form solutions, we assume $P(Q) := 1 - Q$.

[4] This symmetry assumption is not essential and used only to simplify the notation.

The innovator employs a direct mechanism, $M := (t(x,q_2), \gamma(x))$, to license his innovation which consists of an allocation rule $\gamma(x)$ and a transfer rule $t(x,q_2)$, as a function of the reported unit cost, x, and the output q_2 that is observed after the oligopoly game has been played.

We consider incentive compatible mechanisms with a deterministic allocation rule, $\gamma(x) \in \{0,1\}$, where $\gamma(x) = 1$ ($\gamma = 0$) means that the license is (not) awarded.

The transfer rule prescribes firm 2 to pay a royalty rate $r(x)$ per output unit plus a fixed fee, $f(x)$, if the license is awarded and pay nothing if no license is awarded:

$$t(x,q_2) = \begin{cases} f(x) + r(x)q_2 & \text{if} \quad \gamma(x) = 1 \\ 0 & \text{if} \quad \gamma(x) = 0. \end{cases}$$

(Note that this specification does not allow for output dependent royalty rates.) The potential licensee reports his unit cost, and the innovator adopts the allocation and transfers prescribed by the mechanism for the reported cost.

License contracts are regulated by antitrust authorities who interfere if they suspect collusive schemes that are geared to transform the market structure. We capture these regulations by constraints on $r(x)$ that are stated when needed, in Sections 4 and 5.

3 (Non-)drastic innovations and (non-)exclusive licenses

At the outset we state some facts about possible cost profiles induced by an innovation and introduce distinctions between non-exclusive and exclusive licenses, and between (super-)drastic and non-drastic innovations.

Licensing an innovation reduces the unit cost of the licensee from c to x. The bulk of the literature on the inside innovators' licensing problem assumed that the innovator has a cost advantage with probability one, i.e., $d \leq \underline{x}$. However, this is far too restrictive. There are many examples where the licensee can make better use of the innovation and has a lower unit cost after using the innovation. A case in point is the float glass method of glass making, patented by *Pilkington Brothers*, that revolutionized the glass industry. Pilkington offered numerous licenses to glass manufacturers in distant locations. The cost of transporting glass is a major concern. Therefore, after using the innovation, glass manufacturers in distant locations had a cost advantage in serving neighboring markets relative to Pilkington's own plants, which suggested a licensing policy that induced local monopolies, mirrored in prominent antitrust investigations.

If the innovator grants a *non-exclusive license*, he permits the licensee to use that innovation while the innovator remains free to also use the innovation. If the licensor grants an *exclusive license*, after transferring the property right he can no longer use the innovation.[5] The most reliable way to issue an exclusive license is to sell the patent, which is why we refer to exclusive licensing as the sale of the patent.

Evidently, exclusive and non-exclusive licensing induce different cost profiles; non-exclusive licensing induces $(c_1, c_2) = (d, x)$; exclusive licensing induces $(c_1, c_2) = (c, x)$.

An innovation is *drastic* for firm i if that firm's exclusive use of the innovation makes it a monopoly; it is *super-drastic* for firm i if that firm's (exclusive or non-exclusive) use of the innovation makes it a monopoly

[5]If there are several potential licensees, an exclusive license may also restrict the right to license to competitors.

Stated formally: an innovation is *drastic for the innovator* if the best-reply of firm 2 to the monopoly output of firm 1, q_1^M, is equal to zero, i.e., if $q_2^* := \arg\max_{q\geq 0}\left(1-q_2-q_1^M-c\right)q = 0$. It is *drastic for the licensee* if, conditional on having acquired an exclusive license, one has $q_1^* := \arg\max_{q\geq 0}\left(1-q-q_2^M-c\right)q=0$, and it is *super-drastic for the licensee* if $q_1^{**} := \arg\max_{q\geq 0}\left(1-q-q_2^M-d\right)q = 0$.

The following conditions are necessary and sufficient:

$$\begin{aligned} q_2^* = 0 &\iff d \leq 2c-1 \quad \text{(drastic for the innovator)} \\ q_1^* = 0 &\iff x \leq 2c-1 \quad \text{(drastic for the licensee)} \\ q_1^{**} = 0 &\iff x \leq 2d-1 \quad \text{(super-drastic for the licensee).} \end{aligned}$$

Evidently, an innovation is drastic if it is super-drastic, but not *vice versa*.

4 Optimal licensing of non-drastic innovations

In this section we assume that the innovation is non-drastic for both firms, i.e., $d > 2c-1$ and $\underline{x} > 2c-1$, and that the license is non-exclusive. Issuing an exclusive license by selling the patent would only be interesting for the innovator if he could thus implement a monopoly. However, in that case, the antitrust authorities would view the sale of the patent as a collusive scheme that is geared to replace a duopoly by a monopoly.[6]

In oligopoly markets, license contracts are regulated by antitrust authorities who interfere if they suspect collusive schemes that are geared to transform the market structure.[7] We capture these regulations by two constraints that impose upper and lower bounds on $r(x)$:

$$r(x) \leq c - x \tag{1}$$
$$x + r(x) \geq 2d - 1. \tag{2}$$

Constraint (1) requires that royalty rates cannot exceed cost reductions. Without this constraint the innovator could artificially raise his rival's cost. In the extreme this could make the innovator a monopoly by setting a high royalty rate combined with a negative fixed fee, even if the innovation is non-drastic.[8] Constraint (2) requires that the innovator cannot make the licensee a monopoly by artificially lowering its effective unit cost (the form of that constraint will become clear when we state the equilibrium outputs in the section below).

In order to rule out that the license is misused as a "money pump" by producing infinitely large outputs we also require non-negativity of effective unit costs, $x + r(x) \geq 0$. Combining this non-negativity constraint with the lower-bound constraint (2) gives the constraint:

$$x + r(x) \geq \max\{2d-1, 0\}. \tag{3}$$

We refer to the inequalities (1) and (3) as upper- and lower-bound antitrust constraints.

[6]According to the Antitrust Guidelines for the Licensing of Intellectual Property in the U.S.: "Generally, an exclusive license may raise antitrust concerns only if the licensees themselves, or the licensor and its licensees, are in a horizontal relationship. Examples of arrangements involving exclusive licensing that may give rise to antitrust concerns include ... acquisitions of intellectual property rights" (U.S. Department of Justice and Federal Trade Commission, 1995, Section 4.1.2).

[7]Licensing agreements for intellectual property are typically illegal if they involve naked price-fixing, output restrictions, or market division (U.S. Department of Justice and Federal Trade Commission, 1995).

[8] *"Quite appropriately, contracts of this form would likely be held to be illegal by antitrust authorities"* (Katz and Shapiro, 1985, p. 513).

The innovator maximizes his expected payoff (which is the sum of his own expected profit and expected license revenue), subject to incentive compatibility, participation, and antitrust constraints.

In order to solve the set of incentive compatible mechanisms, we proceed as follows: first, we compute the payoffs of the licensee for all combinations of his true cost x and reported cost z. This requires that we solve all oligopoly subgames that may occur on and off the equilibrium path. We are then able to characterize incentive compatible mechanisms.[9]

4.1 Oligopoly subgames on and off the equilibrium path

Suppose the licensee with cost x reports the cost z (possibly different from x) for which a license is awarded. In that case firm 1 believes to play a duopoly game with the cost profile $(c_1, c_2) = (d, z + r(z))$. Denote the equilibrium of that game by $(q_1(z), \tilde{q}_2(z))$, defined as

$$q_1(z) := \arg\max_q \pi_1(q, \tilde{q}_2(z); d), \quad \tilde{q}_2(z) := \arg\max_q \pi_2(q_1(z), q; z + r(z)).$$

However, firm 2 privately knows that the cost profile is $(c_1, c_2) = (d, x + r(z))$ and therefore plays its best reply to $q_1(z) = (1-2d+z+r(z))/3$, which gives:

$$q_2(x,z) = \arg\max_q \pi_2(q_1(z), q; x + r(z)) = \frac{1}{6}\left(2 + 2d - 3x - z - 4r(z)\right). \tag{4}$$

