The Journal of Industrial Economics

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Lurking Patent Claims and Strategic Royalty Contracts

Jay Pil Choi[1,2]

[1]Department of Economics, Michigan State University, East Lansing, Michigan, USA | [2]School of Economics, Yonsei University, Seoul, South Korea

Correspondence: Jay Pil Choi (choijay@msu.edu)

Accepted: 19 May 2026

Keywords: licensing contracts | lurking patents | patent assertion entities | rent-protection device | royalty rate and fixed fees

ABSTRACT

This paper analyzes optimal licensing contracts when a licensee faces the risk of future infringement claims by unknown patent holders. In a setting where a noncompeting licensor contracts with a monopolistic manufacturer, fixed-fee licensing is optimal absent such claims. We show that the possibility of future patent claims makes it optimal to include a per-unit royalty. The royalty reduces the surplus available to potential third-party claimants and serves as a rent-protection device, trading off allocative inefficiency against hold-up risk. Restrictions on the contracting space can be justified when more realistic features of demand uncertainty or stochastic patent arrival are considered.

JEL Classification: D43, L13, O3

1 | Introduction

This paper analyzes the optimal licensing contract when a licensee faces the possibility of future infringement claims by unknown patent holders. We consider a simple environment in which the patent holder is an outsider that does not compete with a monopolistic licensee. In such a setting, it is well known that the optimal contract consists solely of a fixed fee in the absence of future infringement risk. However, we show that once the possibility of unknown future patent claims is introduced, the optimal contractual form entails a positive per-unit royalty. This provides a new rationale for the prevalence of royalty contracts in practice and yields important managerial and policy implications.

To motivate our analysis, we can imagine a situation in which a relevant patent unknown to the licensee is filed and pending in the application process and then issued after the licensee has introduced a product that implements the patented technology. In fact, prior to 1995, when the U.S. signed the TRIPS agreement of the WTO, which changed the patent term from 17 years from the date of issuance to 20 years from the filing date, inventors routinely engaged in strategic delay of patent issuance until a commercial product that implements the patented technology appeared in the market. The strategic use of the so-called "submarine patent"—a patent filed on device or technology that does not exist yet or has not been implemented, and whose issuance is deliberately delayed—also was facilitated under the old US patent system that kept the filing information confidential until issuance [1]. The American Inventors Protection Act (AIPA) enacted in 1999 (37 C.F.R. 1.211) now mandates that the United States Patent and Trademark Office publish patent applications 18 months after they are filed. This new disclosure requirement in the U.S. patent law contributed to lifting the veil of secrecy associated with patent filings and eliminated incentives for "submarine patents" to some extent.[1]

However, this does not imply that licensees are cognizant of all relevant patents and immune from surprises by "lurking patents" [3]: the prospect of unknown infringement claims still looms large for any firm that produces a new product, as we discuss in more detail in the next section.

Absent future infringement risk, the optimal contract between a noncompeting research firm and a monopolistic licensee consists solely of a fixed fee. A royalty distorts the licensee's marginal cost and introduces double marginalization, reducing total surplus. Because the fixed fee can be used to divide the surplus according to bargaining power, the parties jointly maximize total surplus by setting the royalty rate to zero.[2]

The presence of uncertain future patent claims fundamentally alters this result. We show that including a positive royalty becomes optimal because it reduces the surplus available to a potential future patent claimant holding a complementary patent that the product may also infringe upon. The royalty thus serves as a *rent-protection* device for the initial contracting parties. The optimal royalty rate balances the allocative inefficiency it creates against the reduction in surplus that would otherwise be subject to ex post holdup. The optimal royalty rate is increasing in the probability and timing weight of future claims and decreasing in demand elasticity.

Our paper shares a common structure with Aghion and Bolton [4], which analyzes exclusive dealing contracts in the presence of an entry threat: two parties move first by signing an initial contract containing a strategic term that reduces the surplus available to a third party that may appear later. However, the mechanism in our paper differs importantly from that in Aghion and Bolton [4]. In their model, an incumbent and a buyer form a coalition to extract rents from a potential entrant supplying *substitute* products; exclusivity functions as a rent-extraction device. In contrast, in our model, the initial patent holder and the licensee form a coalition to protect rents from a future *complementary* patent holder. The royalty operates as a rent-protection mechanism rather than a rent-extraction mechanism. The distinction arises because the technologies in our baseline setting are complements: the manufacturer requires both patents.[3]

We also show that the main results extend to settings with demand uncertainty and stochastic patent arrival. Incorporating these more realistic features provides a justification for restricting the contracting space to noncontingent royalty rates, as more sophisticated state-contingent contracts are difficult to verify and implement in practice.

This paper contributes to the literature on the optimal patent licensing contract.[4] Early contributions to this literature show that licensing through a fixed fee or an auction is more profitable for an independent R&D only licensor than licensing through a royalty rate (Kamien and Tauman [5, 6], Katz and Shapiro [7]).[5] In practice, licensing contracts often include a royalty component.[6] There are several papers that explain the royalty contract as an optimal contractual form in the presence of market imperfections. For instance, Gallini and Wright [11] and Beggs [12] provide an analysis of licensing in the presence of asymmetric information. They consider a situation in which the licensor has superior information about the value of the technology and explain the royalty rate in the contract as a signaling device for the party that has better information about the value of the licensed technology. Choi [13] considers a licensing environment with moral hazard, in which the effective transmission of knowledge requires costly inputs by both dispensing and receiving parties. He shows that the introduction of inputs that are not contractible and costly explains the prevalence of royalty contracts in the licensing relationship.[7] In all these models, royalty contracts are a response to imperfect contractual environments and a second-best solution to asymmetric information and/or moral hazard problems.[8] Positive royalties can arise under uncertainty in other contexts, such as risk aversion or limited liability [17]. In our model, however, the licensor and the licensee share symmetric information and are risk-neutral. The mechanism in this paper is therefore distinct: the royalty emerges as a strategic device (without any positive effects) to protect surplus from a potential future patent claimant, rather than as a response to asymmetric information or risk-sharing considerations.

Amir et al. [18] analyze a patent holder's choice of licensing mechanism when the validity of the patent is uncertain and subject to potential litigation. They derive sufficient conditions under which a per-unit royalty contract is optimal, showing that these conditions hold across a wide range of environments. Their central insight is that the comparison between fixed-fee and per-unit royalty licensing depends on the balance between two opposing forces, which they term the "efficiency effect" and the "punishment effect." In their setting with multiple competing licensees, per-unit royalties soften competition among licensees, increasing the total profit available to the patent holder.[9] This "efficiency effect" favors the use of per-unit royalties. Conversely, under fixed-fee licensing, a firm that unsuccessfully challenges the patent earns lower profits, which reduces the incentive to litigate against the patent holder. This "punishment effect" favors fixed-fee contracts.[10] Per-unit royalties are preferred when the efficiency effect dominates the punishment effect. While Amir et al. [18] and this paper both demonstrate conditions under which including per-unit royalties is optimal, the underlying mechanisms are fundamentally different. In Amir et al., both the efficiency and punishment effects arise exclusively from the presence of multiple licensees; when there is only a single licensee, these effects vanish. In contrast, the optimality of per-unit royalties in our model emerges even with a single licensee. Moreover, Amir et al. [18] identify a per-unit royalty-only contract as optimal, whereas our model yields a two-part tariff combining a fixed fee with a per-unit royalty due to the distinct underlying rationale. Finally, the differing underlying mechanisms for per-unit royalty contracts in the two papers give rise to opposite comparative static results. While Amir et al. [18], in their extension with a two-part tariff, find that per-unit royalties emerge when the patent is weak, our model predicts an increase in the royalty rate as the likelihood of future patent infringement claims rises.

Our paper is also closely related to Hovenkamp et al. [3], as both studies examine strategic implications for a firm operating amid uncertainty about relevant patents. Hovenkamp et al. [3] explore the patentee's incentives to "lurk," potentially engaging in ex post patent holdup if developers inadvertently infringe on their patents. This possibility can impact developers' technology adoption decisions. In contrast, our paper focuses on the optimal form of licensing contracts with known licensors, specifically under the risk posed by such lurking patent holders. Their papers and ours address different facets of strategic behavior, complementing each other in offering valuable insights and managerial implications regarding the uncertainties surrounding lurking patents.

Finally, our paper has a connection to the literature on royalty stacking and patent pools [21, 22]. While there are superficial similarities particularly the presence of multiple complementary patent holders the mechanism in this paper differs in an important respect. In the royalty stacking literature, inefficiencies arise from multiple patent holders independently setting royalties. In contrast, in our model, the royalty in the initial contract is chosen strategically by the first licensor and the manufacturer to reduce the surplus available to a potential future claimant. The result is therefore driven by a rent-protection motive rather than by uncoordinated pricing among multiple patent holders.

