Investment and Patent Licensing in the Value Chain*

Gerard Llobet
CEMFI and CEPR

Damien Neven
The Graduate Institute Geneva and CEPR

September 14, 2023

Abstract

At which stage in the production chain should patent licensing takes place? In this paper we show that under realistic circumstances a patent holder would be better off by licensing downstream. This occurs when the licensing revenue can depend on the downstream value of the product either directly or through the use of ad-valorem royalties. Downstream licensing is also preferred by the patent holder when, instead, we assume that the downstream licensee is less informed about the validity of the patent. In most cases, downstream licensing increases allocative efficiency. However, it might reduce the manufacturer's incentives to invest and, thereby, decrease welfare. We characterize the circumstances under which a conflict arises between the stage at which patent holders prefer to license their technology and the stage at which it is optimal from a social standpoint that licensing takes place.

JEL codes: L15, L24, O31, O34.
keywords: Royalty Neutrality, Standard Setting Organizations, Patent Licensing, R&D Investment.

*We are grateful to Jorge Padilla, Pekka Sääskilahti, Gregor Langus, two anonymous referees, the editor, as well the audience at the JEI 2022 for comments on earlier version of the paper. The ideas and opinions in this paper, as well as any errors, are exclusively those of the authors. Financial support from ACEA and the Regional Government of Madrid through grant H2019/HUM-5859 is gratefully acknowledged. Comments should be sent to llobet@cemfi.es and damien.neven@graduateinstitute.ch.

1 Introduction

Technology licensing very often takes place along vertical chains in the production of a final product. The sale of a final product which makes use of a protected technology requires that one of the firms in the value chain obtains a license for the corresponding patents. The choice of the level at which licensing takes place has been controversial, specially with respect to Standard Essential Patents (SEPs) that must be licensed for the product to be sold in the final market.[1] One of the most salient examples has been the negotiation between the automobile industry and the holders of patents regarding mobile standards that are implemented in connected cars. Avanci, a patent pool that comprises the vast majority of IP holders, only offers a licence to the car manufacturers. Some car manufacturers, in particular Daimler, Ford and Tesla (and some component makers) originally contented that the licenses should be offered upstream, at the level of the component that implemented the features enabled by the technology. Even if most manufacturers ended up obtaining a licence from Avanci,[2] some component manufacturers remain dissatisfied with the outcome.[3]

To understand the implications of patent licensing at different levels in the supply chain, it is useful to start by identifying circumstances in which it does not matter. Layne-Farrar et al. (2014), develop the so-called *Royalty Neutrality* principle, showing that when information is public, royalties are charged per unit, firms are free to set prices for the goods that they sell, and negotiation among firms jointly maximizes the benefit of the parties involved, the level at which the royalty is set does not affect social welfare. This means that a patent holder cannot use the level at which the royalty is set

[1] See for instance, the Report of the SEP expert group (2020).

[2] Daimler complained to the European Commission claiming that patent holders were violating competition rules in 2019. Daimer was sued by Nokia for patent infringement in the Landgericht Düsseldorf. In the context of a request for a preliminary ruling, the Court asked a number of question to the European Court of Justice (ECJ) regarding the appropriate level at which licensing should take place (Case C-182/21). However, the ECJ never had the opportunity to address the issues as Nokia and Daimer settled in 2021. Daimler as well as Tesla and Ford eventually obtained a licence from Avanci

[3] Continental sued Avanci for refusal to licence in the US but its claim were rejected.

in a opportunistic manner to extract additional rents from downstream competitors or consumers. This result is based on the idea that the price of the input traded between upstream and downstream firms will adjust according to where the royalty is set; if the upstream firm licenses the technology, a part of the royalty rate will be passed through into the price of the input. If, instead, the royalty rate is paid by the producer in the final market, the price of the input will adjust downwards to accommodate the lower net revenues from the downstream buyer.[4]

In this paper, we consider common circumstances under which the previous neutrality result does not hold. We argue that when the value of the product is uncertain the patent holder has an incentive to license downstream. When investment by the downstream and upstream manufacturers are fixed, downstream licensing tends to improve social welfare. This occurs either because of a market expansion effect, resulting from better price discrimination, or because downstream licensing involves a mechanism of rent extraction that involves less distortion. However, when the effect of the choice of the level of licensing on the incentives to invest are taken into account, a conflict may arise; we show that when the sensitivity of investment to a shift towards downstream licensing is neither too high, not too low, society would be better off with upstream licensing even though the patent holder would prefer downstream licensing.