The associated equilibrium profit of firm 2 is

$$\pi_2\left(q_1(z), q_2(x,z); x + r(z)\right) = q_2(x,z)^2. \tag{5}$$

As a special case one obtains the "on the equilibrium path" strategies and profits that apply when firm 2 reports truthfully (i.e., $z = x$):

$$q_1^*(x) := q_1(x) = \frac{1 - 2d + x + r(x)}{3}, \quad \pi_1^*(x) = q_1^*(x)^2 \tag{6}$$

$$q_2^*(x) := q_2(x,x) = \frac{1 + d - 2(x + r(x))}{3}, \quad \pi_2^*(x) = q_2^*(x)^2. \tag{7}$$

In turn, if no license is awarded, the two firms play the equilibrium strategies q_1^N, q_2^N that solve the requirements

$$q_1^N = \arg\max_q \pi_1(q, q_2^N; d) = \frac{1 - 2d + c}{3}, \quad q_2^N = \arg\max_q \pi_2(q_1^N, q; c) = \frac{1 - 2c + d}{3}. \tag{8}$$

The associated equilibrium profits are $\pi_1\left(q_1^N, q_2^N; d\right) = (q_1^N)^2$, $\pi_2\left(q_1^N, q_2^N; c\right) = (q_2^N)^2$. For convenience, we define:

$$\Pi_2^N := \pi_2\left(q_1^N, q_2^N; c\right), \quad \Pi_2^L(x,z) := \pi_2\big(q_1(z), q_2(x,z); x + r(z)\big) - f(z).$$

There, L and N are mnemonic for licensing and no-licensing (exclusion).

[9]This procedure is similar to the solution method employed in the analysis of license auctions with downstream interaction.

4.2 Incentive compatibility

Using the above equilibria of the duopoly subgames for all combinations of z and x, the potential licensee's expected payoff in the licensing game is:

$$\Pi_2(x,z) := \gamma(z)\Pi_2^L(x,z) + (1-\gamma(z))\Pi_2^N.$$

The mechanism is incentive compatible if $x = \arg\max_z \Pi_2(x,z), \forall x$. In the following we show, among other things, that incentive compatible mechanisms prescribe an exclusion threshold level of x, denoted by $\hat{x} \in [\underline{x}, \bar{x}]$, above which no license is awarded.

Lemma 1. *The following conditions are necessary for incentive compatibility,*

$$\gamma(x) = \begin{cases} 1 & \textit{if} \quad x \le \hat{x} \\ 0 & \textit{otherwise} \end{cases} \tag{9}$$

$$f'(x) = -\frac{1}{9}\left(1+d-2x-2r(x)\right)\left(1+4r'(x)\right) \tag{10}$$

$$\frac{r(x)-r(z)}{x-z} \ge -\frac{1}{4}. \tag{11}$$

Proof. To prove that γ is monotone decreasing, as asserted in (9), suppose the mechanism prescribes $\gamma(x') = 1, \gamma(x) = 0$ for some x', x with $x' > x$. Then, the licensee with cost x has an incentive to report $z = x'$, because in that case he obtains the license and earns an even higher profit than type x'.

For all $z \le \hat{x}$ (for which a license is awarded), incentive compatibility requires that $\partial_z \Pi_2^L(x,z)\big|_{z=x} = 0$, for all $x, z \le \hat{x} \le \bar{x}$, which implies (10). Moreover, incentive compatibility also requires that $\Pi_2^L(x,x) - \Pi_2^L(x,z) \ge 0$ and $\Pi_2^L(z,z) - \Pi_2^L(z,x) \ge 0$, and therefore:

$$0 \le \Pi_2^L(x,x) - \Pi_2^L(x,z) + \Pi_2^L(z,z) - \Pi_2^L(z,x) = \frac{1}{6}\left(x-z\right)\left(x-z+4r(x)-4r(z)\right).$$

Rearranging gives (11). □

We stress that these requirements are only necessary conditions for incentive compatibility. In our solution of the optimal mechanism we will also confirm that sufficient conditions are satisfied.

4.3 Participation and antitrust constraints

The mechanism has to assure voluntary participation:

$$\Pi_2(x,x) \ge \Pi_2^N, \quad \forall x. \tag{12}$$

Note that $\Pi_2(x,x)$ is non-increasing:

$$x' > x \Rightarrow \Pi_2(x,x) \ge \Pi_2(x,x') \ge \Pi_2(x',x').$$

by incentive compatibility and the fact that $\Pi_2(x,x')$ is decreasing in x. Therefore, if the participation constraint is satisfied for some x', it is also satisfied for all $x < x'$. It follows that participation constraints can be replaced by:

$$\Pi_2^L(\hat{x},\hat{x}) \ge \Pi_2^N \quad \text{(participation constraint)}. \tag{13}$$

Note that the constraint concerning the slope of the $r(x)$ function, (11), implies $r(x) \le r(\hat{x}) + \frac{\hat{x}-x}{4}$, for all $x \le \hat{x}$ and $r(x) \ge r(\underline{x}) - \frac{x-\underline{x}}{4}$ for all $x \ge \underline{x}$. Combining these with the upper- and lower-bound antitrust constraints, (1), (3), gives us the constraints:

$$k_1(x) := r(x) + \frac{x+3\underline{x}}{4} - \max\{2d-1,0\} \ge 0, \quad \text{for } x \ge \underline{x} \tag{14}$$

$$k_2(x) := c - \frac{3\hat{x}+x}{4} - r(x) \ge 0, \quad \text{for } x \le \hat{x}. \tag{15}$$

We replace the lower- and upper-bound antitrust constraints, (3) and (1), by (14) and (15). Note that these (augmented) antitrust constraints incorporate the constraint concerning the slope of the r function.

4.4 Optimal mechanism

We now characterize the optimal mechanism that maximizes the innovator's expected payoff subject to the necessary conditions for incentive compatibility, (9)-(11), the participation constraint, (13), lower- and upper-bound antitrust constraints, (14), (15). In the proofs it will become clear that sufficient conditions for global incentive compatibility are satisfied.

Proposition 1. *Assume $\left(G(x)/G'(x)\right)' > -2/3$, which is assured if G is log-concave. The optimal mechanism prescribes four regimes: 1) a two-part tariff regime with $r(x)$ decreasing and $f(x)$ positive and constant, 2) a two-part tariff regime with $r(x)$ increasing and $f(x)$ decreasing, 3) a pure royalty regime with $r(x)$ decreasing and $f(x) = 0$, and 4) an exclusion regime: $\gamma(x) = 0$, for $x > \hat{x}$:*

$$r(x) = \begin{cases} r_1(x) := -\frac{x+3\underline{x}}{4} + \max\{2d-1,0\} & \text{if } x \le x_1 \\ r_0(x) := \frac{1-5d+4x}{2} + \frac{3G(x)}{G'(x)} & \text{if } x \in [x_1, x_2] \\ r_2(x) := c - \frac{3\hat{x}+x}{4} & \text{if } x \ge x_2 \end{cases} \tag{16}$$

$$\begin{aligned} f(x) &= f(\hat{x}) - \int_x^{\hat{x}} f'(y)dy, \quad f(\hat{x}) = \pi_2^*(\hat{x}) - \Pi_2^N \\ f'(x) &= -\frac{1}{9}(1+d-2x-2r(x))(1+4r'(x)). \end{aligned} \tag{17}$$

There, $r_1(x)$ and $r_2(x)$ are obtained from $k_1(x) \equiv 0$ and $k_2(x) \equiv 0$.

A license is awarded with positive probability $\hat{x} > \underline{x}$. The thresholds x_1, x_2 that induce regime changes and the optimal exclusion threshold $\hat{x}$ depend upon the parameters and the probability distribution.

Proof. The proof is in five steps : 1) We state the optimal mechanism in the class of mechanisms that satisfy the necessary conditions for incentive compatibility, the participation and antitrust constraints, and state the optimality conditions. 2) We derive the associated r and 3) the associated f functions. 4) We show that the solution satisfies global incentive compatibility and second-order conditions. 5) We show that licensing occurs with positive probability.