The remainder of the paper is organized in the following way. Section 2 briefly describes the risk of lurking patents faced by manufacturers to motivate our analysis. In Section 3, we develop a simple model of sequential patent licensing that demonstrates the optimality of licensing contracts including a per-unit royalty component. Section 4 extends the analysis and examines the robustness of the main results by incorporating demand uncertainty, the stochastic arrival of future patent claimants, endogenous incentives for unknown patent holders to "lurk," substitute patents, and asymmetric bargaining power between the patent holders and the licensee. Section 5 concludes the paper with final remarks and a discussion of the policy implications of our analysis.

2 | "Lurking Patents": Institutional and Industry Background

To motivate our analysis of a setting in which a licensee faces the risk of future infringement claims by unknown patent holders—despite disclosure requirements under U.S. patent law—we explain why such reforms do not eliminate the threat of "lurking patents." Licensees may remain vulnerable to unknown or strategically concealed patent claims because patent holders often have other strategic tools at their disposal to keep potential infringers unaware of relevant patents until the industry becomes dependent on the technology, thereby maximizing their bargaining leverage and potential payouts.

First, patent applicants still have the option of keeping their patent filings secret if they can certify their intention not to seek a patent outside the US. Second, a series of "continuation applications" is another commonly used tactic to engineer a delay through which the claims of the patent can be modified and refined to closely align with the new technology as it develops in the market.[11] Third, even if all patent applications are published, there are simply too many patents to check to avoid any infringement risks. Recent years have seen a dramatic increase in the number of patent applications and patents granted as a result of firms amassing vast patent portfolios [24]. The sheer number of patents held by other firms makes it impractical for firms to develop new products that avoid inadvertent infringement on other firms' patent portfolios with certainty.[12]

For instance, Cotripia and Lemley [26] report that only a very small fraction of patent infringement cases involve defendants who have copied the patented technology, implying that most cases entail inadvertent infringement. Bessen and Meurer [27] also provide empirical evidence suggesting that most defendants in patent litigation are inadvertent infringers rather than firms attempting to copy or invent around patents. This type of situation is particularly pertinent in many high-tech industries where technologies are rapidly advancing and draw upon existing stocks of knowledge. The convergence of digital media and the emergence of the Internet have also blurred the boundaries of the previously separate information and communication technology (ICT) industries. As a result, the development of new products in the ICT industry often requires access to and integration of numerous complementary technologies, as illustrated by smartphones that employ a variety of technologies in the areas of wireless communication, GPS, camera, digital technology, high-speed broadband, and so on.

The semiconductor industry provides another example of an industry that "requires access to a 'thicket' of intellectual property rights in order to advance the technology or to legally produce or sell" new products [28]. This issue is especially problematic for software patents because there is no easily searchable database for software ideas. For instance, Mulligan and Lee [29] estimate that "it would require roughly 2 million patent attorneys, working full-time, to compare every firm's products with every patent issued in a given year" to avoid accidental infringement, which makes eliminating unintentional infringement impractical.

The emergence of patent assertion entities (PAEs) and the business model of "patent privateering" also adds to this type of uncertainty. PAEs exploit secrecy to ambush unsuspecting firms with surprise claims against them. Intellectual Ventures, Acadia, and many others have assigned their patents to thousands of shell companies and subsidiaries, making them hard to track [30]. For instance, Ewing and Feldman [31] identified 1276 shell companies created by Intellectual Ventures. It is their explicit strategy to wait until practicing entities develop products that infringe on their patent portfolios to "surprise them with a suit." High-profile lawsuits, such as Apple's antitrust case against Acacia and other PAEs associated with Nokia, also highlight how the risk and cost of future patent infringement claims may affect a company's strategy. Many of Nokia's patents are used in Apple's products, and Apple has paid Nokia a royalty in order to use the patents. In its case filed on December 20, 2016, Apple claimed that Nokia had transferred patents to third-party companies, including Acacia and Conversant, with the intent to extract excessive royalties from Apple and other patent users. Nokia responded by filing lawsuits against Apple for violating 32 technology patents. The Apple vs. Nokia case demonstrates not only the extent to which Apple is concerned with facing unexpected patent infringement claims, but also the sophistication with which companies may use patents and infringement claims to extract royalties from potential licensees.

The cases discussed above illustrate how lurking patent strategies can impose unexpected costs and disruptions on companies that adopt technologies without full knowledge of all potential patent claims, highlighting the risks posed by lurking patents for manufacturers and motivating our analysis.[13]

3 | Model

Consider a licensing contract between a patent owner ($P1$) and a manufacturing firm (M). Let us assume that the licensee is a monopolist in a market. The monopolist has a constant marginal production cost of c. The monopolist's product infringes on the patent holder's technology and it needs a license from the patent holder to commercialize the product. We, however, entertain the possibility that in the future another patent holder ($P2$) may show up and claim infringement on his technology. Let us assume that the probability of this event is given by β, which is shared by both $P1$ and the licensee. Note that we consider a setting where M's product is highly sophisticated, incorporating numerous unrelated or complementary technologies. The patents held by $P1$ and $P2$ cover distinct, nonoverlapping aspects of the product, so there is no dispute between $P1$ and $P2$. Patent validity is also not in question in our setup. Rather, the potential licensee in the first period is simply uncertain about other patents it may inadvertently infringe upon.

Both patent holders are nonpracticing entities: neither $P1$ nor $P2$ has the ability to commercialize the product that incorporates patented technologies themselves in the relevant market we analyze. For instance, they could be a pure R&D firm without any manufacturing capabilities. Alternatively, they could be a foreign firm that prefers to license its technology rather than entering the domestic market directly. We assume that the patent holders and the monopolist have equal bargaining power and thus share any surplus from the bargaining equally.

3.1 | Benchmark Model Without Future Patent Claimant

To highlight the implications of uncertain future infringement claims by another party on the current licensing contract, we first characterize the optimal linear licensing contract in the absence of such uncertainty. This linear contract, which is the most commonly observed form, comprises two components: a per-unit royalty rate and a lump-sum payment.

Let $\pi(c)$ and $q(c)$ be the (reduced) profit function for the licensee and the corresponding profit-maximizing level of output. In the benchmark case, we derive the standard result that the optimal contract is characterized by fixed fee only contracts.

Lemma 1. *The optimal contract between the patent holder and the monopolist entails a lump-sum payment only contract without any per-unit royalties.*

Proof. Let $P(.)$ be the inverse demand function and $q(c)$ be the profit maximizing level of output when the marginal cost is given by c, that is, $q(c) = argmax[P(q)q - cq]$ and satisfies the first-order condition $P'q + P = c$. Note that by the envelope theorem, $\pi'(c) = -q(c)$. With the royalty rate of r, the licensee's $MC = c + r$. Therefore, the licensee will produce $q(c + r)$ and receive profits of $\pi(c + r)$. As a result, the royalty income is $rq(c + r)$. The remaining profit of the licensee can then be divided between the patent holder and the licensee through a lump-sum payment determined by Nash bargaining. Under equal bargaining power, the Nash bargaining solution satisfies

$$\pi(c + r) - F = rq(c + r) + F.$$

Solving for the fixed fee yields $F = \frac{\pi(c+r) - rq(c+r)}{2}$. The licensor and the licensee will choose r to maximize the joint surplus (S), that is, to maximize the total licensing income $L = \frac{S}{2}$, where $S = \pi(c + r) + rq(c + r)$. The first order condition is given by

$$\frac{dS}{dr} = \pi'(c + r) + q(c + r) + rq'(c + r)$$
$$= -q(c + r) + q(c + r) + rq'(c + r) = rq'(c + r) = 0$$

Thus, the optimal $r = 0$ because $q'(c + r) < 0$. This, in turn, implies that $F = \frac{\pi(c)}{2}$. □

The intuition for Lemma 1 is easy to understand. The per-unit royalty rate determines the licensee's marginal cost and its optimal output decision. As a result, the per-unit royalty affects the size of the pie available between the two contracting parties, whereas the lump-sum component is just a pure transfer between them. The optimal contract requires no royalty rate because any royalty rate artificially distorts the licensee's marginal cost and reduces the size of the available pie. This result is not dependent on any particular bargaining protocol assumptions. As long as the patent holder and the licensee bargain with symmetric information and do not directly compete with each other in the same market, the optimal royalty rate is zero. The optimality of zero royalty rate would still hold even if there are multiple patent owners, and they all bargain together simultaneously. When the surplus is split between multiple patent owners and the licensee, they will maximize the available surplus, which entails zero per-unit royalty rate.

