Field-of-use restrictions in licensing agreements

Schuett, Florian

Toulouse School of Economics

15 July 2007

Field-of-use restrictions in licensing agreements

Florian Schuett*

University of Toulouse

This version: July 2007

Abstract

A widely used clause in license contracts – the field-of-use restriction (FOUR) – precludes licensees from operating outside of the specified technical field. When a technology has several distinct applications, FOUR allow the licensor to slice up his rights and attribute them to the lowest-cost producer in each field of use. This can improve production efficiency. However, with complex technologies, the boundaries of fields of use may be difficult to codify, entailing a risk of overlap of licensees' rights. We explore how this affects the optimal license contract in a moral hazard framework where the licensor's effort determines the probability of overlap. We show that depending on the contracting environment, the license agreement may include output restrictions and nonlinear royalty schemes.

Keywords: licensing, usage restrictions, overlap.

JEL Codes: L24, O3, D23.

* e-mail: florian.schuett@univ-tlse1.fr. I am grateful to Jacques Crémer, Guido Friebel, Patrick Rey, Christian Ruzzier, Paul Seabright, Jean Tirole, seminar participants at the University of Toulouse and participants at the 2007 Spring Meeting of Young Economists in Hamburg, Germany, for helpful comments and suggestions. All errors are mine.

1 Introduction

Field-of-use restrictions (henceforth, FOUR) are clauses that preclude licensees from using the licensed technology in fields other than those specified in the license contract. For example, if a firm patents a new touch-screen technology, it could give an exclusive license to one firm for use of the technology in mobile phones, and exclusively license another firm for use in PDAs. The firms would be prohibited from marketing products using the licensed technology outside of their respective fields. Several empirical studies have found this sort of usage restriction to be extremely common.[1] Nevertheless, their impact on licensing contracts has been left largely unexplored by economists.[2] This is somewhat surprising since practitioners' comments suggest that with complex technologies, fields of use may often be difficult to separate. The licensor must be very careful and may have to exert considerable effort in drafting the field-of-use clauses in order to avoid that licensees' rights overlap. Moreover, overlap may not always be discernible for a court of justice. To go back to the example mentioned above, if the mobile phone company introduces a new phone that includes an organizer function, and the PDA manufacturer launches a new model that allows sending text messages to mobile phones, are they violating the field-of-use restrictions? Under these circumstances, the risk of overlap can have profound implications for the design of license contracts, a fact that the economics literature has ignored so far.

This paper takes a first step towards closing this gap. We look at one particular situation where FOUR arise quite naturally: when a patent holder's technology has several different

[1] Caves *et al.* (1983) survey 22 licensors and 34 licensees from the United States, Canada and the United Kingdom. They find that 34 percent of the license agreements in their sample include "market restrictions", that is, restrictions preventing the licensee from selling outside certain specified markets (no distinction is made between territorial and field-of-use restrictions). Anand and Khanna (2000) conduct a large-scale inter-industry study of licensing behavior. Their sample consists of 1365 licensing deals involving at least one US corporation, as documented by the Securities Data Company. 37 percent of the agreements in their sample are identified as incorporating temporal, product or geographic restrictions, but again no distinction is made between the types of restriction. Anand and Khanna (2000) argue that this figure probably underestimates the actual frequency of such restrictions, because they are not always divulged in public announcements. Bessy and Brousseau (1998) compile a sample of 46 licensing agreements through a survey of French firms. They find that 87 percent of license contracts include some form of usage restriction. Contrary to the first two studies, they distinguish among the different types of restrictions. While geographical restrictions are the most common (58.7 percent), clauses limiting the field of application also turn out to be pervasive. They appear in 50 percent of the licences in the sample.

[2] From an antitrust perspective, usage restrictions are generally considered benign (see, e.g., the *Antitrust Guidelines for the Licensing of Intellectual Property* issued by the US Department of Justice and the Federal Trade Commission in 1995, and Gilbert and Shapiro (1997)).

applications, and different firms are specialized in each of the possible fields of use. Then, production efficiency requires that the lowest-cost firm in each field of use get an exclusive license for that field. Since ex ante, the contracting parties want to maximize the size of the pie, they will agree to slice up the patent holder's rights. The risk of overlap, however, may constrain them in their ability to achieve the efficient outcome. We look at a very simple setup with two symmetric applications of the licensor's technology, and two licensees, each of whom is specialized (i.e., has a competitive advantage in terms of production costs) in one of the fields of use. The licensor's level of care in drafting the field-of-use clauses determines the probability of overlap, and writing a precise contract is costly. If there is no overlap, both licensees are monopolists in their respective field of use. By contrast, in the event of overlap, each of them can produce in both fields of use. We assume that they compete in quantities, so that the equilibrium is asymmetric Cournot.[3] Since the industry's production technology is inefficient under duopoly (high-cost firms contribute to production in equilibrium), overlap reduces joint profits.

We consider three different contracting environments which differ in their degree of "completeness". In the first, overlap can be included as a contingency in the contract. That is, courts can observe whether each licensee has exclusivity. This implies that royalty payments can depend on whether or not there is overlap. In the second, overlap is noncontractible,[4,5] but the licensor's effort – and thus the *probability* of overlap – is observable to the licensees. In the third, overlap is noncontractible and effort is unobservable.

We derive the optimal contract in each of those environments, without imposing any ad-hoc restriction on the permissible royalty schemes, except that they can depend only on each licensee's total quantity. That is, we take a mechanism design approach and look for the feasible quantity-transfer pairs that maximize the licensor's payoff, but we restrict attention to a simple class of mechanisms: the licensor is limited to proposing a menu of contracts and cannot play any more sophisticated message game. From the resulting

[3] The asymmetry is due to the assumption that firm i has a lower marginal cost of producing in field of use i than firm j.

[4] The environments we consider – overlap being perfectly contractible, or not at all – clearly are polar cases. In reality, courts may sometimes find breach of contract when fields of use overlap, and sometimes not. "Contractibility" would thus be probabilistic.

[5] The environment where parties do not include overlap as a contingency could also be interpreted as a shortcut to account for the issues raised in Spier (1992), where contracts are strategically left incomplete by the principal in order to signal his type. In the current context, a clause specifying what happens in the event of overlap might be bad news about the licensor's ability to separate fields of use.

quantity-transfer pairs, we deduce the optimal royalty scheme.

When overlap is contractible, the license agreement consists of a fixed fee equal to the licensees' expected profit given the equilibrium level of effort (the licensor is assumed to have all the bargaining power and thus extracts the entire surplus), as well as an output restriction and a penalty imposed on the licensor, both of which apply only in case of overlap. The optimization problem can be decomposed into two steps: first, the royalty scheme is chosen so as to maximize joint surplus given the realization of overlap. When there is no overlap, the optimal scheme is no royalty since royalties distort the licensees' production decision (this is the standard double marginalization argument). When there is overlap, the royalty scheme is used to soften competition between licensees. One way of doing this is to impose a quantity restriction that is chosen so as to maximize industry profits subject to the licensees' optimizing behavior. Second, the penalty is set so as to induce the licensor to exert the efficient level of effort given the difference between overlap and no-overlap profits. The optimal penalty is equal to the marginal cost of effort evaluated at the efficient level.

When overlap is noncontractible but effort is observable, the royalty scheme can no longer be conditioned on the realization of overlap. The contract must now satisfy incentive compatibility (IC) constraints. We show that implementability requires that licensees produce more under duopoly (overlap) than under monopoly (no overlap). That is, there is a conflict between incentive compatibility and efficiency (which requires that licensees produce less in case of overlap) which is known as *nonresponsiveness* (Guesnerie and Laffont, 1984). Accordingly, a balance needs to be struck between softening competition in case of duopoly and avoiding distortions in case of monopoly. At the same time, there is no moral hazard problem, so that the royalty scheme can be designed with the sole objective of maximizing expected industry profits. The optimal royalty scheme then reduces to an output restriction (quantity rationing). The restriction is typically below the monopoly quantity. Costly effort to separate fields of use is always exerted.

When effort is unobservable, the royalty scheme has to accomplish an additional function on top of the previous one: Incentivizing the licensor to take the appropriate care in drafting the field-of-use clauses. In fact, the patent holder has no incentive to exert effort unless his royalty income depends positively on effort. Notice first that this means that effort is incompatible with the quantity rationing scheme that is optimal under observability. Of course, the licensees will anticipate this and would be willing to pay only their expected profit given the minimum level of effort to obtain a license if they were offered such a scheme. Second, recall that whatever the royalty scheme, monopoly output cannot exceed duopoly output. Hence, a royalty scheme that induces effort must be such that output in case of overlap is strictly greater than in the no-overlap case. Third, the royalty payment associated with the no-overlap quantity must exceed the payment in case of overlap. This implies that in the no-overlap case, licensees will have an incentive to produce more than the monopoly quantity in order to pay lower royalties. Solving for the optimal contract, we find that under some conditions, the resulting royalty scheme is no longer trivial and features royalties that are decreasing with output over some range. We identify a tradeoff between providing incentives to the licensor and producing at the efficient scale. While raising the duopoly quantity relaxes the IC and thereby induces greater effort, it also moves duopoly output away from the efficient level. We show that this can be optimal when the cost of effort is sufficiently convex; otherwise, a nonexclusive license to both firms prevails.