The details of the elaborate proof are spelled out in Appendix A.1. □

An immediate implication of Proposition 1 is:

Corollary. *Suppose G shifts in such a way that the new distribution exhibits a higher reversed hazard rate (which implies that it first-order stochastically dominates G). Then r_0 shifts downwards, and the threshold level x_1 increases.*

Proposition 2. *The optimal exclusion threshold $\hat{x}$ is never in the interval $[x_1, x_2)$, unless $\bar{x} \in [x_1, x_2)$.*

Proof. In the Appendix A.2 we show that $\Pi_1^*(\hat{x})$ is increasing on $[x_1, x_2)$. This proves the assertion. $\square$

Based on this result, the optimal exclusion threshold level $\hat{x}$ can be narrowed down, depending upon $\underline{x}, \bar{x}$ and x_1, x_2, as follows:

$$x_1 \leq \underline{x}, \bar{x} \geq x_2 \Rightarrow \hat{x} \in [x_2, \bar{x}], \qquad x_1 \leq \underline{x}, \bar{x} < x_2 \Rightarrow \hat{x} = \bar{x}$$
$$x_1 > \underline{x}, \bar{x} \geq x_2 \Rightarrow \hat{x} \in [x_2, \bar{x}] \cup [\underline{x}, x_1], \qquad x_1 > \underline{x}, \bar{x} < x_2 \Rightarrow \hat{x} \in [\underline{x}, x_1] \cup \{\bar{x}\}.$$

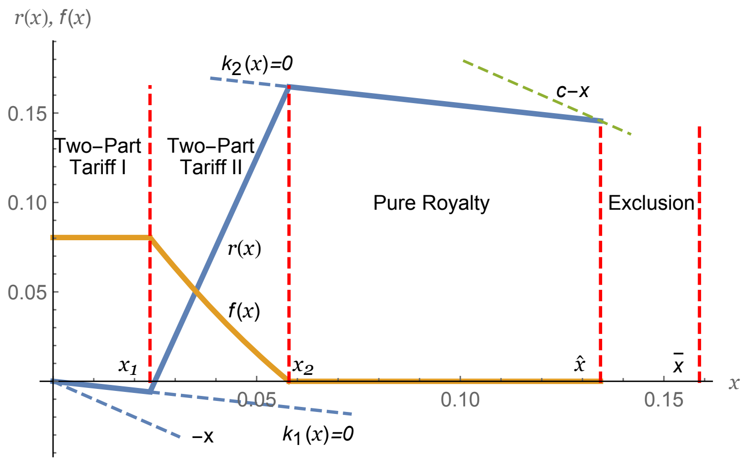

Figure 1: Optimal mechanism

Figure 1 illustrates Proposition 1. There, we plot the boundaries of the lower- and upper-bound constraints, $k_1(x) = 0, k_2(x) = 0$, the optimal r, f functions, and the optimal thresholds at which regime changes occur, $(x_1, x_2, \hat{x})$, assuming a uniform distribution with support $[0, 10/63]$, $d = 1/4$, and $c = 28/100$.[10] The optimal mechanism exhibits three regimes, and one can see clearly how the segments of the $r(x)$ function that apply in the "Two-Part Tariff I" and the "Pure Royalty" regimes are fully determined by the k constraints.[11]

In the "Two-Part-Tariff I" regime, the licensee has a significant cost advantage and the benefit of shifting output to the more efficient firm outweighs the benefit of restricting the output of the licensee. Shifting output to the more efficient firm is achieved by a low, and in this case even negative, royalty rates. Royalty rates are, however, bounded from below by the lower-bound antitrust constraint and the incentive compatibility constraint concerning the slope of the r function. These constraints are captured by the k_1 constraint, which is binding everywhere in this regime.

In the "Pure Royalty" regime, the cost advantage of the licensee is relatively low, and the benefit of restricting the output of the licensee through high royalty rates gains outweighs the benefit of

[10] The exact values of the threshold levels are: $(x_1, x_2, \hat{x}) = (0.024, 0.058, 0.134)$, and one has $(r(x), f(x)) = (-x/4,\ 405169/5040000)$ for $x \leq x_1$, $(r(x), f(x)) = \left(5x - 1/8,\ 12463/80000 - 7x/2 + 14x^2\right)$ for $x_1 \leq x \leq x_2$, and $(r(x), f(x)) = \left(\frac{850 - 3\sqrt{35259}}{1600} - x/4,\ 0\right)$ for $x \geq x_2$.

[11] The only peculiar feature of this example is that the royalty rate is negative in the "Two-Part Tariff I" regime. This does, however, not occur if one increases the parameter d or $\underline{x}$.

shifting output to the more efficient firm. Throughout this regime, the k_2 constraint is binding. More specifically, the incentive compatibility constraint concerning the slope of the r function is binding and the upper-bound antitrust constraint binds at (and only at) the threshold level $\hat{x}$, at which exclusion sets in. This also indicates that more exclusion, i.e., a lower $\hat{x}$, leads to higher royalty rates in that regime.

In the "Two-Part-Tariff II" regime neither one of the two k constraints is binding. There, the r function is equal to $(1-5d+4x)/2$ plus the multiple of the inverse of the reversed hazard rate, $3G(x)/G'(x)$. The first term is the royalty rate that would be optimal under complete information if one ignored antitrust constraints. The second term induces distortions of efficiency, except at $x = \underline{x}$; the latter represents the familiar "no distortion at the top" property, with "top" referring to the lowest cost level.

In order to interpret the optimal distortions, note that the "reversed hazard rate", $G'(x)/G(x)$, is the conditional probability that the licensee's x belongs to the interval $[x-dx, x]$, given that it is known to belong to the interval $[\underline{x}, x]$. Adding a distortion by increasing the royalty rate in the interval $[x-dx, x]$ makes it less attractive to falsely report a cost level from that interval. Therefore, one can extract more surplus from the types whose cost level is below $x-dx$, while maintaining incentive compatibility.

However, this benefit comes at the cost of foregoing some surplus extraction from the types in interval $[x-dx, x]$, whose surplus is reduced by the distortion. This cost is the higher the more probability mass is on that interval $[x-dx, x]$, i.e., the higher the reversed hazard rate.

The optimal distortion trades off these benefits and costs. Therefore, the optimal distortion is decreasing in the reversed hazard rate, and, equivalently, increasing in the inverse of the reversed hazard rate.

If the reversed hazard rate is strictly decreasing, which is the case if G is log-concave, the distortion term is increasing, and hence high cost types are subject to a higher distortion than low cost types. This is why the constraint concerning the slope of the $r(x)$ function never binds in this case.

So far we assumed that the probability distribution function satisfies the requirement that $\left(G(x)/G'(x)\right)' > -2/3$, which holds true if G is log-concave. This assured that the segment of the r function that applies if the constraints $k_1(x) \geq 0, k_2(x) \geq 0$ are not binding, which we denoted by r_0, satisfies the requirement concerning the slope of the $r(x)$ function, (11).

While many frequently used distributions are log-concave (see Bagnoli and Bergstrom, 2005), it is not difficult to find distributions that exhibit log-convexity and even $\left(G(x)/G'(x)\right)' \leq -2/3$ in subsets of the support.[12] Simple examples are obtained for the Kumaraswamy distribution function: $G(x) := 1 - \left(1 - \left((x-\underline{x})/(\bar{x}-\underline{x})\right)^a\right)^b$, introduced by Kumaraswamy (1980), which is log-concave if and only if $a \geq 1, b \geq 1$ (see Jones, 2009), and log-convex in a subset of the support, for example, for $a = b = 3/4$.

If G exhibits log-convexity somewhere, the optimal mechanism may exhibit an r function that is strictly decreasing in a subset of $[x_1, x_2]$. In that case the optimal mechanism under incomplete information differs more radically from that under complete information.

[12] However, the distribution function of a non-negative random variable cannot be log-convex everywhere (Block, Savits, and Singh, 1998, Corollary 2.1).

Finally, we mention that pure royalty contracts are optimal in the special case when the innovator is at least as efficient as the licensee with probability one:[13]

Proposition 3. *Suppose the innovator is more efficient than the licensee with probability one, i.e., $\underline{x} \geq d$. The optimal licensing mechanism is a pure royalty contract.*

Proof. By Proposition 1 it is sufficient to show that $x_2 \leq \underline{x}$. This holds true if and only if $k_2(\underline{x})|_{\eta=0} \leq 0$, which confirms as follows:

$$\begin{aligned} k_2(\underline{x})|_{\eta=0} &= \frac{1}{4}\big(2(2c-1) - 9\underline{x} - 3\hat{x} + 10d\big) \\ &\leq \frac{1}{4}\left(12d - 9\underline{x} - 3\hat{x}\right) \quad \text{(because } d \geq 2c-1 \text{ (non-drastic))} \\ &\leq 0 \quad \text{(because } d \leq \underline{x} \leq \hat{x}\text{).} \end{aligned}$$

□

Note that, in this case, the optimal $r(x)$ function is determined by the boundary of the constraint, $k_2(x) = 0$; optimization plays only a role in the determination of the exclusion threshold level, $\hat{x}$.