3.2 | Baseline Model

We now consider a licensing game in which both the patent owner and the licensee recognize the possibility that the licensee's product may infringe on another patent holder's intellectual property. To analyze this scenario, we consider a two-period model in which the initial negotiation takes place in the first period, but another patent holder may emerge and assert an infringement claim in the second period. The relative weight of the second-period payoff compared to the first is parameterized by δ. When the two periods are of equal length, δ can be interpreted as a discount factor. However, in our context, δ may also capture the relative duration of the second period compared to the first, so it is not necessarily less than one.

The timing of the game is as follows: In the first period, M and $P1$ (whose patent is widely known) negotiate a licensing contract. In the second period, another patent holder ($P2$) appears with the probability of β, in which event M and $P2$ negotiate another bilateral licensing contract given the original contract between M and $P1$ as *given*. If $P2$ does not exist, then the market outcome is the same as in the first period.

In the baseline model analysis of this section, we assume that the initial licensing contract between M and $P1$ cannot specify different royalty rates across periods. In fact, as the analysis below will illustrate, if different royalty rates could be set across periods, the optimal contract would feature a zero royalty rate in the first period, and impose a positive royalty rate only in the second period when $P2$ may appear. However, the feasibility of such a contract depends critically on an artificial two-period, discrete-time framework, where $P2$'s possible entry time is certain and exogenously given. If $P2$'s arrival is uncertain and stochastic, then defining a "second period" based on $P2$'s entry becomes problematic, rendering such a contract infeasible. Although this assumption may seem restrictive in the two-period model, it becomes more reasonable in an infinite-horizon framework where $P2$'s arrival is stochastic without the need for artificial periodization. To address this issue, in Section 4, we analyze an infinite-horizon model with a stochastic arrival time of $P2$ and derive results that are fundamentally the same.

We solve the game by backward induction. Let us consider the bargaining between the monopolistic manufacturer (M) and the second patent holder ($P2$) given any contract (F_1, r_1) between the monopolist and $P1$ in place. Since F_1 is already paid by the monopolist in the first period and sunk, the bargaining between $P2$ and the licensee depends only on r_1.

We assume that $P2$ can shut down the monopolist's business with an injunction if they cannot reach a licensing agreement.[14] By applying the logic of Lemma 1, it can be easily shown that the optimal contract between $P2$ and the licensee is to include only the lump-sum component without any royalty, that is, $F_2 = \frac{\pi(c+r_1)}{2}, r_2 = 0$.

Now consider the optimal licensing contract in the first period in the presence of an uncertain infringement claim in the second period. Suppose that a royalty rate of r_1 is chosen in the first period. Given the uncertainty surrounding the possibility of future infringement claims, and anticipating the optimal contract that would arise in such a contingency, the overall expected surplus available to be shared between $P1$ and the licensee (M) under a royalty rate r_1 is given by

$$\begin{aligned} S_1 &= \left[(1+\delta)\pi(c+r_1) - \delta\beta F_2\right] + (1+\delta)r_1 q(c+r_1) \\ &= (1+\delta)\left[\pi(c+r_1) + r_1 q(c+r_1)\right] - \delta\beta\frac{\pi(c+r_1)}{2} \end{aligned}$$

Then, F_1 is negotiated to split the surplus between $P1$ and the licensee.

$$(1+\delta)\pi(c+r_1) - \delta\beta\frac{\pi(c+r_1)}{2} - F_1 = F_1 + (1+\delta)r_1 q(c+r_1)$$

Thus, $F_1 = \frac{1}{2}[(1+\delta)\pi(c+r_1) - \delta\beta\frac{\pi(c+r_1)}{2} - (1+\delta)r_1 q(c+r_1)]$. The total licensing income for $P1$ is given by

$$L_1 = \frac{S_1}{2} = \frac{(1+\delta)}{2}[\pi(c+r_1) + r_1 q(c+r_1)] - \frac{\delta\beta}{4}\pi(c+r_1)$$

$P1$ and the licensee will choose r_1 to maximize L_1 (or equivalently S_1). By evaluating the derivative of S_1 with respect to r_1 at $r_1 = 0$, we can immediately see that the fixed fee only contract is no longer optimal. More specifically, by using the envelope theorem that $\pi'(c+r) = -q(c+r)$, we obtain

$$\frac{dS_1}{dr_1} = (1+\delta)r_1 q'(c+r_1) + \frac{\delta\beta}{2}q(c+r_1). \tag{1}$$

Evaluating (1) at $r_1 = 0$ yields

$$\left.\frac{dS_1}{dr_1}\right|_{r_1=0} = \frac{\delta\beta}{2}q(c) > 0,$$

implying that any maximum of S_1 is attained with the inclusion of a positive royalty rate r_1. To facilitate a comparative static analysis of r_1 and gain further insight into the properties of the optimal royalty rate, we assume that $q(.)$ is log-concave, which implies that $r_1 q'(c+r_1)/q(c+r_1)$ is decreasing in r_1. The log-concavity of $q(.)$ therefore implies that S_1 is quasi-concave in r_1 and that the first order condition $\frac{dS_1}{dr_1} = 0$ satisfies a single-crossing property.[15] These properties ensure that any solution satisfying the first-order condition is the unique global maximizer, and that the second-order condition holds:

$$\frac{d^2 S_1}{dr_1^2} = \underbrace{(1+\delta+\frac{\delta\beta}{2})q'(c+r_1) + (1+\delta)r_1 q''(c+r_1)}_{[S.O.C.]} < 0 \tag{2}$$

This implies that

$$r_1^* = -\frac{\delta\beta q(c+r_1^*)}{2(1+\delta)q'(c+r_1^*)} > 0, \tag{3}$$

which implicitly defines the optimal royalty rate r_1^* between $P1$ and the licensee.

To derive the comparative statics of the optimal royalty rate with respect to β, we totally differentiate the first order condition (1).

$$\frac{\delta}{4}q(c+r_1)d\beta + [S.O.C]dr_1 = 0$$

Therefore,

$$\frac{dr_1^*}{d\beta} = -\frac{\frac{\delta}{4}q(c+r_1)}{[S.O.C]} > 0$$

Similarly, we can derive

$$\begin{aligned} \frac{dr_1^*}{d\delta} &= -\frac{\frac{1}{2}r_1 q'(c+r_1) + \frac{\beta}{4}q(c+r_1)}{[S.O.C]} \\ &= \frac{\frac{1}{2\delta}r_1 q'(c+r_1)}{[S.O.C]} > 0 \end{aligned}$$

The second equality above comes from the first order condition (1).

Condition (3) that implicitly defines the optimal royalty rate also suggests that the optimal royalty rate is higher when demand is inelastic, that is, when $-\frac{q}{q'}$ is high. We verify this for the case of a constant elasticity of demand function

$$P(q) = Aq^{-\frac{1}{\eta}}.$$

We assume that the elasticity of demand, η, exceeds one. Then, it can be easily verified that the optimal royalty rate is given by

$$r_1^* = \frac{\delta\beta c}{2\eta(1+\delta) - \delta\beta},$$

which is decreasing in the price elasticity of demand η.

Proposition 1. *The optimal contract between P1 and the licensee includes a positive royalty rate. In addition, this royalty rate is increasing in β and δ, and decreasing in the price elasticity of demand.*

The intuition for our result is as follows. In our model, a royalty contract distorts the licensee's true marginal production cost, and introduces inefficiency that reduces the total size of pie available to the licensor and the licensee. Thus, without any potential future IP claims, there is no reason to introduce a royalty component in the licensing contract. However, in the presence of potential future claims, a royalty contract plays a strategic role of reducing surplus available to future IP claimants. Thus, a royalty contract serves as a *rent protection* mechanism by the initial patent holder and the licensee. The optimal royalty rate trades off rent protection from a potential future patent claimant against allocative inefficiency in the market. As β and δ increases, the rent protection becomes more important, and the optimal royalty rate is increased at the expense of allocative efficiency.

There are many papers that explain the royalty contract as an optimal contractual form in the presence of market imperfection. For instance, Gallini and Wright [11] and Beggs [12] provide an analysis of licensing in the presence of asymmetric information. They consider a situation in which the licensor has superior information about the value of the technology and explain the royalty rate in the contract as a signaling device for the party who has better information about the value of the licensed technology. Choi [13] considers a licensing environment in which effective transmission of knowledge requires costly inputs by both dispensing and receiving parties. He shows that the introduction of inputs that are not contractible and costly explains the prevalence of royalty contracts in the licensing relationship. In all these models, royalty contracts are a response to imperfect contractual environments and a second-best solution to asymmetric information and/or moral hazard problems. In our model, a royalty contract is a pure rent protection mechanism without any positive effects.