Related literature This paper is related to two separate strands of literature: The literature on licensing, and particularly on the role of royalties in license agreements, and the literature on vertical restraints, specifically territorial restrictions. Much of the literature on licensing has been concerned with explaining the widespread use of royalties in practice, as documented by, e.g., Taylor and Silberston (1973), Contractor (1981) and Rostoker (1984), which contrasts with the theoretical finding that in a standard setup with risk-neutral firms and symmetric information, royalties are undesirable. This result was first established in the context of a (cost-reducing) process innovation by Kamien and Tauman (1986) who showed that fixed fees dominate royalties for a patentee who licenses to a Cournot oligopoly.[6] Meanwhile, in the context of product innovations, when the inventor is unable to work the patent himself, it is generally optimal to give exclusive rights to a single licensee.[7] As we know from the literature on vertical control, due to the issue of double marginalization (Spengler, 1950), the licensor should then set royalties to zero (i.e., marginal cost) and extract the surplus through a fixed fee. Since the early nineties, scholars have turned their attention to the conflict between theoretical predictions and the

[6] Katz and Shapiro (1986) show that auctioning off a fixed number of licenses, strictly below the total number of firms in the industry, does even better than a simple fixed fee; see Kamien (1992) for a survey of the literature on licensing of cost-reducing innovations.

[7] The argument that is sometimes made according to which, in the presence of increasing marginal costs of production, it can be optimal to license several firms seems unconvincing since a single firm should be able to replicate what several firms are doing, for instance by setting up several production plants.

available empirical evidence. Among the explanations for the use of royalties that have been put forward are adverse selection (Gallini and Wright, 1990; Beggs, 1992), moral hazard (Choi, 2001) and risk aversion (Bousquet *et al.*, 1998). This paper offers an alternative rationale for royalties which is based on the possibility that licensees' fields of use may overlap, meaning that the licensor needs to be given incentives to take the appropriate level of care in drafting the license contract.

The vertical control literature has looked for conditions under which an upstream firm with market power will find it optimal to impose various kinds of vertical restraints on downstream firms, one of which are territorial restrictions.[8] Among the results that have been established is that, when retailers provide pre-sale services (such as advertising or consumer information) which have public-good aspects, competition will lead to free riding, a problem that can be solved by establishing local monopolies (Mathewson and Winter, 1984). When there is uncertainty on demand and/or cost parameters and retailers are better informed than the manufacturer, exclusive territories make better use of the retailers' information than retail competition since the latter drives market prices down to the wholesale price (Rey and Tirole, 1986). This literature generally assumes that downstream firms are identical (so that there is no reason in terms of cost efficiency for exclusive territories) and that defining (and enforcing) territorial restrictions is costless. Our approach is different in that we adopt a setup where there are natural advantages to exclusivity because there is a single most efficient firm in each field. However, carelessness in the definition of fields may lead to overlapping rights. We investigate how the risk of overlap affects the desirability of field-of-use restrictions. While we focus on technical fields, our analysis may also apply to geographical territories in some cases: Retailers with a lot of experience in selling in a particular geographical region may have advantages over competitors who lack knowledge of local characteristics, and territories may sometimes have to be appropriately defined to avoid ambiguity. A recent decision by the French supreme court is a case in point: Overturning the decision of an appeals court, the judges ruled that a manufacturer who had granted an exclusive territory to a retailer did *not* violate the contract by also offering his products through the internet.[9]

The remainder of this paper is organized as follows. Section 2 describes the setup of

[8] See Katz (1989) and Rey and Vergé (2007) for an overview of the economics of vertical restraints.

[9] *Cour de Cassation, Chambre Commerciale, March 14, 2006 - N° du pourvoi: 03-14639*, see http://www.legalis.net/jurisprudence-decision.php3?id_article=1602 (in French).

the model. Section 3 derives the optimal license contract in each of the three contracting environments we consider. Finally, in section 4, we highlight some empirical predictions that follow from our results, and briefly discuss implications for antitrust authorities.

2 A model with two licensees

Consider the following setup. A patent holder (P) wants to license his patented technology in two fields of use (1 and 2) where other firms have cost advantages.[10] Both fields of use are equally profitable. It is common knowledge that firm L_1 is the lowest-cost producer of application 1 and firm L_2 the low-cost producer of application 2. Denoting by c_{ij} the constant unit cost of production of application j by licensee i, we have $c_{ii} < c_{ij}$ for $i = 1, 2$ and $j \neq i$. Assume for simplicity that $c_{ii} = c_L$ and $c_{ij} = c_H$, and denote $\Delta c \equiv c_H - c_L$ the difference in costs. If drafting a contract that clearly separates the licensees' fields of use were costless, it would thus be optimal (i.e., maximize joint profits) to give L_1 an exclusive license restricted to field 1 in exchange for a fixed upfront payment, and analogously for L_2 in field of use 2. Royalties should not be used since, by distorting licensees' decisions, they decrease joint profits. Similarly, no restrictions on licensees' output should be imposed by the licensor.

However, we will assume that precise drafting is costly for the patent holder, and that the precision of the contractual language affects the probability of overlap. Let e be the effort exerted by P when writing the field-of-use clauses. Effort is chosen in the interval $[0, 1]$. The cost of effort is $\psi(e)$, increasing and convex, with $\psi(0) = \psi'(0) = 0$. With probability $1 - e$, the fields of use are so broadly defined that they turn out to overlap. If there is overlap, each firm can produce in *both* fields of use. Thus, in particular, zero effort corresponds to the case of a nonexclusive license to all fields of use for both firms.

The effort variable e can be interpreted in several ways, with different implications in terms of its observability for licensees. Assume there is a list of attributes on which fields could differ (this list could be large – possibly infinite, and include things such as size, color, materials used, ...). Only a few (possibly a single one) of them are relevant for cleanly defining the boundary between the two fields. Effort could consist in the time

[10] We assume that the technology adds sufficient value to existing products to create a completely new market, rather than merely reducing costs. Much of the licensing literature has dealt with cost-reducing innovation, for which it is generally optimal to license many firms (at least for non-drastic innovations). For the purposes of this paper, however, a framework where there are advantages to exclusivity is needed.

and money spent to find out the relevant attributes. Alternatively, effort might be the number of attributes included in the contract (as opposed to those that are left out), in a setting where adding a contingency to the contract is costly, as in Dye (1985) or Bajari and Tadelis (2001). In the former case, effort is observable if licensees can check which amount the licensor has invested, and unobservable otherwise. In the latter case, effort is observable if there is a finite number of attributes, so that the number of attributes included in the contract is a sufficient statistic for effort, and unobservable if there is an infinity of attributes, rendering the number in the contract meaningless. In section 3, we consider both the case where effort is observable and the case where it is unobservable.

Since we will consider the question of whether the use of a royalty can be beneficial when there is some probability of overlap, we have to make assumptions on the observability and verifiability of some other variables. We start by clarifying our notion of overlap. When fields of use overlap, we hold that the descriptions in the field-of-use clauses do not allow a court to establish whether a device produced under the terms of the license is destined for field 1 or 2. This seems natural since it is precisely the object of the contract to define what the fields of use are. Overlap corresponds to the case where the licensor has failed to draw a clear boundary. Thus, a court may be able to rule whether licenses overlap[11] – whether each licensee indeed has exclusivity in his field of use, or whether the licenses are so broad that each licensee can produce in both fields, thus violating exclusivity – but when the licenses turn out to overlap, the court generally cannot guess what the parties' original intention was – which field to attribute exclusively to who. Accordingly, letting q_{ij} denote L_i's output in field of use j, the contract can depend only on the total output of each licensee (that is, $q_{ii} + q_{ij}$) and not on q_{ij} individually.

We assume that total output by each firm is observable and contractible. Effectively, this means we assume that courts can verify whether the terms of the license cover the products produced by the licensees, and thus – among other things – whether royalty payments are due. Prices are unobservable to the licensor. Royalties cannot depend on the other firm's output. Apart from this, any royalty scheme (linear or nonlinear) that depends only on output is possible.

The timing is as follows (see figure 1): at date 0, the contract is drafted by P who chooses a level of care e. The contract specifies a fixed upfront payment F and a royalty scheme $R(q)$ (i.e., $R(q)$ is the royalty payment associated with an output of q units of a

[11] The possibility of using the event of overlap itself as a contingency is discussed in the following section.

product that uses the licensed technology). L_1 and L_2 accept or reject the contract.[12] If the contracts are accepted, each pays the fixed fee to the patent holder. At date 1, the licensees learn whether their license permits them to enter the competitor's field of use. Finally, at date 2, firms produce. If there is no overlap, they are monopolists, otherwise they compete in both fields of use. The discount factor is assumed equal to 1.