5 Extension to (super-)drastic innovations

We now assume the innovation is drastic for the innovator, i.e., $d \leq 2c - 1$.

If the licensee is less efficient (conditional on using the innovation), i.e., if $d \leq \underline{x}$, the innovator will simply not license the innovation and make himself a monopoly. Licensing is only an issue if the licensee is more efficient with positive probability.

In the following we assume that the innovation is drastic and the licensee is more efficient (after licensing) with positive probability. We will characterize the optimal mechanism that makes either the innovator or the potential licensee a monopoly.

In the previous section we argued that the sale of the patent is not likely to be feasible because the antitrust authorities view it as a collusive scheme that transforms a duopoly market into a monopoly. However, this does not apply if the innovation is drastic for the innovator, because in that case the innovation already gives rise to a monopoly if the patent is not sold. If the sale of the patent induces a monopoly, it only replaces an inefficient monopoly by an efficient one. Therefore, the sale of the patent has no anticompetitive effect and is generally permitted.[14]

Suppose the innovator puts his patent up for sale and offers a sales contract $(f(x), r(x), \gamma(x))$ that induces a monopoly, to which we will refer as "patent sale mechanism". Because the innovation is drastic for the innovator, he becomes a monopoly if no sale takes place. However, even though the licensee is more efficient after using the innovation, the mechanism has to assure that if a sale takes place the licensee becomes a monopoly, i.e., firm 1 does not have an incentive to produce. This requires that $r(x)$ must be such that $q_1^*(x) = 0$, i.e,

$$r(x) \leq 2c - 1 - x \quad \text{(monopoly constraint).} \tag{18}$$

[13] This result generalizes Heywood, Li, and Ye (2014) who use a binary model and assume that the innovator is at least as efficient as the potential licensee with probability one.

[14] According to the Antitrust Guidelines: "If the Agencies conclude ... that a restraint in a licensing arrangement is unlikely to have an anticompetitive effect, they will not challenge the restraint" (U.S. Department of Justice and Federal Trade Commission, 1995, Section 4.2).

Lemma 2. *The following conditions are necessary for incentive compatibility of the patent sale mechanism, for all $x, z \leq \hat{x} \leq \bar{x}$:*

$$\gamma(x) = \begin{cases} 1 & \text{if} \quad x \leq \hat{x} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

$$f'(x) = -\frac{1}{2}\left(1 - x - r(x)\right) r'(x) \tag{20}$$

$$\frac{r(x) - r(z)}{x - z} \geq 0. \tag{21}$$

Proof. The proof of (19) is the same as in Lemma 1.

The equilibrium profit of the licensee with cost x and reported cost $z \leq \hat{x}$ is:

$$\Pi_2^L(x,z) = \left(\frac{1 - x - r(z)}{2}\right)^2 - f(z).$$

Incentive compatibility requires $\partial_z \, \Pi_2^L(x,z)\big|_{z=x} = 0$ which implies (20).

Moreover, incentive compatibility requires that r is non-decreasing, as stated in (21), because

$$\Pi_2(x,x) - \Pi_2(x,z) + \Pi_2(z,z) - \Pi_2(z,x) \geq 0 \Rightarrow \frac{1}{2}(x - z)(r(x) - r(z)) \geq 0.$$

□

Proposition 4. *Suppose the licensee is more efficient with positive probability, i.e., $\underline{x} < d$, and the innovation is drastic, i.e., $d \leq 2c - 1$. The optimal patent sale mechanism prescribes:*

$$\gamma(x) = 1 \iff x \leq \hat{x}, \quad r(x) = \frac{G(x)}{G'(x)}, \quad f(x) = q_2^M(\hat{x})^2 + \int_x^{\hat{x}} q_2^M(y) \left(\frac{G(y)}{G'(y)}\right)' dy, \tag{22}$$

where $q_2^M(x)$ denotes the monopoly output of firm 2. It exhibits a strictly increasing r function, a strictly decreasing f function, and induces firm 2 to be a monopoly when $\gamma(x) = 1$. The optimal exclusion threshold, $\hat{x}$, depends on the parameters and the distribution function.

If G is the uniform distribution with support $[0, \bar{x}]$, one has $r(x) = x$, $f(x) = (1-2\hat{x}(1-\hat{x})-2x(1-x))/4$, and $\hat{x} = d/2$ if $d \leq (3-2\bar{x})/2$ and $\hat{x} = \bar{x}$ otherwise, which implies exclusion ($\hat{x} < \bar{x}$) if $d \leq (3-2\bar{x})/2$ and $\bar{x} > d/2$.

Proof. The first-order conditions for incentive compatibility, (20), imply that $\Pi_2(x,x)$ is monotone decreasing in x. Therefore, the participation constraint can be replaced by the constraint:

$$\Pi_2^L(\hat{x}, \hat{x}) \geq 0. \tag{23}$$

By (21) r is non-decreasing. Therefore, the antitrust constraints can be binding at most once, at $x = \hat{x}$. Similarly, the monopoly constraint, (18), can be binding at most once, at $x = \hat{x}$. This allows us to replace these two constraints by the constraint:

$$r(\hat{x}) + \hat{x} \leq \min\{c, 2c - 1\}. \tag{24}$$

It is obviously optimal to adopt the highest possible fixed fee, which allows us to replace the participation constraints by

$$f(\hat{x}) = \left(\frac{1-\hat{x}-r(\hat{x})}{2}\right)^2. \tag{25}$$

Using the incentive compatible allocation rule (19), the optimal patent sale mechanism must solve the following maximization problem (where π_1^M denotes the monopoly profit of firm 1):

$$\max_{r,f,\hat{x}} \Pi_1 = \int_{\underline{x}}^{\hat{x}} \left(r(x)q_2^M(x) + f(x)\right) G'(x)dx + (1-G(\hat{x}))\,\pi_1^M, \quad \text{s.t.} \quad (20), (24), (25). \tag{P2}$$

We first take $\hat{x}$ as given and ignore constraint (24) and the monotonicity constraint (21), which, as we will confirm, are not binding. We solve the thus reduced program as an optimal control problem, and finally characterize the optimal $\hat{x}$, which can be affected by the constraint (24).

The control variable is $u(x) := r'(x)$, the state variables are $r(x), f(x)$, the constraint functions are g_1, g_2, the co-state variables, $\lambda_1(x), \lambda_2(x)$, and the Hamiltonian is $H(x,r,f,u,\lambda_1,\lambda_2)$:

$$H(x,r,f,u,\lambda_1,\lambda_2) := \left(r(x)q_2^M(x) + f(x)\right) G'(x) + \lambda_1 g_1 + \lambda_2 g_2$$
$$q_2^M(x) := \frac{1}{2}(1-x-r(x)), \quad g_1 := u(x), \quad g_2 := -\frac{1}{2}(1-x-r(x))\,u(x).$$

The endpoints $f(\hat{x})$, $r(\hat{x})$ are not free (by (25)), and the endpoints $f(\underline{x})$, $r(\underline{x})$ are free.

The maximum principle requires that the following conditions are satisfied:

$$0 = \partial_u H = \lambda_1(x) - \frac{\lambda_2(x)}{2}(1-x-r(x)) \tag{26}$$
$$\lambda_1'(x) = -\partial_r H = \frac{1}{2}\left((x-1+2r(x))G'(x) - \lambda_2(x)u(x)\right) \tag{27}$$
$$\lambda_2'(x) = -\partial_f H = -G'(x) \tag{28}$$
$$r'(x) = \partial_{\lambda_1} H = u(x) \tag{29}$$
$$f'(x) = \partial_{\lambda_2} H = -\frac{1}{2}(1-x-r(x))\,u(x) \tag{30}$$
$$f(\hat{x}) = \left(\frac{1-\hat{x}-r(\hat{x})}{2}\right)^2 \tag{31}$$
$$\lambda_1(\underline{x}) = \lambda_2(\underline{x}) = 0. \tag{32}$$

(28) and (32) yield: $\lambda_2(x) = -G(x)$; inserting this into (26) yields

$$\lambda_1(x) = -\frac{G(x)}{2}(1-x-r(x)). \tag{33}$$

Differentiating (33) and equating with (27), after inserting $\lambda_2(x)$, yields the asserted r function.