As in this paper, Bousquet et al. [35] assume symmetric information between the licensor and the licensee, and analyze the design of license contracts under demand or cost uncertainty. They show that the optimal contracts generally consist of a mix of a fixed fee and royalties. In their paper, the licensee is assumed to be risk-averse, and a royalty enables the two licensing parties share the risk as it is conditional on output. However, our paper assumes all parties are risk-neutral. The role of royalty for the two first movers is to protect future rents from the second mover ($P2$) as a Stackelberg leader. In this sense, the nature of the contract between $P1$ and the licensee is reminiscent of Aghion and Bolton [4], in which an incumbent and a buyer form a coalition through an exclusive contract that enables them to extract part of the entrant's rent. The key difference is that in our setup the licensing contract between $P1$ and the licensee is designed to protect their rents from an outsider rather than to extract rents from an entrant. In particular, the outsider ($P2$) in our model attempts to extract rents from the initial contracting parties without creating additional surplus, whereas the entrant in Aghion and Bolton [4] generates additional surplus. This also explains why, unlike in Aghion and Bolton [4], renegotiation does not eliminate the strategic value of the royalty in our setting. Even if the parties can renegotiate ex post, the initial inclusion of a royalty still serves to reduce the surplus available to $P2$ and thereby mitigates the hold-up problem.

4 | Discussion and Extensions of the Model

A natural question to ask is if there is any other commitment mechanism other than the royalty contract that serves the same purpose but without sacrificing any efficiency. In our simple model, the answer turns out to be affirmative. We have assumed that the fixed fee for licensing is paid in the first period and considered sunk in the second period. However, we can consider a scheme in which the fixed component is paid over time in installments out of each period's profits over time, (f_1, f_2), where f_i is the fixed fee to be paid by the monopolist to $P1$ in period i. With such a contract allowed, we can easily verify that a fixed fee licensing contract paid over time is superior to the one analyzed above. To see this, let (F_1, r_1) be the optimal licensing contract when the whole fixed fee needs to be paid in the first period. As an alternative licensing contract with fixed fees paid over time, consider a contract with $f_2 = r_1 q(c + r_1)$, which yields the same licensing income to $P1$ in the second period as the optimal fixed fee cum royalty contract considered above. It will yield a higher payoff to the licensee in the second period due to the absence of a royalty rate. In addition, the profit to be shared by $P1$ and the licensee in the first period is also higher. As a result, such a contract strictly dominates the one we considered above, with the assumption that the fixed fee component is paid fully in the first period. In fact, our logic tells us that the optimal fixed fee contract should be end-loaded to protect rents from future infringement claims. If a negative payment of $P1$ to the licensee is feasible in the first period, it would be possible to construct a contract that completely shields $P1$ and M from future payment claims by pledging all second-period profits as f_2 to P1, leaving no surplus for $P2$ to extract.[16]

However, the logic above crucially depends on the certainty of future market demand and on the artificial structure of the two-period model, which makes the timing of future infringement claims certain. In the next subsections, we show that introducing more realistic features, such as demand uncertainty or stochastic arrival of future claims, restores the optimality of royalty-rate contracts.

4.1 | Model With Demand Uncertainty

We imagine a situation in which demand in the second period is uncertain. More specifically, assume that demand in the second period can be either high (H) or low (L) with probabilities of α and $(1 - \alpha)$, respectively, where $P_H(q) > P_L(q)$ for all q. Let $q_i(c)$ be the profit-maximizing level of output for demand state i when the marginal cost is given by c, that is, $q_i(c) = argmax[P_i(q)q - cq]$, where $i = L, H$. Let $\pi_i(c)$ be the corresponding (reduced) profit function for the licensee. With the possibility of the fixed fee component being paid in installments, we consider the contractual form of (f_1, f_2, r). We assume that the future fixed payment f_2 cannot be contingent on the demand state because it cannot be verified in court. When the second patent holder approaches the monopolist in the second period, the realized market demand is known to both. Note that we implicitly assume that the per-unit royalty rate is constant over time. Once again, we justify this assumption on the idea of stochastic arrival time of $P2$, as the two-period model imposes an artificial entry time of $P2$ (see the next subsection for stochastic arrival time).

As usual, we solve the game by backward induction. Let us consider the bargaining between the monopolist and the second patent holder ($P2$) given any contract (f_1, f_2, r). For demand uncertainty to be relevant to our analysis, we assume that it is optimal for M to operate even in the low-demand state.[17] Note that the optimality of M's operation under the low-demand state requires that $f_2 \leq \pi_L(c+r)$.[18] By the same logic as before, the optimal contract between $P2$ and the licensee is to include only the lump-sum component without any royalty, that is, $F_{2i} = \frac{\pi_i(c+r)-f_2}{2}, r_2 = 0$.

Now consider the optimal licensing contract in the first period between the monopolist and $P1$. Given a royalty rate of r, (f_1, f_2) will be chosen to satisfy the following condition in Nash bargaining:

$$\pi(c+r) + \delta E[\pi_i(c+r)] - \delta\beta\left[\frac{E[\pi_i(c+r)] - f_2}{2}\right] - (f_1 + \delta f_2)$$
$$= (f_1 + \delta f_2) + rq(c+r) + \delta r E[q_i(c+r)],$$

where $E[\pi_i(c+r)] = \alpha\pi_H(c+r) + (1-\alpha)\pi_L(c+r)$ and $E[q_i(c+r)] = \alpha q_H(c+r) + (1-\alpha)q_L(c+r)$.

The logic for end-loading incentives in the contract implies that the constraint for f_2 is binding with the optimal $f_2 = \pi_L(c+r)$ for any choice of r.

Lemma 2. *Suppose that the fixed fee can be paid in installments over two periods. The optimal contract between P1 and the licensee requires that $f_2 = \pi_L(c+r)$ for any choice of r.*

Proof. Suppose not and the optimal contract is given by (f_1, f_2), where $f_2 < \pi_L(c+r)$. Then we can devise another contract $(\hat{f}_1, \hat{f}_2)$ such that $\hat{f}_1 = f_1 - (1 - \frac{\beta}{4})\delta\epsilon$ and $\hat{f}_2 = f_2 + \epsilon$, where ε is an arbitrarily small positive number. This contract equally splits the surplus between P1 and the licensee, and improves both parties' payoffs by $\frac{\beta}{4}\delta\epsilon$. We thus have a contradiction. □

Once we pin down $f_2 = \pi_L(c+r)$, the total licensing income for $P1$ can be written as

$$L_1 = \frac{1}{2}\left[\pi(c+r) + \delta\left(1 - \frac{\beta}{2}\right)E[\pi_i(c+r)] + rq(c+r) + \delta r E[q_i(c+r)] + \frac{\delta\beta}{2}\pi_L(c+r)\right]$$

$P1$ and the licensee will choose r to maximize L_1. By using the envelope theorem that $\pi_i'(c+r) = -q_i(c+r)$, where $i = L, H$, the first order condition is given by

$$\frac{dL_1}{dr} = \frac{\delta\beta}{2}E[q_i(c+r)] + r[q'(c+r) + \delta[\alpha q_H'(c+r) + (1-\alpha)q_L'(c+r)]] - \frac{\delta\beta}{2}q_L(c+r) = 0 \quad (4)$$

We thus have

$$r^* = -\frac{\delta\beta\alpha[q_H(c+r) - q_L(c+r)]}{2[q'(c+r) + \delta[\alpha q_H'(c+r) + (1-\alpha)q_L'(c+r)]]} > 0, \quad (5)$$

which implicitly defines the optimal royalty rate r^* between $P1$ and the licensee.

Proposition 2. *With demand uncertainty, the optimal contract between P1 and the licensee includes a positive royalty rate, even if the fixed fee can be paid in installments.*

Condition (3) that implicitly defines the optimal royalty rate also suggests that the optimal royalty rate is higher when demand uncertainty is greater. To be more specific, consider a demand function with constant elasticity.

$$P_i(q) = A_i q^{-\frac{1}{\eta}},$$

where $i = H, L$, $A_H > A_L (> 0)$, and $\eta > 1$. To analyze the effects of greater demand variation on the optimal royalty rate, we consider a mean-preserving spread in $A_i(\Delta_A = A_H - A_L)$ such that $\alpha A_H + (1-\alpha)A_L = A$. With a constant elasticity demand function, it can be shown that the optimal royalty rate with delayed fixed fee payments is given by

$$r_1^* = \frac{\xi c}{\eta - \xi}, \text{ where } \xi = \frac{\delta}{1+\delta}\beta\alpha\Delta_A$$

The optimal royalty rate is increasing in δ, β, and $\Delta_A = A_H - A_L$, and decreasing in η.