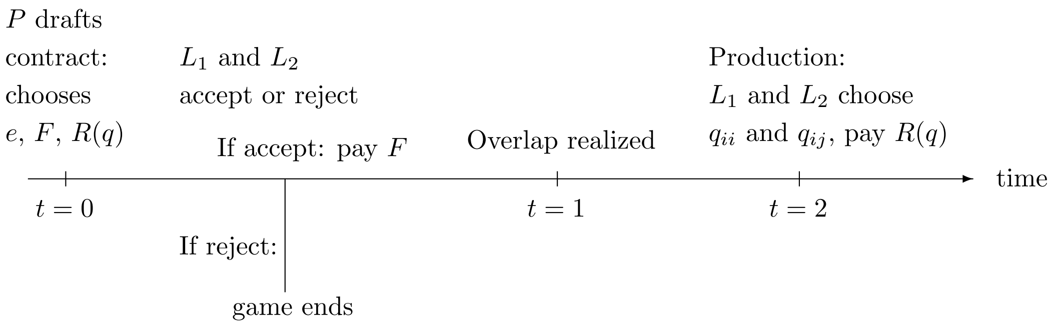

Figure 1: Timing of the game

If there is overlap, firms compete à la Cournot.[13] Inverse demand functions are identical and given by $p(q) = a - q$ where q is the total quantity sold in a given field. That is, if L_i sells q_{ii} and L_j sells q_{ji} in market i, inverse demand is $a - q_{ii} - q_{ji}$.[14]

[12] We assume that the contract offers are public. The literature on vertical control has sometimes used public contracts (see, e.g., Mathewson and Winter, 1984), and at other times secret contracts (see, e.g., Hart and Tirole, 1990). The issue that arises with secret offers in the context of licensing is one of commitment: while ex ante, the patent holder would like to commit himself, for instance, to license only one firm, once the contract is signed he is tempted to hand out additional licenses, eventually leading to a flooding of the market which erodes profits (Rey and Tirole, 2007). This is an example of contracting with externalities; see Segal (1999) for an excellent synthesis of many existing models which exhibit this feature. While an exclusivity clause may solve the problem in the particular example mentioned above, the issue is in fact much more general. In this paper, we abstract from the complexities that result from secret contract offers and focus on public offers.

[13] When firms compete à la Bertrand and their only strategic variable is price, P can use a simple royalty scheme such as a unit royalty, set the royalty rate such that the duopoly quantity is exactly equal to the monopoly quantity at a zero royalty, and choose $e = 0$, thereby achieving the most desired (joint profit maximizing) outcome. There is no inefficiency since, under price competition with asymmetric costs, the high-cost firm's output is zero. Hence, in that extreme case there is no reason for the use of FOUR.

[14] This setup – Cournot competition with linear demand – has been widely used in the economic analysis of licensing.

3 The optimal design of the license contract

We will consider several scenarios concerning the verifiability of overlap and observability of effort. To establish a benchmark, we will derive the contract that would obtain if overlap were contractible.[15] It is likely, though, that the event of overlap is not a perfectly contractible variable.[16] In practice, P can perhaps be held responsible for careless drafting (overlap amounts to a breach of contract since licensees do not get the exclusivity they signed up for), but whether a court will actually find him liable is uncertain at best. Therefore, we will proceed to consider the other polar case where overlap is noncontractible. In a first step, we examine what the optimal contract is when licensees observe the amount of effort that P invests in the definition (and clean separation) of fields of use. In a second step, we investigate the opposite case where licensees do not observe e, so that a moral hazard problem arises. To summarize, in what follows we study the optimal license contract in three different contracting environments: (1) when overlap is contractible, (2) when overlap is noncontractible and effort is observable, and (3) when overlap is noncontractible and effort is unobservable.

Throughout the paper, we will take a mechanism design approach to the problem. That is, we look for pairs of quantities and associated royalty payments for each state of nature (overlap and exclusivity) which are uniquely implementable in Nash equilibrium, and from this set, we choose the ones that maximize the licensor's expected profit.[17] The resulting "dots" in the quantity-royalty space allow us to deduce the features of an optimal royalty scheme. We denote the state of nature by $s \in \{m, d\}$ where m (monopoly) corresponds to exclusivity and d (duopoly) corresponds to overlap, and the associated quantities and royalty payments by $q(s)$ and $r(s)$, respectively.

3.1 Benchmark: contracting on overlap is possible

This section derives a sort of "first best" solution that obtains if the contract can be contingent on the state of nature s (i.e., on overlap). We will deal with the case where effort

[15] If it may seem implausible that real-world contracts would actually be contingent on overlap, it is interesting to note that they sometimes are: for an example of a license agreement which includes an overlap clause from the SEC info database, see http://www.secinfo.com/dRqWm.8vWq.a.htm.

[16] Meanwhile, note that making the contract contingent on there being competition may be impractical because it is hard to establish that firms are actually competing with each other.

[17] As stated in the introduction and explained in more detail below, there is one qualification: we will restrict the available mechanisms to the simple class of incentive compatible contracts.

is observable first before considering the more general case where effort is unobservable. Since parties are risk neutral, the efficient solution will be attained regardless of whether effort is observable, so that the second part merely consists in finding a way to implement the efficient level of effort.

With contractible overlap and observable effort, the licensor's problem is

$$\max_{e;\{F_i\}_{i=1,2};\{(q_i(m),r_i(m));(q_i(d),r_i(d))\}_{i=1,2}} \sum_i F_i + e\sum_i r_i(m) + (1-e)\sum_i r_i(d) - \psi(e)$$

$$\text{subject to} \quad F_i \leq e\big[\Pi_m\big(q_i(m)\big) - r_i(m)\big] + (1-e)\big[\Pi_d\big(q_i(d), q_j(d)\big) - r_i(d)\big], \quad i = 1,2;\ j \neq i$$

where $\Pi_s(\cdot)$ is the equilibrium (gross of royalty) profit attained by a licensee given the market structure s and each firm's total output (this will be defined more precisely below). Notice that, consistent with the assumption that the royalty scheme cannot depend on q_{ij} in case of overlap, the mechanism only specifies a total quantity $q_i(d) = q_{ii} + q_{ij}$.

Since the licensor makes a take-it-or-leave-it offer, he will choose F as large as possible; hence, the constraint must be binding at the optimum. The problem becomes

$$\max_{e;\{(q_i(m),r_i(m));(q_i(d),r_i(d))\}_{i=1,2}} \quad e\sum_i \Pi_m\big(q_i(m)\big) + (1-e)\sum_i \Pi_d\big(q_i(d), q_j(d)\big) - \psi(e). \quad (1)$$

Clearly, the choice of quantities is independent of e, so the problem of finding the optimal license contract can be decomposed into two steps: First, determining the optimal quantity for any given realization of overlap, and deducing the royalty scheme that achieves it (note that since the event of overlap can be included as a contingency in the contract, the royalty scheme can depend on its realization so that there can be two different royalty schemes $R_m(\cdot)$ and $R_d(\cdot)$); second, finding the level of effort that maximizes expected joint profit given the difference between licensee profits in the overlap and no-overlap cases. Note also that, since P can extract the licensees' profit through the fixed fee, the levels of royalty payments, $r(m)$ and $r(d)$, are irrelevant. The use of a royalty will be motivated solely by the desire to influence the licensees' decisions.

No overlap: monopoly. When both fields of use are cleanly separated, each licensee is a monopolist in his market. Each licensee's monopoly profit as a function of quantity q is

$$\Pi_m(q) = [a - q - c_L]q. \quad (2)$$

The profit-maximizing quantity is $q_m^* = \frac{a-c_L}{2}$. It is obvious from (1) that the mechanism should implement $q_i(m) = q_m^*\ \forall i$.

Thus, any royalty scheme that induces the licensee to produce the monopoly quantity q_m^* is a solution. One particular solution that stands out for its simplicity is no royalty at all, $R_m(q) = 0 \quad \forall q$. This is the standard double marginalization argument (Spengler, 1950): in a vertical relationship, the upstream firm should charge a price to the downstream firm that equals the marginal cost of the input supplied, and extract the surplus through a fixed upfront payment. Here, the input is simply information (technological knowledge), which has zero marginal cost, so that the optimal price is zero too.

Overlap: duopoly. When fields of use overlap, both licenses are effectively nonexclusive: licensee L_i faces competition in field i from (less efficient) licensee L_j and vice versa. While in the case of monopoly, royalties are undesirable because they create a distortion, in the duopoly case royalties can be useful as a means of softening competition. We will proceed as follows: first, we derive licensees' equilibrium behavior given that each of them is required to produce a total quantity $q_i(d)$; second, we determine the quantity that maximizes industry profit subject to the licensees' optimizing behavior; third, we deduce the R that achieves those quantities.

To begin, we make an assumption on c_H.

Assumption 1 *The inefficient firm's marginal cost satisfies* $c_H < \dfrac{a + c_L}{2}$.

In words, c_H must be lower than the monopoly price that would prevail without royalties. This assumption makes sure that L_i enters field of use j in the absence of a royalty, i.e., that in case of overlap, licensees are actually a threat to each other.

Since, by assumption, the firms compete in quantities, L_i maximizes over q_{ii} and q_{ij}

$$\max_{q_{ii}, q_{ij}} [a - q_{ii} - q_{ji} - c_L]q_{ii} + [a - q_{jj} - q_{ij} - c_H]q_{ij} \quad \text{subject to} \quad q_{ii} + q_{ij} = q_i(d) \tag{3}$$

while taking q_{ji} and q_{jj} as given. The following lemma characterizes the equilibrium values of q_{ii} and q_{ij}.