To find $f(x)$ we use the fact that $f(x) \equiv f(\hat{x}) - \int_x^{\hat{x}} f'(y)dy$ together with (30) and (31) and find the asserted f function.

Evidently, the assumed log-concavity of G assures that r is strictly increasing and f strictly decreasing.

We now confirm that incentive compatibility is satisfied and then characterize the optimal $\hat{x}$.

Because $r'(x) > 0$ one has $\partial_{zx}\Pi_2(x,z) = r'(z)/2 > 0$. Therefore, $\Pi_2(x,z)$ is pseudoconcave in z, which assures that the solution is incentive compatible.

The Arrow Sufficiency Theorem (Seierstad and Sydstæter, 1977, Theorem 3) applies, because, at the optimum, $\lambda_1 g_1 + \lambda_2 g_2 = 0$, and $H^* := \max_u H$ is concave in r and f.

Define ξ as the unique solution of the equation

$$x + r(x) = \min\{c, 2c-1\}, \tag{34}$$

and denote $\Pi_1^*(\hat{x}) = \max_{r,f}\Pi_1$. The optimal $\hat{x}$, denoted by $\hat{x}^*$, maximizes $\Pi_1^*(\hat{x})$ over the constraint set: $[\underline{x}, \min\{\xi, \bar{x}\}]$. If G is the uniform distribution with support $[0,\bar{x}]$, the innovator's expected profit is equal to: $\Pi_1^*(\hat{x}) = \big(3\bar{x}(1-d)^2 + \hat{x}\big(6d - 3d^2 + 2\hat{x}(2\hat{x}-3)\big)\big)/(12\bar{x})$. Maximizing Π_1 yields the optimal exclusion threshold level $\hat{x} = \min\{d/2, \bar{x}\}$ because $\Pi_1^*(\hat{x})$ has a local maximum at $\hat{x} = d/2$ which is also the global maximum over $[0,\xi]$. □

Evidently, it is not optimal to sell the patent for a pure cash price. Instead, it is optimal to supplement the cash price by output based royalties. This can be interpreted as an implication of the "linkage principle". According to that principle, linking the license fee to a variable that is correlated with the licensee's private information, such as output, tends to lower information rents (as shown in a different context by Milgrom, 1987).

6 Discussion

The present paper considers the inside innovator's optimal licensing of non-drastic and drastic innovations under incomplete information. We consider exclusive and non-exclusive licenses, taking into account the conditions under which exclusive licenses are compatible with antitrust rules. Due to the nature of the licensing problem and its constraints, we analyze optimal licensing as solutions of optimal control problems. Our results indicate that optimal licensing typically exhibits a complex pattern of state dependent royalty rates and fixed fees.

In the case of non-drastic innovations we find that the optimal mechanism generally exhibits four regimes: a two-part tariff regime with a decreasing royalty rates and a positive constant fixed fees, a two-part tariff regime with increasing royalty rates and decreasing fixed fees, a pure royalty regime with decreasing royalty rates and zero fixed fees, and an exclusion regime where no license is awarded. The optimal mechanism always includes state dependent royalty rates, because a mechanism that includes only state dependent fixed fees cannot be incentive compatible. The pure royalty mechanism is optimal if the innovator's unit cost induced by the innovation is never greater than that of the potential licensee.

In the case of drastic innovations licensing occurs only if the licensee is more efficient conditional on using the innovation with positive probability; otherwise, the innovator will not license the innovation and enjoy the benefits of being a monopoly. We consider patent sale mechanisms in which the innovator either issues an exclusive license by selling the patent, choosing the royalty rate in such a way that the licensee becomes a monopoly, or does not issue a license and himself becomes a monopoly.

We show that the optimal patent sale mechanism exhibits increasing royalty rates and decreasing fixed fees; exclusion tends to occur, in which case the innovator becomes a monopoly. The optimal patent sale mechanism always includes state dependent royalty rates. Unlike in the case of non-drastic innovations, this property is not driven by incentive compatibility concerns. Instead, it can be interpreted as an implication of the fact that linking the license fee to a variable that is correlated with the licensee's private information tends to lower information rents.

The present analysis has covered innovations that are either drastic or non-drastic with probability one. In future research, one may wish to extend the analysis to innovations that can be either drastic and non-drastic with positive probability. On should also extend the analysis to multiple potential licensees, which raises the issue of the optimal number of licenses.

A Appendix

A.1 Proof of Proposition 1

As a preliminary note that it is optimal to adopt the highest possible fixed fee, which allows us to replace the participation constraints by

$$f(\hat{x}) = \pi_2(q_1^*(\hat{x}), q_2^*(\hat{x}); \hat{x} + r(\hat{x})) - \Pi_2^N. \tag{35}$$

Step 1: Consider the relaxed maximization problem (relaxed by ignoring the constraint $r'(x) \geq -1/4$):

$$\begin{aligned} \max_{r,f,\hat{x}} \Pi_1 = & \int_{\underline{x}}^{\hat{x}} (\pi_1^*(x) + r(x)q_2^*(x) + f(x))\, dG(x) + \pi_1\left(q_1^N, q_2^N; d\right)(1 - G(\hat{x})) \\ \text{s.t.} \quad & (14), (15), (35). \end{aligned} \tag{P1}$$

We solve (P1) as an optimal control problem for given $\hat{x}$ and then characterize the optimal $\hat{x}$ and confirm that the ignored constraint is never binding.

To translate (P1) into an optimal control problem define the control variable, $u(x) := r'(x)$, the state variables $r(x), f(x)$, the constraint functions g_1, g_2, k_1, k_2, the co-state variables, $\lambda_1(x)$, $\lambda_2(x)$, $\eta_1(x)$, $\eta_2(x)$ and the Hamiltonian, $H(x, r, f, u, \lambda_1, \lambda_2, \eta_1, \eta_2)$:

$$\begin{aligned} H(x,r,f,u,\lambda_1,\lambda_2,\eta_1,\eta_2) &:= h(x,r,f) + \lambda_1(x)g_1 + \lambda_2(x)g_2 + \eta_1(x)k_1 + \eta_2(x)k_2 \\ h &:= \left(\left(\frac{1-2d+x+r(x)}{3} \right)^2 + r(x)\frac{1-2x-2r(x)+d}{3} + f(x) \right) G'(x) \\ g_1 &:= u(x), \quad g_2 := -\frac{1}{9}(1+d-2x-2r(x))(1+4u(x)) \\ k_1 &:= r(x) + \frac{1}{4}(x+3\underline{x}) - \max\{2d-1, 0\}, \quad k_2 := c - \frac{3\hat{x}+x}{4} - r(x). \end{aligned}$$

There, g_2 corresponds to the incentive compatibility conditions (10), and k_1, k_2 to the constraints, (14), (15). The endpoints $f(\hat{x})$ and $r(\hat{x})$ are not free (see (35)) and the endpoints $f(\underline{x})$, $r(\underline{x})$ are free.