A positive royalty rate naturally arises under demand uncertainty if the licensee is risk averse. In such a case, a royalty rate can be used as a risk-sharing mechanism because the licensee's royalty payment will co-vary with the demand realizations, thus shielding the licensee from negative demand shocks. Note that here we assume risk-neutrality for both the licensor and the licensee. Thus, our explanation for a positive royalty rate is not driven by risk sharing.

4.2 | A Stochastic Arrival Model

As we discussed earlier, we have assumed that the royalty rate is constant over time. This would be a restrictive assumption in the two-period model we considered. However, this would be a more realistic assumption when the arrival time of $P2$ is not restricted to the second period and the royalty rates cannot be fine-tuned to $P2$'s arrival time. To address this issue, we now develop an infinite horizon continuous time model in which $P2$ appears at a random time from the perspectives of $P1$ and the licensee at the time of the initial licensing contracting. We represent this uncertainty as a random Poisson arrival time with a parameter of λ. Once again, the optimal contract between $P2$ and the licensee will be characterized by $F_2 = \frac{\pi(c+r_1)}{2\rho}, r_2 = 0$, where ρ is the discount rate. The following proposition for a stochastic arrival time qualitatively replicates the optimal contract in a two-period model.

Proposition 3. *The optimal contract between P1 and the licensee includes a positive royalty rate. In addition, this royalty rate is increasing in λ and decreasing in ρ.*

Proof. Suppose that a royalty rate of r_1 is chosen in the initial contract between $P1$ and the licensee. Given a royalty rate of r_1 and a random arrival time of $P2$, the expected profit for the licensee gross of a lump-sum payment to $P1$ is given by:

$$V(r_1,\lambda) = \int_0^\infty e^{-\rho t}e^{-\lambda t}\left[\pi(c+r_1) + \lambda\frac{\pi(c+r_1)}{2\rho}\right]dt$$
$$= \frac{\pi(c+r_1) + \lambda\frac{\pi(c+r_1)}{2\rho}}{\rho+\lambda}$$

Then, F_1 is negotiated to split the expected surplus between $P1$ and the licensee.

$$V(r_1,\lambda) - F_1 = F_1 + \frac{r_1 q(c+r_1)}{\rho}$$

Thus, $F_1 = \frac{1}{2}[V(r_1,\lambda) - \frac{r_1 q(c+r_1)}{\rho}]$. The total licensing income for $P1$ is given by

$$L_1 = \frac{1}{2}[V(r_1,\lambda)] + \frac{r_1 q(c+r_1)}{2\rho}$$

$P1$ and the licensee will choose r_1 to maximize L_1. Once again, by using the envelope theorem that $\pi'(c+r) = -q(c+r)$, the first order condition is given by

$$\frac{2dL_1}{dr_1} = \frac{\lambda q(c+r_1)}{2\rho(\rho+\lambda)} + \frac{r_1 q'(c+r_1)}{\rho} = 0 \tag{6}$$

This implies that

$$r_1^* = -\frac{\lambda q(c+r_1^*)}{2(\rho+\lambda)q'(c+r_1^*)} > 0$$

By totally differentiating (6)with respect to ρ and λ, we can easily verify that

$$\frac{dr_1^*}{d\lambda} > 0 \text{ and } \frac{dr_1^*}{d\rho} < 0$$

□

Remark 1. We have assumed that the manufacturing licensee (M) and $P2$ negotiate with the existing contract between M and $P1$ as given. We can entertain the possibility of renegotiation to have a three-party lump-sum sharing of profits in the event until $P2$'s arrival, similar to what is optimal in the case of two parties. However, the qualitative results do not change, and the incentive to include a royalty rate in the initial contract between M and $P1$ still remains. Furthermore, we have assumed that only one additional patentee lurks. If multiple potential patent claimants exist, then even the contract between M and $P2$ would entail a royalty component.

4.3 | Endogenous Incentives to "Lurk"

In the baseline model of Section 3, we assumed an exogenous probability for $P2$'s emergence to assert its patent rights. However, it can be easily verified that $P1$'s licensing income is higher than $P2$'s, which raises the question of why $P2$ would wait until the second period to approach the manufacturing firm. One plausible scenario is that $P1$'s patent is publicly well-known, allowing the manufacturing firm to negotiate an ex ante licensing contract with $P1$, while $P2$'s patent remains unknown to the firm. Only after the product is introduced does $P2$ discover the manufacturing firm's infringement of its patent and subsequently approach the firm to negotiate a licensing contract.

In this subsection, we consider an alternative scenario in which $P2$ is aware of the manufacturing firm's product but strategically waits until after its introduction before asserting its patent rights. This ex post hold-up incentive against the manufacturing firm can be derived by assuming a sunk cost of product development. More specifically, suppose that developing the product incurs a sunk cost K. In this case, an ex ante contract before the development of the product would divide the manufacturing firm's total profits among three parties: the manufacturing firm, $P1$, and $P2$. Using the same logic as in Section 3, under these conditions the optimal contract would consist solely of a fixed fee, with no per-unit royalty rate, which results in the total monopoly profit of $(1+\delta)\pi(c) - K$. This arrangement yields an equal split of the total profit, with each party receiving one-third of the monopoly profit (i.e., $\frac{(1+\delta)\pi(c)-K}{3}$), assuming equal bargaining power. This outcome is consistent with the Shapley value solution.

Now suppose that $P2$ keeps its patent in the dark and approaches the manufacturing firm (M) only after the introduction of the product in period 2. Anticipating this possibility, M and $P1$ negotiate as in the baseline model. As the development cost is already sunk, the optimal contract between $P2$ and the licensee (M) is, once again, to include only the lump-sum component without any royalty, that is, $F_2 = \frac{\pi(c+r_1)}{2}, r_2 = 0$. As in the analysis of the baseline model, suppose that a royalty rate of r_1 is chosen in the first period. Then, F_1 is negotiated to split the surplus between $P1$ and the licensee. Accounting for the development cost of K, we have

$$(1+\delta)\pi(c+r_1) - \delta\beta\frac{\pi(c+r_1)}{2} - F_1 - K = F_1 + (1+\delta)r_1 q(c+r_1),$$

which yields $F_1 = \frac{1}{2}[(1+\delta)\pi(c+r_1) - \delta\beta\frac{\pi(c+r_1)}{2} - (1+\delta)rq(c+r_1) - K]$. The total licensing income for $P1$ is given by

$$L_1 = \frac{(1+\delta)}{2}[\pi(c+r_1) + r_1 q(c+r_1)] - \frac{K}{2} - \frac{\delta\beta}{4}\pi(c+r_1)$$

We can easily see that the optimal royalty rate does not change in comparison to the baseline model because the development cost is independent of the output: it is borne by the two parties equally and the only effect is to reduce each part's payoff by $\frac{K}{2}$. Thus, the optimal royalty rate is implicitly determined by the same condition in (3), which is independent of K. Thus, $P2$'s payoff from "lurking" is given by $\frac{\pi(c+r_1^*)}{2}$. Thus, "lurking" is more profitable for $P2$ if

$$F_2 = \delta\frac{\pi(c+r_1^*)}{2} > \frac{(1+\delta)\pi(c) - K}{3},$$

where r_1^* is implicitly defined by (3), that is $r_1^* = -\frac{\delta\beta q(c+r_1^*)}{2(1+\delta)q'(c+r_1^*)} > 0$. This condition can be expressed as

$$K > (1+\delta)\pi(c) - \frac{3\delta}{2}\pi(c+r_1^*), \tag{7}$$

indicating that when K is sufficiently large, $P2$ has an incentive to keep its patent undisclosed and makes a surprise licensing demand, thereby holding up the practicing entity. In addition, however, there is also an upper bound on K to ensure that M

has an incentive to develop the product in the first place, which is given by

$$K < (1+\delta)[\pi(c+r_1^*) + r_1 q(c+r_1^*)] - \frac{\delta\beta}{2}\pi(c+r_1^*). \qquad (8)$$

Taken together, we obtain an equilibrium in which incentives for $P2$ to lurk endogenously arise if the following condition holds:

$$(1+\delta)\pi(c) - \frac{3\delta}{2}\pi(c+r_1^*) < K < (1+\delta) \times [\pi(c+r_1^*) + r_1 q(c+r_1^*)] - \frac{\delta\beta}{2}\pi(c+r_1^*). \qquad (9)$$

Note that the condition above is written in terms of the endogenous variable r_1^*. Nevertheless, we can see that there exists a range of K for which $P2$ prefers to lurk while M still prefers to contract with $P1$ when β is relatively small. To see this, observe that as $\beta \to 0$, $r_1^* \to 0$, and $[\pi(c+r_1^*) + r_1 q(c+r_1^*)]$ approaches $\pi(c)$, which ensures the existence of range for K satisfying condition (9).