Lemma 1 *Let $q_i(d)$ be the total quantity that licensee i should produce in case of overlap. If $q_i(d) \geq \Delta c$ and $q_j(d) \geq \Delta c$, the Cournot-Nash equilibrium of the game has*

$$\begin{aligned} q_{ii} &= \frac{q_i(d) + \Delta c}{2} \\ q_{ij} &= \frac{q_i(d) - \Delta c}{2} \end{aligned}$$

for all i, $j \neq i$ and for any $q_j(d) \geq \Delta c$.

If $q_i(d) \geq \Delta c$ and $q_j(d) < \Delta c$, the equilibrium is

$$\begin{aligned} q_{ii} &= \frac{2q_i(d) + q_j(d) + \Delta c}{4} \\ q_{ij} &= \frac{2q_i(d) - q_j(d) - \Delta c}{4} \\ q_{jj} &= q_j(d) \\ q_{ji} &= 0. \end{aligned}$$

Finally, if $q_i(d) < \Delta c$ and $q_j(d) < \Delta c$, the equilibrium has

$$\begin{aligned} q_{ii} &= q_i(d) \\ q_{ij} &= 0 \end{aligned}$$

for all i, $j \neq i$ and for any $0 \leq q_j(d) < \Delta c$.

Proof: Rewriting (3) using $q_i(d)$ in the constraint and taking into account that L_j faces an analogous problem so that $q_{ji} = q_j(d) - q_{jj}$, L_i's problem becomes

$$\max_{q_{ii}} [a - q_{ii} - (q_j(d) - q_{jj}) - c_L]q_{ii} + [a - q_{jj} - (q_i(d) - q_{ii}) - c_H](q_i(d) - q_{ii}).$$

The first-order condition of the problem is

$$a - 2q_{ii} - q_j(d) + q_{jj} - c_L = a - q_{jj} - 2(q_i(d) - q_{ii}) - c_H.$$

Solving for q_{ii} and replacing q_{jj} from the first-order condition of L_j's problem, we obtain

$$q_{ii} = \frac{\Delta c - q_j(d) + 2q_i(d) + [\Delta c - q_i(d) + 2q_{ii} + 2q_j(d)]/2}{4},$$

which can be solved for the equilibrium q_{ii} which, in turn, can be replaced in the constraint to obtain q_{ij}, yielding the expressions claimed in the first part of the lemma (for the case where $q_i(d) > \Delta c$ for all i). These are valid solutions as long as $q_{ij} \geq 0$ for all i which is the case as long as $q_i(d) \geq \Delta c$. When $q_i(d) < \Delta c$, L_i's output in field j, q_{ij}, is zero, while $q_{ii} = q_i(d)$. Modifying the first-order condition accordingly yields the second part of the lemma. ■

Thus, the difficulty for the licensor is that he can only control the total quantity produced by each licensee. He can hardly influence the distribution of output between the two fields of use – except when the total quantity is so restrictive as to deter a licensee from entering his competitor's market altogether. Otherwise, the Cournot-Nash equilibrium has each licensee producing in both fields of use, with the share of the more efficient firm determined by the difference in marginal costs. The consequence of this is that, in equilibrium, some of the units sold are generally contributed by high-cost producers, making the industry's production technology inefficient. The licensor will optimally react to this by inducing less than monopoly output.

Note also that Lemma 1 implies that it can never be optimal for the licensor to choose a quantity $q_i(d)$ that is strictly below Δc. For any given total output such that $q_i(d)+q_j(d) \geq 2\Delta c$, the proportion that is produced at low cost (c_L) is always greater if each licensee produces at least Δc. Moreover, as long as $q_i(d) \geq \Delta c$ for all i, all that matters for industry profits is aggregate output, $q_i(d) + q_j(d)$, and not how it is shared between the licensees. Therefore, we will from now on suppose that the licensor treats both licensees symmetrically, so that $q_i(d) = q_j(d) = q(d)$, which greatly simplifies notation.

Formally, the licensor's problem then is to choose $q(d)$ so as to

$$\max_{q(d)} [a - q(d)]q(d) - c_L \min\{q(d), (q(d) + \Delta c)/2\} - c_H \max\{0, (q(d) - \Delta c)/2\}. \tag{4}$$

The next lemma describes the solution to this problem.

Lemma 2 *If $c_H \leq (2a + 3c_L)/5$, the quantity q_d^* that maximizes industry profits is given by*

$$q_d^* = \frac{a - (c_L + c_H)/2}{2}$$

and both firms contribute to production with $q_{ii} = (2a-5c_L+3c_H)/8$ and $q_{ij} = (2a-5c_H+3c_L)/8$ for all i.

If $c_H > (2a + 3c_L)/5$, only the low-cost producer is active, i.e. $q_{ij} = 0$, and $q_d^* = \Delta c$. In both cases, the aggregate output is lower than the monopoly quantity: $q_d^* < q_m^*$.

Proof: The optimization program

$$\max_q [a - q]q - c_L \frac{q + \Delta c}{2} - c_H \frac{q - \Delta c}{2}$$

leads to the first-order condition $a - 2q - (c_L + c_H)/2 = 0$, the solution of which is $q_d^* = \frac{a-(c_L+c_H)/2}{2}$. This is a valid solution to (4) as long as it is greater than Δc, that is

$$\frac{a - (c_L + c_H)/2}{2} \geq \Delta c \Leftrightarrow c_H \leq (2a + 3c_L)/5.$$

One can then determine q_{ii} and q_{ij} using Lemma 1. If $c_H > (2a + 3c_L)/5$, we have a corner solution so that the second part of Lemma 2 applies. As for the last claim, $\frac{a-(c_L+c_H)/2}{2} < \frac{a-c_L}{2}$ since $c_H > c_L$, while $\Delta c < \frac{a-c_L}{2}$ follows from Assumption 1. ■

In order to limit the damage caused by overlap, there are two possibilities. One can shut down the inefficient firm by reducing output to Δc. Alternatively, one can be less restrictive but at the expense of involving the inefficient firm in production. The intuition for Lemma 2 is that, when c_H is low, one needs to sacrifice a lot of output to deter the high-cost firm from entering, while at the same time the loss in production efficiency from involving it is not too important. Therefore, it is optimal to allow both firms to be active, albeit at an aggregate level of activity that is below q_m^*. By contrast, when c_H is high, having the high-cost firm contribute to production is very inefficient, and at the same time it is not too costly to exclude it. Therefore, restricting output to Δc is optimal. Figure 2 below illustrates the case where $c_H < (2a + 3c_L)/5$ so that letting both firms produce is optimal. q_d^C denotes the Cournot duopoly output in the absence of royalties.

There are many royalty schemes that allow the licensor to implement $q(d) = q_d^*$. The simplest is a quantity restriction that limits output to q_d^*, that is,

$$R_d(q) = \begin{cases} 0 & \text{for} \quad q \leq q_d^* \\ \infty & \text{for} \quad q > q_d^*. \end{cases}$$

He can also use a unit royalty, $R_d(q) = \rho q$ with

$$\rho = \begin{cases} \frac{a-(c_H+c_L)/2}{4} & \text{for} \quad c_H < (2a + 3c_L)/5 \\ a - 2c_H + c_L & \text{for} \quad c_H \geq (2a + 3c_L)/5. \end{cases}$$

The licensor's effort choice. When effort is observable, the licensor chooses his level of care so as to maximize expected profits. In the preceding analysis, we have derived the optimal quantity for each realization of overlap; let the associated levels of (gross of

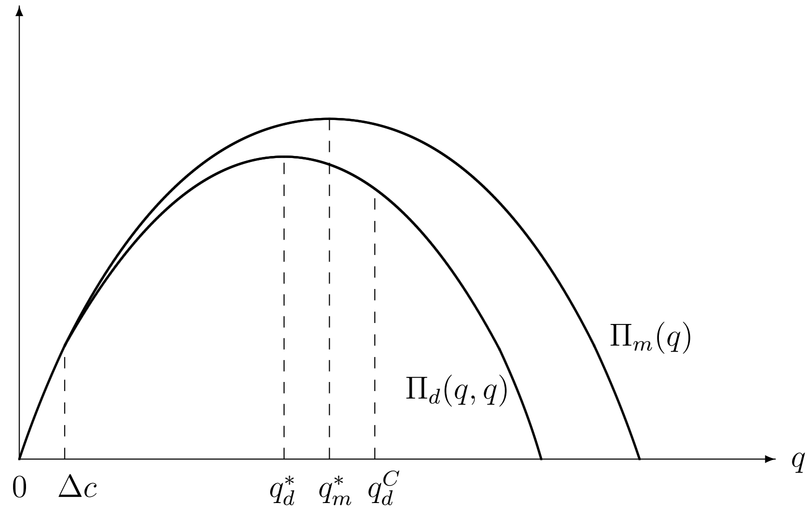

Figure 2: Licensees' profits with and without overlap

royalty) profit be $\pi_m^* \equiv [a - q_m^* - c_L]q_m^*$ and $\pi_d^* \equiv [a - q_d^*]q_d^* - c_L(q_d^* + \Delta c)/2 - c_H(q_d^* - \Delta c)/2$.[18] The licensor's choice of e then simply maximizes

$$e\pi_m^* + (1-e)\pi_d^* - \psi(e)/2. \tag{5}$$

When effort is unobservable, the licensor needs to be given incentives to actually choose the efficient level of effort, but this can be easily achieved through an appropriate penalty clause.[19] Proposition 1 summarizes the features of the optimal contract when the realization of overlap is verifiable for a court of justice. (In terms of notation, we have $r_i(s) = r_j(s) = r(s)$ and $F_i = F_j = F$ because of symmetric treatment of licensees.)