The maximum principle requires that the following conditions are satisfied:[15]

$$
\begin{aligned}
0 = \partial_u H &= \lambda_1(x) - \frac{4\lambda_2(x)}{9}(1+d-2x-2r(x)) && (36)\\
\lambda_1'(x) = -\partial_r H &= -\eta_1 + \eta_2 - \frac{2\lambda_2(x)(1+4u(x)) + (5-d-4x-10r(x))G'(x)}{9} && (37)\\
\lambda_2'(x) = -\partial_f H &= -G'(x) && (38)\\
r'(x) = \partial_{\lambda_1} H &= u(x) && (39)\\
f'(x) = \partial_{\lambda_2} H &= -\frac{1}{9}(1+d-2x-2r(x))(1+4u(x)) && (40)\\
f(\hat{x}) &= \left(\frac{1-2(\hat{x}+r(\hat{x}))+d}{3}\right)^2 - \left(\frac{1-2c+d}{3}\right)^2 && (41)\\
\lambda_1(\underline{x}) &= \lambda_2(\underline{x}) = 0 && (42)\\
\eta_1(x) \geq 0, \quad k_1 &= r(x) + \frac{1}{4}(x+3\underline{x}) - \max\{2d-1,0\} \geq 0, \quad \eta_1(x)k_1 = 0. && (43)\\
\eta_2(x) \geq 0, \quad k_2 &= c - \frac{3\hat{x}+x}{4} - r(x) \geq 0, \quad \eta_2(x)k_2 = 0, && (44)
\end{aligned}
$$

Step 2: Deriving $r(x)$.
(38) and (42) yield $\lambda_2(x) = -G(x)$. Inserting this into (36) yields

$$\lambda_1(x) = -\frac{4G(x)}{9}(1+d-2x-2r(x)). \tag{45}$$

Differentiating (45) and equating with (37) yields:

$$r(x) = \frac{1-5d+4x}{2} + \frac{6G(x)+9(\eta_1-\eta_2)}{2G'(x)}. \tag{46}$$

Define $\eta := (\eta_1, \eta_2)$ and:

$$r_0(x) := r(x)|_{\eta=0} = \frac{1-5d+4x}{2} + \frac{3G(x)}{G'(x)}. \tag{47}$$

Consider the function $k_1(x)|_{\eta=0} = r_0(x) + \frac{1}{4}(x+3\underline{x}) - \max\{2d-1,0\}$. Because $k_1'(x)|_{\eta=0} \geq 9/4$, the equation $k_1(x)|_{\eta=0} = 0$ has a unique solution denoted by x_1 and:[16]

$$k_1(x)|_{\eta=0} \gtreqless 0 \iff x \gtreqless x_1. \tag{48}$$

Similarly, consider the function $k_2(x)|_{\eta=0} = c - \frac{3\hat{x}+x}{4} - r_0(x)$. Because $k_2'(x)|_{\eta=0} \leq -9/4$, the equation $k_2(x)|_{\eta=0} = 0$ has a unique solution denoted by x_2 (which may be above $\hat{x}$ or even $\bar{x}$) and:

$$k_2(x)|_{\eta=0} \gtreqless 0 \iff x \lesseqgtr x_2. \tag{49}$$

(43) and (44) imply that $\hat{x}$ must be chosen to satisfy

$$\hat{x} \leq \underline{x} + \frac{4}{3}(c - \max\{2d-1,0\}). \tag{50}$$

[15] The present control problem is subject to inequality constraints concerning the state variables and some free and fixed end-points. The corresponding conditions for optimality can be found in Kamien and Schwartz (1991, p. 230 ff.).

[16] We extend G outside its support in the obvious way. Therefore the equation has a solution, which may however be below $\underline{x}$.

By the definition of x_1 we have $r_0(x_1) = -\frac{1}{4}(x_1 + 3\underline{x}) + \max\{2d-1,0\}$; hence

$$k_2(x_1)|_{\eta=0} = c - \frac{3\hat{x}+x_1}{4} - r_0(x_1) = c - \frac{3(\hat{x}-\underline{x})}{4} - \max\{2d-1,0\} \geq 0 \quad \text{(by (50))}$$

which implies $x_1 \leq x_2$ with equality if and only if $\hat{x} = \underline{x} + \frac{4}{3}(c - \max\{2d-1,0\})$.

Also we have

$$k_1(x) = k_1(x)|_{\eta=0} + \frac{9(\eta_1 - \eta_2)}{2G'(x)} \tag{51}$$

$$k_2(x) = k_2(x)|_{\eta=0} - \frac{9(\eta_1 - \eta_2)}{2G'(x)}. \tag{52}$$

We now derive the asserted $r(x)$ function.

$$\begin{aligned}
x < x_1 &\Rightarrow k_1(x)|_{\eta=0} < 0 \Rightarrow k_1(x) = 0 \text{ (because } k_1(x) > 0 \Rightarrow \eta_1 = 0 \Rightarrow k_1(x) < 0)\\
&\Rightarrow r(x) = -\frac{1}{4}(x + 3\underline{x}) + \max\{2d-1,0\} = r_1(x)\\
x = x_1 &\Rightarrow k_1(x)|_{\eta=0} = 0 \text{ and } \eta_1 - \eta_2 \geq 0 \Rightarrow k_1\eta_1 = 9(\eta_1-\eta_2)\eta_1/(2G'(x)) = 0\\
&\Rightarrow \eta_1 - \eta_2 = 0, k_1(x) = 0 \Rightarrow r(x) = -\frac{1}{4}(x + 3\underline{x}) + \max\{2d-1,0\} = r_1(x)\\
x_1 < x < x_2 &\Rightarrow k_1(x)|_{\eta=0} > 0, k_2(x)|_{\eta=0} > 0\\
&\Rightarrow k_1(x) > 0 \text{ (because } k_1 = 0 \Rightarrow k_2 > 0 \Rightarrow \eta_2 = 0 \Rightarrow \eta_1 < 0\text{) and similarly } k_2(x) > 0\\
&\Rightarrow \eta = 0 \Rightarrow r(x) = r_0(x)\\
x = x_2 &\Rightarrow k_2(x)|_{\eta=0} = 0 \text{ and } \eta_1 - \eta_2 \leq 0 \Rightarrow k_2\eta_2 = -9(\eta_1-\eta_2)\eta_2/(2G'(x)) = 0\\
&\Rightarrow \eta_1 - \eta_2 = 0, k_2(x) = 0 \Rightarrow r(x) = c - \frac{3\hat{x}+x}{4} = r_2(x)\\
x > x_2 &\Rightarrow k_2(x)|_{\eta=0} < 0 \Rightarrow k_2(x) = 0 \text{ (because } k_2(x) > 0 \Rightarrow \eta_2 = 0 \Rightarrow k_2(x) < 0)\\
&\Rightarrow r(x) = c - \frac{3\hat{x}+x}{4} = r_2(x).
\end{aligned}$$

$r(x)$ is decreasing on $[\underline{x}, x_1]$, increasing on $[x_1, x_2]$, which confirms that the ignored constraint is not binding there, and decreasing on $[x_2, \hat{x}]$, as asserted.

Step 3: Deriving $f(x)$.

Denote the unique solution to the equation $c - x = r_0(x)$ by x_0, then we have

$$\hat{x} \gtreqless x_0 \iff k_2(x_0)|_{\eta=0} = \frac{3}{4}(x_0 - \hat{x}) \lesseqgtr 0 \iff x_0 \gtreqless x_2.$$

We distinguish three cases:

Case 1: $\hat{x} \geq x_0$. In this case we have $x_1 \leq x_2 \leq \hat{x}$ and $r(\hat{x}) = c - \hat{x}$. Thus $f(\hat{x}) = 0$.

$$
\begin{aligned}
x \in [x_2, \hat{x}] &\Rightarrow 1 + 4r'(x) = 0 \Rightarrow f'(x) = 0 \Rightarrow f(x) = f(\hat{x}) - \int_x^{\hat{x}} f'(y)dy = f(\hat{x}) = 0 \\
x \in [x_1, x_2] &\Rightarrow f'(x) = -\frac{1}{9}(1 + d - 2x - 2r_0(x))\left(1 + 4r_0'(x)\right) = -q_2^*(x)\left(3 + 4\left(\frac{G(x)}{G'(x)}\right)'\right) \\
&\Rightarrow f(x) = f(x_2) + \int_x^{x_2} q_2^*(y)\left(3 + 4\left(\frac{G(y)}{G'(y)}\right)'\right)dy = \int_x^{x_2} q_2^*(y)\left(3 + 4\left(\frac{G(y)}{G'(y)}\right)'\right)dy \\
x \in [\underline{x}, x_1] &\Rightarrow 1 + 4r'(x) = 0 \Rightarrow f'(x) = 0 \\
&\Rightarrow f(x) = f(x_1) - \int_x^{x_1} f'(y)dy = f(x_1) = \int_{x_1}^{x_2} q_2^*(y)\left(3 + 4\left(\frac{G(y)}{G'(y)}\right)'\right)dy.
\end{aligned}
$$

In this case $f(x)$ is positive and constant on $[\underline{x}, x_1]$, decreasing on $[x_1, x_2]$, and equal to zero on $[x_2, \hat{x}]$.