To illustrate the point above, consider again the case of a constant elasticity of demand function, $P(q) = Aq^{-\frac{1}{\eta}}$, with elasticity η (>1), and the associated optimal royalty rate

$$r_1^* = \frac{\delta\beta c}{2\eta(1+\delta) - \delta\beta}.$$

We can find a range of K that satisfies condition (9) whenever $(1+\delta)[\pi(c+r_1^*) + r_1 q(c+r_1^*)] - \frac{\delta\beta}{2}\pi(c+r_1^*) > (1+\delta)\pi(c) - \frac{3\delta}{2}\pi(c+r_1^*)$. Under constant elasticity demand, we have

$$\pi(x) = Kx^{-(\eta-1)} \text{ and } q(x) = (\eta-1)Kx^{-\eta}.$$

Therefore, the above condition can be written as

$$\left(1 + \frac{5}{2}\delta - \frac{\delta\beta}{2\eta}\right)\left(1 - \frac{\delta\beta}{2\eta(1+\delta)}\right)^{\eta-1} > 1+\delta$$

The left-hand side of the above inequality is satisfied at $\beta = 0$ and is strictly decreasing in β. This implies that either there exists a unique cutoff $\beta^* \in (0,1)$ such that the condition is satisfied for all $\beta < \beta^*$, or the condition holds for the entire range $\beta \in [0,1]$.

4.4 | Substitute Patents

So far, we have analyzed the possibility of future infringement claims involving complementary technologies, that is, a setting in which M requires both patents held by $P1$ and $P2$. We now turn to a different scenario in which future infringement claims arise from a substitute technology, meaning that M needs only one of the two technologies or that the incremental value of the second technology is lower than its stand-alone value. We show that the nature of the relationship between the two patents matters and that complementarity is essential for our main result. If the two technologies are substitutes, the rationale for a royalty rate as a rent-protection mechanism disappears, and the optimal royalty rate is zero.

To formalize this idea, we modify the model to incorporate substitute patents. Let $\pi(c)$ and $q(c)$ denote the (reduced-form) profit function of the licensee and the corresponding profit-maximizing level of output when its constant marginal cost is c. M's marginal cost is c_1 when $P1$'s technology is used. In the second period, another patent holder ($P2$) appears with probability β. In contrast to the baseline model, the two patents are substitutes, and $P2$'s technology allows M to produce the same product at cost c_2, where c_2 is distributed according to $G(.)$ on $[0,\infty)$. If the initial contract between M and $P1$ includes a royalty rate r_1, then entry by $P2$ occurs only if $c_2 \le c_1 + r_1$.

This implies that the inclusion of a positive royalty rate increases the probability of entry by $P2$. When such an entry occurs, $P1$ loses its licensing income in the second period unless the initial contract between M and $P1$ includes an exclusive licensing provision, as in Aghion and Bolton [4]. It therefore follows immediately that any contract with a positive royalty rate is strictly dominated by one without a royalty rate: the royalty sacrifices allocative efficiency in the first period while simultaneously increasing the likelihood of entry by a substitute patent holder.

Proposition 4. *When the patents held by P1 and P2 are substitutes, the optimal contract between P1 and the licensee (M) includes no positive royalty rate.*

Proof. See the Appendix. □

4.5 | Asymmetric Bargaining Power Between the Manufacturer and Patent Holders

We have so far assumed equal bargaining power between the patent holders and the manufacturer. In this subsection, we show that our main results are robust to the introduction of asymmetric bargaining power. To this end, consider the more general case in which the relative bargaining power between the patent holder Pi (for $i = 1,2$) and the manufacturer is given by $\mu_i \in (0,1)$ and $(1-\mu_i)$, respectively, representing their respective shares of the surplus generated relative to the disagreement point.

Our main qualitative results are robust to these modifications. In the second period, if $P2$ appears and there is no further risk of subsequent infringement claims, the optimal contract consists solely of a fixed fee, with the surplus divided between M and $P2$ according to their relative bargaining power. In the first period, however, anticipating the risk of future patent claims in the second period, the optimal contract includes a positive royalty rate. This royalty rate is increasing in $P2$'s bargaining power μ_2, since its primary function is to protect rents from $P2$. Both mathematically and intuitively, this effect is isomorphic to an increase in β in Proposition 1, as shown in the Appendix. By contrast, $P1$'s bargaining power μ_1 affects only the determination of the fixed fee F_1. Hence, allocative efficiency is influenced by μ_2, while μ_1 has purely distributional consequences.

Proposition 5. *With asymmetric bargaining power between the manufacturer and the patent holders, the optimal contract between P1 and the licensee includes a positive royalty rate. Moreover, the optimal royalty rate is increasing in P2's bargaining power,* μ_2.

Proof. See the Appendix. □

5 | Concluding Remarks With Policy Implications

We have analyzed the optimal licensing contract in an environment under which the manufacturing licensee can be subject to uncertain future infringement claims by unknown patent holders. We propose structured licensing contracts with royalty components as a mechanism to protect firm profits from the holdup problem associated with lurking patents. Ideally, licensing agreements should anticipate the possibility of future lurking patents, incorporating contingency clauses or royalty arrangements that account for the cost of potential additional licenses if lurking patents emerge. However, these contracts can be challenging to implement due to ex post disputes over contingency interpretation when issues arise. While including a royalty component may introduce some inefficiency, it offers a more robust and easier-to-implement solution. The licensing literature has identified several reasons why royalty contracts can be optimal, including benefits like risk-sharing and addressing market imperfections due to information asymmetries between licensors and licensees, which may lead to adverse selection [11] or moral hazard [13]. This paper highlights an additional channel through which royalty contracts can protect licensee profits from holdup risks.

The environment we analyze in this paper has become increasingly relevant in recent years, driven by the surge in patent applications and grants, as well as the rise of patent assertion entities (PAEs) as a major force in patent litigation. PAEs often pejoratively referred to as "patent trolls" are organizations whose primary business model is to acquire and enforce patents in order to extract licensing revenues, frequently relying on strategies of secrecy and surprise infringement claims against practicing entities. Although PAEs are not a new phenomenon, their prominence has grown substantially.[19] Indeed, one study reports that patent assertion entities were responsible for approximately 67 percent of all patent lawsuits [36]. In such an environment, we show that a licensing contract with a royalty component can be used as a mechanism to shield rents from future extraction of rents by a third party. Our model provides a new rationale for the use of royalty rates in licensing contract, despite the apparent inefficiency associated with such a contract.

Our paper also has implications for policymakers. The holdup problem related to submarine or lurking patents has long been recognized, prompting past reforms to patent laws. In 1995, the United States changed patent terms to 20 years from the filing date (rather than from the grant date). However, the publication of most patent applications did not begin until 2000, when the American Inventors Protection Act of 1999 mandated the publication of most applications 18 months after filing. These changes were implemented as part of the TRIPS Agreement under the WTO. Nevertheless, applicants may still keep their filings secret if they do not seek patent protection outside the United States. Closing this loophole would further mitigate the risk of patent hold-up. Additionally, antitrust authorities should closely monitor the use of "continuation patents" to prevent abusive practices (see [23]).

Our analysis also considers environments with potentially multiple patent holders a particularly relevant issue in the high-tech sector, which involves many complementary technologies. In an increasingly interconnected world, interoperability among devices has become essential, making standards crucial to ensure compatibility among products from different vendors. One solution is for standard-setting organizations (SSOs) to mandate fair, reasonable, and nondiscriminatory (FRAND) licensing for standard essential patents (SEPs).[20] Nonetheless, lurking patents remain a potential issue, as seen in the case of Rambus v. F.T.C, 522 F.3d 456 (D.C. Cir. 2008). Rambus was accused of not disclosing relevant patents or applications during DRAM standard development. Once the standards gained widespread adoption, Rambus pursued infringement claims, engaging in strategic holdup.