Proposition 1 *Assume that overlap is contractible. The optimal contract* $\big(e^c, F^c, (q^c(m), r^c(m)), (q^c(d), r^c(d))\big)$ *then takes the following form:*

[18] Although not of much relevance for the analysis, profits can be computed. They are

$$\pi_m^* = \frac{(a-c_L)^2}{4}$$
$$\pi_d^* = \begin{cases} [4a(a - c_H - c_L) + 9(c_H^2 + c_L^2) - 14c_L c_H]/16 & \text{for} \quad c_H < (2a + 3c_L)/5 \\ (a - c_H)\Delta c & \text{for} \quad c_H \geq (2a + 3c_L)/5. \end{cases}$$

[19] Note that, even though it is the licensor who proposes the contract, this sort of penalty clause is in his own best interest. In the absence of such a clause, licensees correctly anticipate that P's effort will be low, and their willingness to pay is reduced accordingly.

(i) *The level of care exerted by P is e^c, determined by $\psi'(e^c) = 2(\pi_m^* - \pi_d^*)$;*

(ii) *the fixed fee is $F^c = \pi_d^*$;*

(iii) *when there is no overlap, output is $q^c(m) = q_m^*$ and $r^c(m) = \pi_m^* - \pi_d^*$;*

(iv) *when there is overlap, output is $q^c(d) = q_d^*$ and $r^c(d) = 0$.*

This can be implemented through a royalty scheme $R_m^c(q) = \pi_m^* - \pi_d^*$ (independent of q) in case of exclusivity and

$$R_d^c(q) = \begin{cases} 0 & \textit{for} \quad q \leq q_d^* \\ \infty & \textit{for} \quad q > q_d^* \end{cases}$$

for the case of overlap.

Proof: The previous analysis has shown that, in the no-overlap case, the royalty scheme should be non distortive (which is the case here because $R_m^c(\cdot)$ doesn't depend on output), while in case of overlap, Lemma 2 has shown that the optimal royalty scheme should induce q_d^* (which can be achieved through quantity rationing, as in the case of $R_d^c(\cdot)$). When effort is observable, maximizing (5) leads to the first-order condition

$$\pi_m^* - \pi_d^* = \psi'(e)/2. \tag{6}$$

Each licensee's expected profit then is

$$e^c\big(\pi_m^* - r^c(m)\big) + (1 - e^c)\big(\pi_d^* - r^c(d)\big) = e^c\big(\pi_m^* - (\pi_m^* - \pi_d^*)\big) + (1 - e^c)\pi_d^* = \pi_d^*,$$

which P extracts through the fixed fee F^c. His total payoff, given by the fixed fee plus expected royalty revenue less the cost of effort, is $2\big(\pi_d^* + e^c(\pi_m^* - \pi_d^*)\big) - \psi(e^c)$, i.e. the entire aggregate surplus of the relationship.

Turning to the case where effort is unobservable, we have to show that, given the royalty payments $r^c(m)$ and $r^c(d)$, the licensor wants to choose e^c. His preferred level of effort is obtained by solving

$$\max_e \; e\, r^c(m) + (1 - e)\, r^c(d) - \psi(e)/2 = e(\pi_m^* - \pi_d^*) - \psi(e)/2,$$

the first-order condition of which is $\pi_m^* - \pi_d^* = \psi'(e)/2$ which coincides with (6). ■

Proposition 1 says that the optimal level of effort is such that the marginal cost of effort equals the marginal benefit of effort, the latter being given by twice (because there are two fields of use) the difference between monopoly and duopoly profits. The fixed fee is equal to the duopoly profit. When there is no overlap, the licensor is rewarded through a second fixed payment that doesn't affect the licensees' choice of quantity but gives him incentives to take the appropriate care. (Alternatively, this might be achieved through a penalty clause, as mentioned above.)

There are several reasons why licensing contracts may not include an overlap clause. The licensor may be reluctant to insert such a clause because it may be bad news about his ability to cleanly separate fields of use. That is, he may leave the contract incomplete for signaling purposes, as in Spier (1992).[20] Alternatively, overlap may be difficult to verify for a court of justice, or even for the licensor himself: to establish overlap, what needs to be proved is that a good produced by licensee i *within* the terms of his license *competes* with one of licensee j's products. This can be less than straightforward, especially when the licensees have private information on market prices, as we have assumed. Moreover, as argued by Cestone and White (2003), contracting parties may want to complement legal incentives by financial incentives when enforcing certain clauses is expensive and highly uncertain. For these reasons, the following sections consider the case where the license contract cannot be contingent on overlap.

3.2 Overlap noncontractible but effort observable

We now drop the assumption that overlap can be included as a contingency in the contract. We still assume the courts to be able to ascertain whether a product falls within the terms of the license, but they can no longer determine whether all products that can possibly be produced under the license serve only one of the two markets.[21] Thus, when fields of use overlap and L_i sells a product in field j, the court is unable to hold the licensor responsible for the lack of exclusivity enjoyed by L_j. When there is no overlap, meanwhile, there is no problem because the terms of the license do not allow L_i to launch a product that competes with L_j's.

Overlap being nonverifiable, the mechanism design problem is no longer as straightforward. Nevertheless, if we assume that all three contracting parties learn the state of the

[20] Of course, this is outside of the model since we have assumed no information asymmetry with respect to the licensor's cost of effort.

[21] In a sense, courts are unaware of consumers' preferences: they cannot determine whether two products are substitutes, complements or unrelated.

world before the production stage, the fundamental result by Maskin (1977) tells us that the first-best allocation derived in Proposition 1 is Nash implementable because it satisfies monotonicity. And even if we assume that only the licensees learn the state of the world, it is still possible to implement the first best through a direct revelation mechanism where the allocation depends on both players' messages, as we show in the Appendix. Here, we will keep the problem interesting by restricting the set of mechanisms available to the licensor. Specifically, we suppose that the licensor can ask the licensees to report the state of the world, but that each licensee's output and royalty payment can only depend on his own report, and not on the other licensee's message. (Nor can any other sophisticated message game be played.) This is equivalent to having the licensor offer each licensee a menu of contracts. By doing this, we meet a much voiced concern with implementation theory according to which mechanisms are often excessively complex.[22]

Formally, we must now add incentive compatibility constraints to the problem which make sure that it is in the licensees' interest to choose the contract corresponding to the underlying state of the world. In addition, we have to make sure that there is no nontruthful equilibrium. Therefore, the licensor's problem becomes

$$\max_{e;F;(q(m),r(m));(q(d),r(d))} \quad F + e\,r(m) + (1-e)\,r(d) - \psi(e)/2$$

subject to

$$\begin{aligned} F &\leq e\big[\Pi_m\big(q(m)\big) - r(m)\big] + (1-e)\big[\Pi_d\big(q(d), q(d)\big) - r(d)\big] && (7)\\ \Pi_m\big(q(m)\big) - r(m) &\geq \Pi_m\big(q(d)\big) - r(d) && (8)\\ \Pi_d\big(q(d), q(d)\big) - r(d) &\geq \Pi_d\big(q(m), q(d)\big) - r(m) && (9)\\ \Pi_d\big(q(d), q(m)\big) - r(d) &\geq \Pi_d\big(q(m), q(m)\big) - r(m). && (10) \end{aligned}$$

In words, the licensees must prefer reporting m when they have exclusivity, and each licensee must prefer reporting d regardless of what the other reports when there is overlap. From Lemma 1, we can derive the equilibrium profit of licensee i in case of overlap as a

[22] See, e.g., Dewatripont (1992), who notes that "the search for positive implementation results has led to a series of excessively sophisticated games. While there is no reason for *a priori* restricting the set of acceptable mechanisms, it is of some concern that the presumed outcomes of these games rely on extremely subtle equilibrium behavior. If one were to actually apply these games in practice, one can doubt that the agents would play the equilibrium strategies." Dewatripont therefore recommends constraining from the start the set of possible games.

function of total quantities q_i and q_j assuming $q_i \geq \Delta c \quad \forall i$:

$$\begin{aligned}\Pi_d(q_i, q_j) &= \left[a - \tfrac{q_i+\Delta c}{2} - \tfrac{q_j-\Delta c}{2} - c_L\right]\tfrac{q_i+\Delta c}{2} + \left[a - \tfrac{q_j+\Delta c}{2} - \tfrac{q_i-\Delta c}{2} - c_H\right]\tfrac{q_i-\Delta c}{2} \\ &= \left[a - \tfrac{q_i+q_j}{2} - \tfrac{c_L+c_H}{2}\right] q_i + \tfrac{\Delta c^2}{2}. \end{aligned} \tag{11}$$

As the following lemma shows, the incentive constraints (8) and (9) severely restrict the set of implementable allocations.