Case 2: $\hat{x} < x_0$ and $\hat{x} > x_1$. In this case we have $x_1 < \hat{x} < x_2$. Thus, $k_2(\hat{x}) > 0$, which implies $r(\hat{x}) < c - \hat{x}$, and $f(\hat{x}) = \pi_2(q_1^*(\hat{x}), q_2^*(\hat{x}); \hat{x} + r(\hat{x})) - \Pi_2^N > 0$.

$$
\begin{aligned}
x \in [x_1, \hat{x}) &\Rightarrow f'(x) = -\frac{1}{9}(1 + d - 2x - 2r_0(x))\left(1 + 4r_0'(x)\right) = -q_2^*(x)\left(3 + 4\left(\frac{G(x)}{G'(x)}\right)'\right) \\
&\Rightarrow f(x) = f(\hat{x}) + \int_x^{\hat{x}} q_2^*(y)\left(3 + 4\left(\frac{G(y)}{G'(y)}\right)'\right)dy \\
x \in [\underline{x}, x_1] &\Rightarrow 1 + 4r'(x) = 0 \Rightarrow f'(x) = 0 \\
&\Rightarrow f(x) = f(x_1) - \int_x^{x_1} f'(y)dy = f(x_1) = f(\hat{x}) + \int_{x_1}^{\hat{x}} q_2^*(y)\left(3 + 4\left(\frac{G(y)}{G'(y)}\right)'\right)dy.
\end{aligned}
$$

In this case $f(x)$ is positive and constant on $[\underline{x}, x_1]$ and decreasing on $[x_1, \hat{x}]$.

Case 3: $\hat{x} < x_0$ and $\hat{x} \leq x_1$ ($\leq x_2$). In this case we have $f(\hat{x}) = \pi_2(q_1^*(\hat{x}), q_2^*(\hat{x}); \hat{x} + r(\hat{x})) - \Pi_2^N$.

$$
x \in [\underline{x}, \hat{x}] \Rightarrow 1 + 4r'(x) = 0 \Rightarrow f'(x) = 0 \Rightarrow f(x) = f(\hat{x}) - \int_x^{\hat{x}} f'(y)dy = f(\hat{x}).
$$

In this case $f(x)$ is positive and constant.

Step 4: Global incentive compatibility and second-order conditions.
Next, we confirm incentive compatibility.[17]

(a) For $x, z \geq x_2$, we have $\Pi_2(x,x) = \Pi_2(x,z) = \frac{1}{36}(2 - 4c + 2d + 3\hat{x} - 3x)^2$.

(b) For $x, z \in [x_1, x_2]$, we have $\partial_{x,z}\Pi_2(x,z) = \frac{1}{6}(1 + 4r'(z)) > 0$.

(c) For $x_1 \leq z < x_2 \leq x$, we have $\Pi_2(x_2, x_2) - \Pi_2(x_2, z) \geq 0$ from (b) and

$$
(\Pi_2(x,x) - \Pi_2(x,z)) - (\Pi_2(x_2,x_2) - \Pi_2(x_2,z)) = \frac{2}{3}(x - x_2)k_2(z) \geq 0,
$$

which implies $\Pi_2(x,x) - \Pi_2(x,z) \geq 0$.

[17] In one case we prove incentive compatibility by showing that $\Pi_2(x,z)$ is pseudoconcave in z. The function $\Pi_2(x,z)$ is "pseudoconcave" in z if, for all x, it is increasing to the left of its stationary point and decreasing to the right. Due to the first-order condition for incentive compatibility $\Pi_2(x,z)$ has a stationary point at $z = x$, i.e., $\partial_z\Pi_2(x,z)|_{z=x} = 0$ for all x. Evidently, pseudoconcavity assures that the stationary points are global maxima.

(d) For $x_1 \le x < x_2 \le z$, we have $\Pi_2(x,x) \ge \Pi_2(x,x_2) = \Pi_2(x,z)$.

(e) For $x,z \le x_1$, we have $\Pi_2(x,x) = \frac{1}{36}(2+2d+3\underline{x}-3x)^2 - f(x_1) = \Pi_2(x,z)$.

(f) For $x \le x_1 < z \le x_2$, we have $\Pi_2(x_1,x_1) - \Pi_2(x_1,z) \ge 0$ from (b) and

$$(\Pi_2(x,x) - \Pi_2(x,z)) - (\Pi_2(x_1,x_1) - \Pi_2(x_1,z)) = \frac{2}{3}(x_1 - x)k_1(z) \ge 0,$$

which implies $\Pi_2(x,x) - \Pi_2(x,z) \ge 0$.

(g) For $z < x_1 \le x \le x_2$, we have $\Pi_2(x,x) \ge \Pi_2(x,x_1) = \Pi_2(x,z)$.

(h) For $x \le x_1, z \ge x_2$, we have $\Pi_2(x_1,x_1) - \Pi_2(x_1,z) \ge 0$ from (d) and

$$(\Pi_2(x,x) - \Pi_2(x,z)) - (\Pi_2(x_1,x_1) - \Pi_2(x_1,z)) = \frac{1}{6}(3\underline{x} + 4c - 3\hat{x})(x_1 - x) \ge 0,$$

which implies $\Pi_2(x,x) - \Pi_2(x,z) \ge 0$.

(i) For $x \ge x_2, z \le x_1$, we have $\Pi_2(x_2,x_2) - \Pi_2(x_2,z) \ge 0$ from (g) and

$$(\Pi_2(x,x) - \Pi_2(x,z)) - (\Pi_2(x_2,x_2) - \Pi_2(x_2,z)) = \frac{1}{6}(3\underline{x} + 4c - 3\hat{x})(x - x_2) \ge 0,$$

which implies $\Pi_2(x,x) - \Pi_2(x,z) \ge 0$.

Therefore, incentive compatibility is satisfied.

The Arrow Sufficiency Theorem (Seierstad and Sydstæter, 1977, Theorem 3) applies, because, at the optimum, $\lambda_1 g_1 + \lambda_2 g_2 = 0$ and $H^* := \max_u H$ is concave in r and f.

Step 5: $\hat{x} > \underline{x}$.
Denote the maximum expected payoff of the innovator by $\Pi_1^*(\hat{x})$. We now prove that a contract is awarded with positive probability by showing

$$\Pi_1^{*\prime}(\hat{x})\big|_{\hat{x}=\underline{x}} = \Big(\pi_1^*(\underline{x}) + r(\underline{x})q_2^*(\underline{x}) + f(\underline{x}) - \pi_1\left(q_1^N, q_2^N; d\right)\Big)G'(\underline{x}) > 0.$$

We consider three cases: (i) $\underline{x} > x_2$, (ii) $\underline{x} \in [x_1, x_2)$, (iii) $\underline{x} < x_1$.

Case (i): If $\underline{x} \ge x_2$, then $r(x) = r_2(x)$ and we have

$$\Pi_1^{*\prime}(\hat{x})\big|_{\hat{x}=\underline{x}} = \frac{1}{3}(1-2c+d)(c-\underline{x})G'(\underline{x}) > 0, \quad \text{because } d > 2c-1 \text{ (non-drastic).}$$

Case (ii): If $\underline{x} \in [x_1, x_2)$, then $r(x) = r_0(x)$ and we have

$$\Pi_1^{*\prime}(\hat{x})\big|_{\hat{x}=\underline{x}} = \Big((d-\underline{x})^2 + \frac{1}{36}\left(1+8c-20c^2-(10-32c)d-11d^2\right)\Big)G'(\underline{x}) > 0,$$

because the second term in the big bracket is a concave function of d which takes the minimum value 0 at $d = 2c-1$ on $[2c-1, c]$.

Case (iii): If $\underline{x} < x_1$, then $r(x) = r_1(x)$ and we have

$$\Pi_1^{*\prime}(\hat{x})\big|_{\hat{x}=\underline{x}} = \frac{1}{9}\Big(c(2-5c+8d) - 3\underline{x}(1+d) + (1+6a-5d-\max\{2d-1,0\})\max\{2d-1,0\}\Big)G'(\underline{x}).$$

If $2d-1 \leq 0$, Then

$$\begin{aligned}
\Pi_1^{*\prime}(\hat{x})\big|_{\hat{x}=\underline{x}} &= \frac{1}{9}\Big(c(2-5c+8d)-3\underline{x}(1+d)\Big)G'(\underline{x}) \\
&> \frac{1}{9}\Big(c(2-5c+8d)+\frac{1}{2}(1-5d)(1+d)\Big)G'(\underline{x}), \quad \text{because } k_1(\underline{x})<0 \Leftrightarrow \underline{x}<\frac{5d-1}{6} \\
&= \frac{1}{9}\Big(\frac{1}{2}+2c-5c^2-2(1-4c)d-\frac{5d^2}{2}\Big)G'(\underline{x})>0,
\end{aligned}$$

because the function in the big bracket of the last line is a concave function of d which takes the minimum value c^2 at $d=2c-1$ on $[2c-1,c]$.