Policymakers and courts are encouraged to mandate stricter disclosure requirements for SSOs, requiring participants to identify any relevant patents or pending applications when engaging in standards development.[21] Policies could further limit injunctive relief and require compulsory licensing at reasonable rates for SEP holders who fail to disclose relevant patents during the standard-setting process, restricting their ability to demand excessive royalties. Antitrust authorities might also issue guidelines on anticompetitive practices related to SEPs, emphasizing that intentional nondisclosure and holdup can constitute antitrust violations.

Acknowledgements

I thank Joseph Farrell and Carl Shapiro for useful comments and discussions. I am also grateful to two anonymous referees and Editor Andrew Rhodes for their constructive comments and valuable guidance, which significantly improved the paper. All remaining errors are mine. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (Grant No. NRF-2026S1A5A2A01004892).

Funding

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (Grant No. NRF-2026S1A5A2A01004892).

Endnotes

[1] See Hegde and Luo [2] for an empirical analysis of AIPA's impact on the timing of licensing deals in the biomedical industry. Their findings indicate that, on average, AIPA's enactment shortened licensing delays by approximately 10 months. This suggests that AIPA facilitates transactions in the market for ideas by lowering information costs.

[2] If the licensor is also a practicing entity competing with the licensee in the same market, a royalty contract will be optimal for mitigating competitive pressures. To highlight a new rationale for royalty contracts in this paper, we intentionally consider a scenario in which no other factors justify the use of royalty contracts.

[3] In Section 4, we show that when the patents are substitutes rather than complements, the optimal initial contract does not include a royalty component.

[4] See Kamien [5] for an excellent survey on patent licensing.

[5] See also Sen [8] and Sen and Tauman [9] for more recent analyses in the context of cost-reducing innovation in oligopoly.

[6] See Shapiro [10] for an analysis of royalty rate *cross-licensing* contracts as a collusive device for oligopolistic firms.

[7]See also Bhattacharyya and Lafontaine [14] for a model of double-sided moral hazard that explains revenue/profit-sharing arrangements in the framework of business-format franchising.

[8]See Hegde [15] for empirical evidence showing that the structure of license contracts in the biomedical industry aligns broadly with theoretical predictions. Kotha et al. [16] also demonstrate how an effective contract payment structure a combination of upfront and royalty payments can help mitigate asymmetric information problems in the absence of proven results. By analyzing university technology licensing contracts, they show that structuring payments to trade lower upfront fees for higher royalty payments serves as a signaling mechanism for value.

[9]See also Farrell and Shapiro [19] and Jeon and Lefouili [20] for discussions on how per-unit royalties can serve to relax competition among licensees.

[10]Farrell and Shapiro [19] also rely on coordination externalities requiring multiple licensees to show how royalties can be excessive under uncertainty about a patent's validity.

[11]Jerome Lemelson, an American inventor, was known for filing numerous patents and intentionally delaying their approval through continuation applications. Many of his patents covered emerging technologies, such as bar code scanners and machine vision. Once these technologies became essential, Lemelson enforced his patents, resulting in costly settlements from companies already using the technology. Continuation patents are also widely used in the pharmaceutical industry to increase the parent patent's expected patent scope with additional uncertainty. See, for instance, European Commission [23] for more details.

[12]Another contributing factor is that there is no legal obligation for a firm to investigate whether other companies hold patents. However, once a firm becomes aware of a patent that it might infringe, legal duties arise, and if infringement is found, it could be seen as "willful infringement" and the firm may be liable for treble damages. This situation creates perverse incentives for practicing entities, with some lawyers even suggesting that it might be better not to search for other firms' patents [25].

[13]Other prominent cases of "lurking patents" include Rambus Inc. and the DRAM Standards, TiVo vs. DISH Network, NTP vs. Research In Motion (RIM), and Honeywell vs. Minolta (Auto-Focus Camera), among others.

[14]For an analysis of injunctions in patent enforcement, see Shapiro [32]. We can easily accommodate the possibility that $P2$'s patent right is probabilistic and its validity can be challenged in the court. In such a case, our parameter β can also capture the strength of $P2$'s patent. See Choi and Gerlach [33, 34] for such an analysis.

[15]Note that the log-concavity assumption on $q(.)$ is not primitive, since $q(c)$ is defined as the profit-maximizing quantity given marginal cost c. A sufficient condition, expressed in terms of primitives of the model, is that the marginal revenue function $MR(q)$ is strictly decreasing and satisfies $MR'(q) + qMR''(q) \leq 0$ for all relevant q. It is satisfied by many commonly used demand specifications, including linear and constant-elasticity demand.

[16]If the two periods are of equal length and $\delta(< 1)$ represents the discount factor, such a contract is feasible even without a reverse payment from $P1$ to the licensee in the first period.

[17]This condition is satisfied provided that neither the probability of the low-demand state nor demand in that state is negligible that is, when α is not too close to 1 and $P_L(q)$ is not near zero. For instance, suppose α is sufficiently high, with $\alpha = 1 - \epsilon$, where ϵ is arbitrarily small. In that case, the optimal contract may prescribe exit by M in the low-demand state, which arises only with probability ϵ. Demand uncertainty then becomes effectively irrelevant for the design of the optimal contract, resulting in a fixed-fee-only contract (with installment payments).

[18]If this condition is not satisfied, the licensee will simply exit the market in the low-demand state. This condition can also be justified by appealing to the licensee's liquidity constraints, which may prevent continued operation under unfavorable demand realizations.

[19]For example, George Selden, a late-19th-century American patent lawyer and inventor, is often regarded as one of the earliest "patent trolls." He delayed the issuance of his 1895 "road engine" automobile patent for approximately 16 years, allegedly to maximize royalty extraction once the automobile industry began to expand. The term "patent troll" itself gained prominence around 2001 after Intel was sued for defamation for referring to a patent litigant as an "extortionist." The expression is widely attributed to Peter Detkin, then Assistant General Counsel at Intel. I thank an anonymous referee for bringing this historical context to my attention.

[20]See, however, Ganglmair et al. [37] for an analysis of the limits and potential unintended consequences of RAND commitments as a solution to the hold-up problem when innovators' incentives are taken into account. In contrast to our setup with a potentially unknown patent claimant, their model considers a setting in which the manufacturer is fully aware of the innovator's patent. As a result, the manufacturer would never find it optimal to use the patented technology without first obtaining a license, since infringement would be willful and expose the firm to treble damages.

[21]Disclosure requirements also have important limitations [38], particularly because they may create incentives for over-declaration. For instance, studies by Goodman and Myers [39] and van Audenrode et al. [40] suggest that only about 20 to 40 percent of the patents disclosed to ETSI (European Telecommunications Standards Institute) are in fact truly essential.

References

1. G. Quinn, *Submarine Patents Alive and Well: Tivo Patents DVR Scheduling* (IPWatchdog, 2010), http://www.ipwatchdog.com/2010/02/19/submarine-patents-alive-and-well-tivo-patents-dvr-scheduling/.

2. D. Hegde and H. Luo, "Patent Publication and the Market for Ideas," *Management Science* 64, no. 2 (2017): 652–672.

3. E. Hovenkamp, J. Lemus, and J. L. Turner, "Product Development With Lurking Patentees," *RAND Journal of Economics* 56 (2025): 163–183.

4. P. Aghion and P. Bolton, "Contracts as a Barrier to Entry," *American Economic Review* 77 (1987): 388–401.

5. M. I. Kamien, "Patent Licensing," in *Handbook of Game Theory*, vol. 1, ed. R. J. Aumann and S. Hart (Elsevier, 1992), 331–354.

6. M. I. Kamien and Y. Tauman, "Fees Versus Royalties and the Private Value of a Patent," *Quarterly Journal of Economics* 101 (1986): 471–491.

7. M. Katz and C. Shapiro, "On the Licensing of Innovations," *RAND Journal of Economics* 16 (1985): 504–520.

8. D. Sen, "Fee Versus Royalty Reconsidered," *Games and Economic Behavior* 53 (2005): 141–147.

9. D. Sen and Y. Tauman, "General Licensing Scheme for a Cost-Reducing Innovation," *Games and Economic Behavior* 59 (2007): 163–186.

10. C. Shapiro, "Patent Licensing and R&D Rivalry," *American Economic Review* 75 (1985): 25–30.

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

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

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

14. S. Bhattacharyya and F. Lafontaine, "Double-Sided Moral Hazard and the Nature of Share Contracts," *RAND Journal of Economics* 26 (1995): 761–781.

15. D. Hegde, “Tacit Knowledge and the Structure of License Contracts: Evidence From the Biomedical Industry,” *Journal of Economics and Management Strategy* 23 (2014): 568–600.