Lemma 3 *When overlap is nonverifiable, a necessary and sufficient condition for any pair of outputs* $\big(q(m) \geq \Delta c, q(d) \geq \Delta c\big)$ *to be implementable is*

$$q(d) \geq q(m).$$

Proof: Adding up (8) and (9) and rearranging, we have

$$\Pi_m\big(q(m)\big) - \Pi_m\big(q(d)\big) \geq \Pi_d\big(q(m), q(d)\big) - \Pi_d\big(q(d), q(d)\big).$$

Substituting from (2) and (11), this is

$$\begin{aligned}[a - q(m) - c_L]q(m) - [a - q(d) - c_L]q(d) \;\geq\; & \left[a - \frac{q(m)+q(d)}{2} - \frac{c_L+c_H}{2}\right] q(m) - \\ & - \left[a - q(d) - \frac{c_L+c_H}{2}\right] q(d),\end{aligned}$$

which, after simplification, yields

$$\big(q(d) - q(m)\big)\big(q(m) - \Delta c\big) \geq 0.$$

As for the sufficiency part, we now show that (10) is implied by (9) when the monotonicity condition, $q(d) \geq q(m)$, holds. The constraint (10) is satisfied whenever (9) is if

$$\Pi_d\big(q(d), q(m)\big) - \Pi_d\big(q(d), q(d)\big) \geq \Pi_d\big(q(m), q(m)\big) - \Pi_d\big(q(m), q(d)\big).$$

This is true if

$$\frac{d}{dq}\left[\Pi_d(q, q_0) - \Pi_d(q, q_1)\right] > 0,$$

which is the case for any $q_0 < q_1$. ■

The intuition for this result is related to the fact that total output in a Cournot duopoly is higher than under monopoly. In duopoly, increasing output represents an externality:

it reduces the price on all units, but firms only take into account the effect on their own output.

The monotonicity condition in Lemma 3 is important because it highlights the fact that the licensor faces a phenomenon of *nonresponsiveness* (Guesnerie and Laffont, 1984): While efficiency requires $q(d) < q(m)$ (recall Lemma 2), he can only implement output pairs satisfying $q(d) \geq q(m)$. A direct implication of Lemma 3 therefore is that an upper bound to what the licensor can achieve is to have the licensees produce the same quantity whether or not there is overlap; we get a pooling allocation. This is not surprising since nonresponsiveness is frequently associated with pooling of types.

How should the quantity on which to pool, $\bar{q}$, be chosen? The optimal level of effort and $\bar{q}$ are, of course, interdependent: The lower the probability of overlap, the more should licensees' production approach q_m^*. The higher production, the higher is the difference between monopoly and duopoly profits, and the stronger are the incentives to avoid overlap. Proposition 2 characterizes the optimal combination of effort and $\bar{q}$.

Proposition 2 *Assume that overlap is noncontractible and that effort is observable. Then, the optimal license contract* $\big(e^o, F^o, (q^o(m), r^o(m)), (q^o(d), r^o(d))\big)$ *has the following properties (assuming an interior solution):*

(i) *Output is the same irrespective of overlap:* $q^o(m) = q^o(d) = \bar{q}^o$ *given by*

$$\bar{q}^o = \frac{1}{2}\big[a - \big(e^o c_L + (1-e^o)(c_H + c_L)/2\big)\big].$$

(ii) *Effort* e^o *satisfies* $\psi'(e^o) = \Delta c(\bar{q}^o - \Delta c)$*;*

(iii) *the fixed fee is* $F^o = [a - \bar{q}^o - e^o c_L]\bar{q}^o - \frac{(1-e^o)}{2}\big(c_L(\bar{q}^o + \Delta c) + c_H(\bar{q}^o - \Delta c)\big)$.

This can be implemented through a royalty scheme that takes the form of an output restriction, that is,

$$R(q) = \begin{cases} 0 & \textit{for} \quad q \leq \bar{q}^o \\ \infty & \textit{for} \quad q > \bar{q}^o, \end{cases}$$

Proof: The fact that a pooling contract is optimal is a corollary of Lemma 3. Using the fact that the ex ante participation constraint (7) binds, the optimization program determining the level of effort and the output restriction then is

$$\max_{e,\bar{q}} \; e\Pi_m(\bar{q}) + (1-e)\Pi_d(\bar{q},\bar{q}) - \frac{\psi(e)}{2}$$

$$= e[a - \bar{q} - c_L]\bar{q} + (1-e)\left[(a-\bar{q})\bar{q} - \frac{1}{2}\big(c_L(\bar{q}+\Delta c) + c_H(\bar{q}-\Delta c)\big)\right] - \frac{\psi(e)}{2}.$$

The first-order conditions are

$$\begin{aligned} [a - \bar{q} - c_L]\bar{q} - \left[(a-\bar{q})\bar{q} - \frac{1}{2}\big(c_L(\bar{q}+\Delta c) + c_H(\bar{q}-\Delta c)\big)\right] &= \frac{\psi'(e)}{2} \\ e[a - 2\bar{q} - c_L] + (1-e)\left[a - 2\bar{q} - \frac{1}{2}(c_H + c_L)\right] &= 0 \end{aligned}$$

which can be simplified to the claimed expressions which together determine e^o and $\bar{q}^o$. ■

From the expression determining $\bar{q}^o$, it is easy to see that $\bar{q}^o \in [q_d^*, q_m^*]$. Moreover, when $c_H > (2a + 3c_L)/5$, the following argument allows us to infer that e^o must be strictly positive. Suppose $e^o = 0$ and thus $\bar{q}^o = q_d^*$. Increasing e slightly generates only second-order losses (in terms of duopoly profit, and also effort costs since, by assumption, $\psi'(0) = 0$) but first-order gains (in terms of monopoly profit). Thus, we have the result that $0 < e^o \leq 1$, which means that, in spite of the inefficiency caused by nonresponsiveness when overlap is noncontractible, field-of-use restrictions are still preferred to nonexclusive licenses. Inducing the licensees to produce the optimal quantity can most easily be achieved through an output restriction. For other distortive royalty schemes, such as a constant per-unit royalty, monopoly output would be reduced to a level strictly below duopoly output, which is undesirable.

3.3 Overlap noncontractible and effort unobservable

We now turn to the case where the level of care exercised by the patent holder when drafting the field-of-use clauses is unobservable. What consequence does this have on the optimal contract? Notice first that the patent holder has no incentive to exert effort unless his royalty income (as opposed to the fixed fees which are paid upfront, before the realization of overlap) depends positively on e. Thus, provision of effort is incompatible with the bunching scheme that is optimal under observability. Of course, the licensees will anticipate this and, if offered such a scheme, will be willing to pay only their expected profit given $e = 0$ to obtain a license. Second, recall that implementability requires that monopoly output not exceed duopoly output (Lemma 3). Hence, a scheme that induces effort must be such that output in case of overlap is strictly greater than in the no-overlap case. Third, the royalty payment associated with the no-overlap quantity must exceed the payment in case of overlap. We can thus guess that the relevant incentive constraint will be the one for the no-overlap "type" who may be tempted to mimic the overlap "type".

The optimization problem is

$$\max_{e;F;(q(m),r(m));(q(d),r(d))} \quad F + e\,r(m) + (1-e)\,r(d) - \psi(e)/2$$

subject to

$$\begin{aligned} F &\leq e^*\big[\Pi_m\big(q(m)\big) - r(m)\big] + (1-e^*)\big[\Pi_d\big(q(d), q(d)\big) - r(d)\big] && (12)\\ \Pi_m\big(q(m)\big) - r(m) &\geq \Pi_m\big(q(d)\big) - r(d) && (13)\\ \Pi_d\big(q(d), q(d)\big) - r(d) &\geq \Pi_d\big(q(m), q(d)\big) - r(m) && (14)\\ \Pi_d\big(q(d), q(m)\big) - r(d) &\geq \Pi_d\big(q(m), q(m)\big) - r(m). && (15) \end{aligned}$$

where e^* is the equi1librium level of effort (rationally anticipated by the licensees). Unlike in the previous section, the menu of outputs and royalty payments now has to accomplish two things: organizing production efficiently, and providing incentives to the licensor to exert effort. While the problem of nonresponsiveness persists, it may now sometimes be optimal to induce separation, precisely in order to incentivize the licensor. To do so, $q(d)$ must be raised above $q(m)$, which is inefficient. Thus, the parties may accept to sacrifice some efficiency in exchange for higher effort to avoid overlap. However, this can only be optimal when the cost of effort is not too important (or more precisely, when it doesn't increase too rapidly with e, at least for small values of e). Otherwise, a bunching contract will be preferred. Bunching, meanwhile, means that no effort can be sustained, so that the optimal solution becomes a nonexclusive license given to both firms; that is, the license contract no longer features field-of-use restrictions. The following proposition characterizes the optimal contract assuming $\psi(\cdot)$ is such that an interior solution exists and dominates a contract with pooling at q_d^*.