If $2d-1>0$, Then

$$\begin{aligned}
\Pi_1^{*\prime}(\hat{x})\big|_{\hat{x}=\underline{x}} &= \frac{1}{9}\Big(c(2-5c+8d)-9\underline{x}(1-d)-2+11d-14d^2\Big)G'(\underline{x}) \\
&> \frac{1}{9}\Big(c(2-5c+8d)-\frac{9}{2}(3d-1)(1-d)-2+11d-14d^2\Big)G'(\underline{x}), \\
&\quad \text{because } k_1(\underline{x})<0 \Leftrightarrow \underline{x}<\frac{3d-1}{2} \\
&= \frac{1}{9}\Big(\frac{5}{2}+2c-5c^2-(7-8c)d-\frac{d^2}{2}\Big)G'(\underline{x})>0,
\end{aligned}$$

because the function in the big bracket of the last line is a concave function of d which takes the minimum value $\frac{5}{18}(1-c)^2$ at $d=c$ on $[2c-1,c]$.

Depending upon the parameters, the optimal $\hat{x}$ is either a corner solution, $\hat{x}=\bar{x}$ (no-exclusion), or an interior solution (exclusion).

A.2 Supplement to the proof of Proposition 2

Here we supplement the proof of Proposition 2 and show that $\Pi_1^*(\hat{x})$ increases on $[x_1,x_2)$.

On the interval $[x_1,x_2)$ one has $r(x)=r_0(x)$. First, suppose $x_1<\underline{x}$. Then, the innovator's expected revenue can be written as

$$\Pi_1^*(\hat{x})=\int_{\underline{x}}^{\hat{x}}\Big(r_0(y)q_2^*(y)+f(y)+\pi_1^*(y)\Big)dG(y)+(1-G(\hat{x}))\Pi_1^N.$$

Its derivative with respect to $\hat{x}$ is

$$\begin{aligned}
\Pi_1^{*\prime}(\hat{x}) &= \frac{1}{36G'(\hat{x})}\Big(36G(\hat{x})^2-72(d-\hat{x})G(\hat{x})G'(\hat{x}) \\
&\quad +\big(1-20c^2+25d^2+8c(1+4d)+36\hat{x}^2-2d(5+36\hat{x})\big)G'(\hat{x})^2\Big) \\
&= \frac{1}{36G'(\hat{x})}\Big(36\big(G(\hat{x})-(d-\hat{x})G'(\hat{x})\big)^2+(1+10c-11d)(d-2c+1)G'(\hat{x})^2\Big),
\end{aligned}$$

which is positive because $d>2c-1$.

Next,suppose $x_1 \geq \underline{x}$. Then, $r(x)=r_1(x)$ on $[\underline{x},x_1]$, and Π_1 can be written as

$$\begin{aligned}
\Pi_1^*(\hat{x}) &= \int_{\underline{x}}^{x_1}\Big(r_1(y)q_2^*(y)+f(x_1)+\pi_1^*(y)\Big)dG(y) \\
&\quad +\int_{x_1}^{\hat{x}}\Big(r_0(y)q_2^*(y)+f(y)+\pi_1^*(y)\Big)dG(y)+(1-G(\hat{x}))\Pi_1^N.
\end{aligned}$$

The derivative of $\Pi_1^*(\hat{x})$ with respect to $\hat{x}$ is exactly the same as above, which completes the proof.

References

Bagnoli, M. and T. Bergstrom (2005). "Log-Concave Probability and its Applications". *Economic Theory* 26, pp. 445–469.

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

Block, H., T. Savits, and H. Singh (1998). "The Reversed Hazard Rate Function". *Probability in the Engineering and Informational Sciences* 12, pp. 69–90.

Brousseau, E., R. Coeurderoy, and C. Chaserant (2007). "The Governance of Contracts: Empirical Evidence on Technology Licensing Agreements". *Journal of Institutional and Theoretical Economics* 163, pp. 205–235.

Calvert, R. (1964). *The Encyclopedia of Patent Practice and Invention Management*. Reinhold, New York.

Choi, J.-P. (2001). "Technology Transfer with Moral Hazard". *International Journal of Industrial Organization* 19, 249266.

Fan, C., B. Jun, and E. Wolfstetter (2013). "Licensing Process Innovations when Losers' Messages Determine Royalty Rates". *Games and Economic Behavior* 82, pp. 388–402.

– (2014). "Licensing a Common Value Innovation in Oligopoly when Signaling Strength may Backfire". *International Journal of Game Theory* 43, pp. 215–244.

Gallini, N. and B. Wright (1990). "Technology Transfer under Asymmetric Information". *RAND Journal of Economics* 21.1, pp. 147–160.

Giebe, T. and E. Wolfstetter (2008). "License Auctions with Royalty Contracts for (Winners and) Losers". *Games and Economic Behavior* 63, pp. 91–106.

Heywood, J., J. Li, and G. Ye (2014). "Per Unit vs. ad Valorem Royalties under Asymmetric Information". *International Journal of Industrial Organization* 37, pp. 38–46.

Jones, M. C. (2009). "Kumaraswamy's Distribution: A Beta-Type Distribution with some Tractability Advantages". *Statistical Methodology* 6, 7081.

Kamien, M. I. (1992). "Patent Licensing". *Handbook of Game Theory*. Ed. by R. Aumann and S. Hart. Vol. I. Elsevier Science, pp. 331–354.

Kamien, M. I. and N. L. Schwartz (1991). *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. second. North-Holland.

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

– (1986). "Fee versus Royalties and the Private Value of a Patent". *Quarterly Journal of Economics* 101, pp. 471–491.

Katz, M. and C. Shapiro (1985). "On the Licensing of Innovations". *RAND Journal of Economics* 16.4, pp. 504–520.

Kumaraswamy, P. (1980). "A Generalized Probability Density Function for Double-Bounded Random Processes". *Journal of Hydrology* 46, pp. 79–88.

Macho-Stadler, I., X. Martínez-Giralt, and D. Pérez-Castrillo (1996). "The Role of Information in Licensing Contract Design". *Research Policy* 25, pp. 25–41.

Milgrom, P. (1987). "Auction Theory". *Advances in Economic Theory*. Ed. by T. F. Bewley. Cambridge: Econometric Society Monographs No. 12, Cambridge University Press, pp. 1–32.

Poddar, S. and U. Sinha (2010). "Patent Licensing from a High-Cost Firm to a Low-Cost Firm". *The Economic Record* 86, pp. 384–395.

Rostoker, M. (1984). "A Survey of Corporate Licensing". *IDEA* 24, pp. 59–92.

Seierstad, A. and K. Sydstæter (1977). "Sufficient Conditions in Optimal Control Theory". *International Economic Review* 18, pp. 367–391.

Sen, D. and Y. Tauman (2007). “General Licensing Schemes for a Cost-Reducing Innovation”. *Games and Economic Behavior* 59, pp. 163–186.

Taylor, C. and Z. Silberston (1973). *The Economic Impact of the Patent System*. Cambridge University Press.

U.S. Department of Justice and Federal Trade Commission (1995). *Antitrust Guidelines for Licensing of Intellectual Property*. Washington, D. C.

Vishwasrao, S. (2007). “Royalties vs. Fees: How do Firms Pay for Foreign Technology?” *International Journal of Industrial Organization* 25, pp. 741–759.

Wang, K. A., W. Liang, and P. Chou (2013). “Patent Licensing under Cost Asymmetry among Firms”. *Economic Modelling* 31, pp. 297–307.

Wang, X. H. (1998). “Fee versus Royalty Licensing in a Cournot Duopoly Model”. *Economics Letters* 60, pp. 55–62.