16. R. Kotha, P. Crama, and P. H. Kim, “Experience and Signaling Value in Technology Licensing Contract Payment Structures,” *Academy of Management Journal* 61, no. 4 (2018): 1307–1342.

17. P. Rey and J. Tirole, “The Logic of Vertical Restraints,” *American Economic Review* 76 (1986): 921–939.

18. R. Amir, D. Encaoua, and Y. Lefouili, “Optimal Licensing of Uncertain Patents in the Shadow of Litigation,” *Games and Economic Behavior* 88 (2014): 320–338.

19. J. Farrell and C. Shapiro, “How Strong Are Weak Patents?,” *American Economic Review* 98 (2008): 1347–1369.

20. D.-S. Jeon and Y. Lefouili, “Cross-Licensing and Competition,” *RAND Journal of Economics* 49 (2018): 656–671.

21. M. A. Lemley and C. Shapiro, “Patent Holdup and Royalty Stacking,” *Texas Law Review* 85 (2007): 1991–2049.

22. J. Lerner and J. Tirole, “Efficient Patent Pools,” *American Economic Review* 94 (2004): 691–711.

23. European Commission, “Pharmaceutical Sector Inquiry Final Report, DG Competition,” (2009), http://ec.europa.eu/competition/sectors/pharmaceuticals/inquiry/staff_working_paper_part1.pdf.

24. B. H. Hall, “Exploring the Patent Explosion,” *Journal of Technology Transfer* 30 (2005): 35–48.

25. J. R. Harris, Willful Patent Infringement and Treble Damages: The Reason for Legal Opinions (2004), https://www.mmmlaw.com/files/documents/publications/article_234.pdf.

26. C. A. Cotripia and M. A. Lemley, “Copying in Patent Law,” *North Carolina Law Review* 87 (2009): 1421–1466.

27. J. Bessen and M. J. Meurer, “The Direct Costs From NPE Disputes,” *Cornell Law Review* 99 (2014): 387–424.

28. B. H. Hall and R. H. Ziedonis, “The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995,” *RAND Journal of Economics* 32 (2001): 101–128.

29. C. Mulligan and T. B. Lee, “Scaling the Patent System,” *NYU Annual Survey of American Law* 68 (2012): 289–317.

30. C. V. Chien, “From Arms Race to Marketplace: The Complex Patent Ecosystem and Its Implications for the Patent System,” *Hastings Law Journal* 62 (2010): 297–355.

31. T. Ewing and R. Feldman, “The Giants Among Us,” *Stanford Technology Law Review* 1 (2011).

32. C. Shapiro, “Injunctions, Hold-Up, and Patent Royalties,” *American Law and Economics Review* 12 (2010): 1–39.

33. J. P. Choi and H. Gerlach, “A Theory of Patent Portfolios,” *American Economic Journal: Microeconomics* 9 (2017): 315–351.

34. J. P. Choi and H. Gerlach, “A Model of Patent Trolls,” *International Economic Review* 59 (2018): 2075–2106.

35. A. Bousquet, H. Cremer, M. Ivaldi, and M. Wolkowicz, “Risk Sharing in Licensing,” *International Journal of Industrial Organization* 16 (1998): 535–554.

36. F. S. Morton and C. Shapiro, “Strategic Patent Acquisitions,” *Antitrust Law Journal* 79 (2014): 463–499.

37. B. Ganglmair, L. M. Froeb, and G. J. Werden, “Patent Hold-Up and Antitrust: How A Well-Intentioned Rule Could Retard Innovation,” *Journal of Industrial Economics* 60 (2012): 249–273.

38. R. Bekkers, C. Catalini, A. M. Martinelli, C. Righi, and T. Simcoe, “Disclosure Rules and Declared Essential Patents,” *Research Policy* 52 (2023): 104628.

39. D. Goodman and R. Myers, “3G Cellular Standards and Patents,” in *2005 International Conference on Wireless Networks, Communications and Mobile Computing* (IEEE, 2005), 415–420.

40. M. Van Audenrode, J. Royer, R. Stitzing, and P. Saaskilahti, Over-Declaration of Standard Essential Patents and Determinants of Essentiality. Working Paper (2017).

Appendix A

Proof of Proposition 4. In line with the baseline model, we assume that when M and $P2$ negotiate another bilateral licensing contract given the original contract between M and $P1$ as *given* if $P2$ enters. However, the result does not change even if we assume that $P1$ and $P2$ compete in Bertrand fashion and M chooses the best contract. If $P2$ does not exist, then the market outcome is the same as in the first period.

Suppose that the initial licensing contract signed between M and $P2$ in the first period contains a royalty rate of r_1. Given a royalty rate of r_1, $P2$ can induce the manufacturer to switch to its technology if $c_2 < c_1 + r_1$. By applying the logic of Lemma 1, it can be easily shown that the optimal contract between $P2$ and the licensee is to include only the lump-sum component without any royalty, that is,

$$F_2 = \frac{\pi(c_2) - \pi(c_1 + r_1)}{2}, r_2 = 0.$$

Now consider the optimal licensing contract in the first period between M and $P1$ in the presence of potential substitute patent in the second period. Suppose that a royalty rate of r_1 is chosen in the first period. Given the uncertainty surrounding the possibility of future infringement claims, and anticipating the optimal contract that would arise in such a contingency, the overall expected surplus available to be shared between $P1$ and the licensee under a royalty rate r_1 is given by

$$\begin{aligned} S_1 = \Big[&(1+\delta)\big[\pi(c_1 + r_1) + r_1 q(c_1 + r_1)\big] \\ &+\delta \int_0^{c_1+r_1} \frac{\pi(c_2) - \big[\pi(c_1 + r_1) + r_1 q(c_1 + r_1)\big]}{2} dG(c_2)\Big] \end{aligned}$$

By using the Leibniz integral rule, we can verify that

$$\frac{dS_1}{dr_1} = \left[1 + \delta\left(1 - \frac{G(c + r_1)}{2}\right)\right] r_1 q'(c + r_1) - \delta r_1 q(c + r_1),$$

which is negative for all $r_1 > 0$. Thus, the optimal royalty rate is zero and the initial contract between M and $P1$ only includes the fixed fee, which is given by

$$F_1 = \frac{S_1}{2}\bigg|_{r_1=0} = \frac{(1+\delta)}{2}\pi(c_1) + \frac{\delta}{2}\int_0^{c_1} \frac{\pi(c_2) - \pi(c_1)}{2} dG(c_2).$$

□

Proof of Proposition 5. We show that our analysis is robust to general bargaining power and the strategic royalty rate as a rent protection is increasing in μ_2. With the same logic in the case of equal bargaining power, the optimal contract between $P2$ and the licensee is to include only the lump-sum component without any royalty, that is, $F_2 = \mu_2 \pi(c + r_1)$, $r_2 = 0$, where μ_2 represents $P2$'s bargaining power.

Suppose that a royalty rate of r_1 is chosen in the first period. Anticipating the risk of rent extraction by potential future infringement claims F_2, the overall expected surplus available to be shared between $P1$ and the licensee under a royalty rate r_1 is given by

$$\begin{aligned} S_1 &= \big[(1+\delta)\pi(c + r_1) - \delta\beta F_2\big] + (1+\delta) r_1 q(c + r_1) \\ &= (1+\delta)\big[\pi(c + r_1) + r_1 q(c + r_1)\big] - \delta\beta\big[\mu_2 \pi(c + r_1)\big] \end{aligned}$$

Then, $P1$ and the licensee will choose r_1 to maximize S_1 while F_1 is negotiated to split the maximized surplus between $P1$ and the licensee according to their relative bargaining power. By proceeding as in the case of equal bargaining power in the main text, the first order condition can be written as

$$\frac{dS_1}{dr_1} = (1+\delta)r_1 q'(c+r_1) + \delta\beta\mu_2 q(c+r_1) = 0 \tag{A1}$$

Thus, we have

$$\left.\frac{dS_1}{dr_1}\right|_{r_1=0} = \delta\beta\mu_2 q(c) > 0,$$

implying that any maximum of S_1 is attained with the inclusion of a positive royalty rate r_1. The optimal royalty rate is implicitly defined by

$$r_1^* = -\frac{\delta\beta\mu_2 q(c+r_1^*)}{(1+\delta)q'(c+r_1^*)} > 0, \tag{A2}$$

and we have

$$\frac{dr_1^*}{d\mu_2} = -\frac{\delta\beta q(c+r_1)}{[S.O.C]} > 0.$$

□