Proposition 3 *Assume that overlap is noncontractible and effort unobservable. Assume also that the cost of effort, $\psi(e)$, is sufficiently convex and that $\psi(1)/2 > \pi_m^* - \pi_d^*$. Then, the optimal contract $\big(e^u, F^u, (q^u(m), r^u(m)), (q^u(d), r^u(d))\big)$ has the following features:*

(i) *e^u is defined by $\psi'(e^u) = 2\big[\pi_m^* - \Pi_d\big(q^u(d), q^u(d)\big)\big]$;*

(ii) *$F^u = e^u[\pi_m^* - r^u(m)] + (1-e^u)\big[\Pi_d\big(q^u(d), q^u(d)\big) - r^u(d)\big]$;*

(iii) *$q^u(m) = q_m^*$ and $q^u(d)$ is determined by*

$$\frac{\partial e(q_m^*, q(d))}{\partial q(d)} \left[\pi_m^* - \hat{\Pi}_d(q(d)) - \frac{\psi'\big(e(q_m^*, q(d))\big)}{2} \right] + \big[1 - e(q_m^*, q(d))\big] \frac{d\hat{\Pi}_d(q(d))}{dq} = 0$$

where $e(q(m), q(d)) \equiv (\psi')^{-1}\big[2\big(\Pi_m(q(m)) - \Pi_m(q(d))\big)\big]$ and $\hat{\Pi}_d(q) \equiv \Pi_d(q, q)$.

(iv) *$(r^u(m), r^u(d))$ satisfy $r^u(m) - r^u(d) = \pi_m^* - \hat{\Pi}_d\big(q^u(d)\big)$.*

This can be implemented with a royalty scheme $R(\cdot)$ which is such that $R'(q_m^*) = 0$ and $R(q_m^*) = r^u(m)$ while $\partial \Pi_d\big(q^u(d), q^u(d)\big)/\partial q_i = R'(q^u(d))$ and $R(q^u(d)) = r^u(d)$.

Proof: We proceed as follows. Based on the above discussion and earlier results, we ignore the second incentive compatibility constraint (14) and the uniqueness condition (15) and guess that (12) and (13) are binding. We derive the solution and then check whether a) it satisfies the implementability condition $q(d) \geq q(m)$, and b) whether it makes the licensor's payoff higher than with the non-exclusive license (bunching) solution $\big(e = 0, q(m) = q(d) = q_d^*\big)$, which delivers a payoff of π_d^*.

Note first that equilibrium effort is determined by

$$r(m) - r(d) = \psi'(e)/2.$$

Using (13), we have

$$e^* = (\psi')^{-1}\big[2\big(\Pi_m(q(m)) - \Pi_m(q(d))\big)\big] \equiv e(q(m), q(d)).$$

Using (12) and the fact that expectations concerning e must be correct in equilibrium, the program becomes

$$\max_{q(m), q(d)} \; e(q(m), q(d))\Pi_m(q(m)) + \big[1 - e(q(m), q(d))\big]\Pi_d(q(d), q(d)) - \psi\big(e(q(m), q(d))\big)/2.$$

Letting $\hat{\Pi}_d(q) \equiv \Pi_d(q, q)$, so that $d\hat{\Pi}_d/dq = \partial \Pi_d/\partial q_i + \partial \Pi_d/\partial q_j$, and let $e_m \equiv \partial e/\partial q(m)$ and $e_d \equiv \partial e/\partial q(d)$, the first-order conditions are:

$$\begin{aligned} \frac{\partial e}{\partial q_m} \left[\Pi_m(q(m)) - \hat{\Pi}_d(q(d)) - \frac{\psi'\big(e(q(m), q(d))\big)}{2} \right] + e(q(m), q(d)) \frac{d\Pi_m(q(m))}{dq} &= 0 \\ \frac{\partial e}{\partial q(d)} \left[\Pi_m(q(m)) - \hat{\Pi}_d(q(d)) - \frac{\psi'\big(e(q(m), q(d))\big)}{2} \right] + \big[1 - e(q(m), q(d))\big] \frac{d\hat{\Pi}_d(q(d))}{dq} &= 0. \end{aligned}$$

Noticing that $\frac{\partial e}{\partial q_m} = \frac{d\Pi_m(q(m))/dq}{\psi''\left[(\psi')^{-1}\left[2\left(\Pi_m(q(m))-\Pi_m(q(d))\right)\right]\right]}$, the first equation yields $q(m) = q_m^*$. Thus, $q(m)$ is independent of $q(d)$ and we can replace it in the second equation which becomes

$$\frac{\partial e(q_m^*, q(d))}{\partial q(d)} \left[\pi_m^* - \hat{\Pi}_d(q(d)) - \frac{\psi'\big(e(q_m^*, q(d))\big)}{2}\right] + \left[1 - e(q_m^*, q(d))\right] \frac{d\hat{\Pi}_d(q(d))}{dq} = 0.$$

Depending on the precise functional form of $\psi(\cdot)$, this can have several solutions, one of which typically is a local maximum that is not implementable (i.e. smaller than q_m^*), and (at least) one a local minimum; both can be discarded. If ψ is sufficiently convex, there also exists a local maximum to the right of q_m^*, which is a candidate for the optimal $q(d)$; let us call it $q^u(d)$. A second candidate is a corner solution that consists in increasing $q(d)$ up to the point where $e(q_m^*, q(d)) = 1$; this, however, can be the optimal $q(d)$ only if $\psi(1)/2 \leq \pi_m^* - \pi_d^*$, which we have ruled out by assumption. Therefore, we are left with only one candidate, $q^u(d)$, which can be shown to deliver a payoff larger than π_d^* for $\psi(\cdot)$ sufficiently convex and not too large.[23] ■

Proposition 3 shows under some assumptions on $\psi(e)$ that, when effort is unobservable, the optimal royalty scheme is no longer trivial and features royalties that are decreasing with output over some range. There are a number of other interesting observations. First, we get a variant of the well-known "no distortion at the top" result: In the absence of overlap, the contract provides for the efficient output, q_m^*. This is natural since a separating contract aims at incentivizing the licensor. The binding incentive constraint (13) means that effort is increasing in the difference between the no-overlap profits at $q(m)$ and $q(d)$. This difference is best maximized by setting $q(m)$ at the profit-maximizing level.

Second, there is a tradeoff between providing incentives to the licensor and producing at the efficient scale. In fact, while raising $q(d)$ relaxes the incentive constraint and thereby induces greater effort, it also moves duopoly output away from the efficient level, q_d^*. The intuition for why this can be optimal is the following. Starting from q_m^*, moving $q(d)$ upwards leads to an increase in effort and to a decrease in Π_d. Because the increase in effort is second order at $q(m)$, while the loss in Π_d is first order, this reduces the licensor's payoff for small deviations from $q(m)$. However, if ψ doesn't increase too fast, a small difference between $q(d)$ and $q(m)$ translates into a large rise in effort (we are in the flat

[23] Details are available from the author upon request.

part of the cost curve). Therefore, e quickly grows sufficiently large to compensate the losses incurred at the beginning. As $q(d)$ continues to increase, e enters the steeper part of the ψ curve, and eventually the negative effect on Π_d again comes to outweigh the positive effect on effort, so that the licensor's payoff peaks at some quantity – namely, $q^u(d)$.

There are some degrees of freedom in the choice of $r^u(m)$ and $r^u(d)$. What is important is the difference between them. The fixed fee can again be used to extract the remaining (expected) surplus from the licensees.

4 Discussion

Two corollaries of the results obtained in the case where overlap is noncontractible and effort is unobservable are the following. First, the licensor cannot be given incentives to exert effort with a linear (constant per-unit) royalty. Second, if the licensor measures only production and not sales, and there is free disposal, the monopolist licensee will produce the duopoly quantity to pay lower royalties, and then sell only the (lower) monopoly quantity. Thus, the above royalty scheme is not incentive compatible, implying that effort is not sustainable in equilibrium. Since this reduces the attractiveness of using FOUR in the first place, we might expect them to be used less when sales are hard to measure relatively to production.

From a more general perspective, the extent to which effort is observable can be interpreted as a measure of the information gap between licensor and licensees, and/or the complexity of the patented technology. Along these lines, the model predicts that as the technology becomes more complex, license contracts with FOUR should make greater use of royalties.

Our results are also roughly consistent with the observation by Taylor and Silberston (1973) and Kamien (1992) that royalty rates often decrease with output,[24] at least when these royalty rates are interpreted as average rates. Since in case of exclusivity, the royalty payment is greater and the output lower than in case of overlap, the model predicts that implied royalty rates per unit of output decrease with the quantity produced.[25]

[24] Kamien (1992, p. 346) states that "in the case of licensing the sale of a new product, the patentee often offers a lower royalty rate if sales exceed a certain prespecified level."

[25] There are, of course, alternative explanations for decreasing royalty rates. The patentee may want to incentivize the licensee to push sales above those of other products (Kamien, 1992). Private information on the part of licensees concerning the value of the patented technology may also lead to the observed

To conclude, we briefly discuss some implications of our results for antitrust authorities. A direct implication of the model is that quantity restrictions and non linear royalty schemes should not be considered per se violations of the antitrust laws when used in conjunction with FOUR. In the framework we examine, they lead to improvements in terms of production efficiency. One qualification that needs to be made in this respect is the possibility that FOUR may be misused to restrain competition between firms that would otherwise be horizontal competitors. If the licensed technology represents only a minor improvement, competition is harmed by the agreement without any offsetting benefits in terms of higher quality. However, a simple rule that antitrust authorities could use to determine whether an agreement is likely to be welfare-enhancing is to check whether the licensees continue to sell the original product that does not incorporate the improved technology. In this way, if the improvement is minor, the improved product will not sell if its price greatly exceeds the price of the original product, as they will be close substitutes. And since the original product is supplied competitively, welfare is not likely to be harmed by the agreement.

pattern (Beggs, 1992).

Appendix

A direct revelation mechanism that implements first best

Assume that only the licensees learn the state of the world (i.e., the realization of overlap) at date $t = 1$ of the game. Consider the following direct revelation mechanism: Both licensees are asked to report the state of the world s. The allocation that results depends on both licensees' reports and is determined by the following table where each cell contains a quadruplet $(q_i, r_i), (q_j, r_j)$, that is, quantity-transfer pairs for licensees L_i (row player) and L_j (column player):

	m	d
m	$(q_m^*, \pi_m^* - \pi_d^*), (q_m^*, \pi_m^* - \pi_d^*)$	$(\hat{q}, \hat{r}), (\tilde{q}, \tilde{r})$
d	$(\tilde{q}, \tilde{r}), (\hat{q}, \hat{r})$	$(q_d^*, 0), (q_d^*, 0)$

Table 1: Allocations resulting from the message game

Can we find values for $(\hat{q}, \hat{r}, \tilde{q}, \tilde{r})$ such that truthtelling is the unique Nash equilibrium of the game? For a truthful equilibrium to exist, it must be the case that

$$\Pi_m(q_m^*) - (\pi_m^* - \pi_d^*) \geq \Pi_m(\tilde{q}) - \tilde{r} \tag{16}$$

$$\Pi_d(q_d^*, q_d^*) \geq \Pi_d(\hat{q}, \tilde{q}) - \hat{r}. \tag{17}$$

Moreover, in order not to get a nontruthful equilibrium, the following constraints must be satisfied:

$$\Pi_m(q_d^*) < \Pi_m(\hat{q}) - \hat{r} \tag{18}$$

$$\Pi_d(q_m^*, q_m^*) < \Pi_d(\tilde{q}, \hat{q}) - \tilde{r}. \tag{19}$$

Let us set $\hat{q} = q_d^* + \varepsilon$, where ε is arbitrarily small, and $\hat{r} = 0$. This satisfies (18), and also ensures that (17) is satisfied for any $\tilde{q} \geq q_d^*$. Adding up the remaining two constraints (16) and (19), and noting that $\Pi_m(q_m^*) = \pi_m^*$ and $\Pi_d(q_d^*, q_d^*) = \pi_d^*$, we obtain the following condition:

$$\Pi_m(\tilde{q}) - \pi_d^* \leq \tilde{r} < \Pi_d(\tilde{q}, q_d^*) - \Pi_d(q_m^*, q_m^*).$$

For such an $\tilde{r}$ to exist, it must be the case that, for some $\tilde{q} \geq q_d^*$,

$$\Pi_m(\tilde{q}) - \Pi_d(\tilde{q}, q_d^*) < \pi_d^* - \Pi_d(q_m^*, q_m^*).$$

The right-hand side of this expression is necessarily strictly positive. Thus, it suffices to find a $\tilde{q}$ such that

$$\Pi_m(\tilde{q}) = \Pi_d(\tilde{q}, q_d^*). \tag{20}$$

We show that such a $\tilde{q}$ exists for the case $c_H < (2a + 3c_L)/5$, i.e., where both licensees are involved in production. Solving (20), we obtain

$$\tilde{q} = (2a + 3c_H - 5c_L + \sqrt{A})/8,$$

where $A = (2a + 11c_H - 13c_L)(2a + 3c_H - 5c_L)$. Given our assumptions, A is sure to be positive, and thus $\tilde{q}$ exists.

References

Anand, Bharat N. and Khanna, Tarun (2000): The Structure of Licensing Contracts. *Journal of Industrial Economics*, Vol. 48(1), pp. 103–135.

Bajari, Patrick and Tadelis, Steven (2001): Incentives versus Transaction Costs: A Theory of Procurement Contracts. *RAND Journal of Economics*, Vol. 32(3), pp. 387–407.

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

Bessy, Christian and Brousseau, Eric (1998): Technology Licensing Contracts: Features and Diversity. *International Review of Law and Economics*, Vol. 18, pp. 451–489.

Bousquet, Alain, Cremer, Helmuth, Ivaldi, Marc, and Wolkowicz, Michel (1998): Risk Sharing in Licensing. *International Journal of Industrial Organization*, Vol. 16, pp. 535–554.

Caves, Richard E., Crookell, Harold, and Killing, J. Peter (1983): The Imperfect Market for Technology Licenses. *Oxford Bulletin of Economics and Statistics*, Vol. 45, pp. 249–267.

Cestone, Giacinta and White, Lucy (2003): Anticompetitive Financial Contracting: The Design of Financial Claims. *Journal of Finance*, Vol. 58(5), pp. 2109–2141.

Choi, Jay Pil (2001): Technology Transfer With Moral Hazard. *International Journal of Industrial Organization*, Vol. 19, pp. 249–266.

Contractor, Farok J. (1981): *International Technology Licensing: Compensation, Costs, and Negotiation*. Lexington, MA: Lexington Books.

Dewatripont, Mathias (1992): Implementation Theory: Discussion. In: J.J. Laffont (ed.), *Advances in Economic Theory. Sixth World Congress*, Vol. 1, pp. 324–327. Cambridge: Cambridge University Press.

Dye, Ronald A. (1985): Costly Contract Contingencies. *International Economic Review*, Vol. 26(1), pp. 233–250.

Gallini, Nancy T. and Wright, Brian D. (1990): Technology Transfer Under Asymmetric Information. *RAND Journal of Economics*, Vol. 21(1), pp. 147–160.

Gilbert, Richard and Shapiro, Carl (1997): Antitrust Issues in the Licensing of Intellectual Property: The Nine No-Nos Meet the Nineties. In: *Brookings Papers on Economic Activity: Microeconomics*, pp. 283–336. Washington, DC: The Brookings Institution.

Guesnerie, Roger and Laffont, Jean-Jacques (1984): A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm. *Journal of Public Economics*, Vol. 25, pp. 329–369.

Hart, Oliver and Tirole, Jean (1990): Vertical Integration and Market Foreclosure. *Brookings Papers on Economic Activity*, pp. 205–276. Special issue.

Kamien, Morton I. (1992): Patent Licensing. In: R.J. Aumann and S. Hart (eds.), *Handbook of Game Theory with Economic Applications*, pp. 331–354. North Holland: Elsevier.

Kamien, Morton I. and Tauman, Yair (1986): Fees versus Royalties and the Private Value of a Patent. *Quarterly Journal of Economics*, Vol. 101(3), pp. 471–492.

Katz, Michael L. (1989): Vertical Contractual Relationships. In: R. Schmalensee and R.D. Willig (eds.), *Handbook of Industrial Organization*, Vol. 1. North Holland: Elsevier.

Katz, Michael L. and Shapiro, Carl (1986): How to License Intangible Property. *Quarterly Journal of Economics*, Vol. 101(3), pp. 567–590.

Maskin, Eric (1977): Nash Implementation and Welfare Optimality. Mimeo, Massachusetts Institute of Technology.

Mathewson, G. Frank and Winter, Ralph A. (1984): An Economic Theory of Vertical Restraints. *RAND Journal of Economics*, Vol. 15(1), pp. 27–38.

Rey, Patrick and Tirole, Jean (1986): The Logic of Vertical Restraints. *American Economic Review*, Vol. 76(5), pp. 921–939.

Rey, Patrick and Tirole, Jean (2007): A Primer on Foreclosure. In: M. Armstrong and R. Porter (eds.), *Handbook of Industrial Organization*, Vol. 3, chap. 7. North Holland: Elsevier.

Rey, Patrick and Vergé, Thibaud (2007): The Economics of Vertical Restraints. In: P. Buccirossi (ed.), *Handbook of Antitrust Economics*. Cambridge, MA: MIT Press.

Rostoker, Michael D. (1984): A Survey of Corporate Licensing. *IDEA: The Journal of Law and Technology*, Vol. 24, pp. 59–92.

Segal, Ilya (1999): Contracting with Externalities. *Quarterly Journal of Economics*, Vol. 114(2), pp. 337–388.

Spengler, Joseph J. (1950): Vertical Integration and Anti-trust Policy. *Journal of Political Economy*, Vol. 58(4), pp. 347–352.

Spier, Kathryn E. (1992): Incomplete Contracts and Signalling. *RAND Journal of Economics*, Vol. 23(3), pp. 432–443.

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