We propose a very simple setup where an upstream firm trades with a downstream producer for an homogeneous input required for a product in the final market. The quality of this product is determined by two components. Part of the value is due to the investment that both the upstream and downstream firm carry out. The value also has an exogenous and idiosyncratic component, which is unknown at the time at which investments take place. Firms negotiate the price at which they trade the input based on their bargaining power.

In this very simple setup, we derive our main trade-offs from the analysis of a simple

[4] The intuition for this effect is very closely related to the results on tax incidence in public finance (Dalton, 1936).

case in which licensing downstream might allow the patent holder to offer a different royalty rate depending on the value of the final product. This means that higher value products are associated with a higher payment. This kind of price discrimination is more costly to be carried out upstream, as the same input might be used for products with a different final-market value. As previously emphasized (i.e., Layne-Farrar et al. (2014)), allowing the royalty rate to vary according to the value of the product improves allocative efficiency. Lower value goods are only offered if a lower royalty rate is charged by the patent holder and this is more likely to occur when licensing takes place downstream. For a given level of investment by manufacturers, the patent holder will naturally prefer to license downstream and discriminate prices.

The previous basic insight is instrumental in understanding the effects of an ad-valorem royalty rate. As opposed to a per-unit royalty rate, it yields a revenue that depends on the value of the product even when this royalty rate cannot be directly adjusted to different realizations of the uncertain value of the product. Both upstream and downstream licensing allow for such discrimination. However, except when the downstream firm can make take-it-or-leave-it offer, the extraction of rents through upstream licensing is less effective than downstream licensing as it is mediated by the bargaining between the upstream and downstream manufacturer. Hence, downstream licensing allows the licensor to obtain a return that is more closely related to the value of the final product than if licensing took place upstream. This implies that downstream licensing leads to more effective price discrimination and increases patent holder profits. The royalty neutrality result thus fails when ad-valorem rates are used. Furthermore, consistent with what we observe in practice, the patent holder will also prefer ad-valorem over (constant) per-unit royalties.

Finally, we also explore the presence of asymmetric information regarding the validity of the patent within the vertical chain. Specifically, we consider the realistic case where the upstream firm is more likely to have the technological knowledge required to evaluate the validity of a patent than the firm that operates in the downstream market. This last firm often aggregates different components in the final product and, therefore, it might not have a specific knowledge of each of the underlying technologies.

If the technology is licensed upstream, a royalty will be paid when the patent is likely to be infringed, essential, and valid. In contrast, to the extent that the downstream firm cannot assess the validity of the patent, it has to decide whether to unconditionally pay the royalty rate or not. In that calculation, the downstream firm anticipates that if it refused to pay for the license and the patent turned out to be valid, production would not take place (i.e., an injunction would be imposed) and some legal costs would also be incurred. In those circumstances, the patent holder will offer a royalty rate such that downstream firm will prefer not to take the gamble that the patent might be invalid but be exposed to the risk of an injunction.

We show that under those circumstances, the patent holder also prefers to license downstream.[5] Licensing at that level implies a higher probability of obtaining licensing revenues but a lower royalty rate than what would be obtained with upstream licensing. The first effect dominates because a low royalty rate implies a lower expected distortion in the quantity sold than a much higher royalty rate that is paid only in some circumstances. In other words, licensing downstream provides a more efficient way to extract surplus from production and this benefits the patent holder.

Our analysis allows us to draw some implications from the preference of the patent holder to license downstream on welfare. For a given level of investment, the incentives of the patent holder to license downstream are mostly aligned with social welfare. Adapting the royalty payment to the value of the product improves allocative efficiency. Similarly, when the upstream firm is in a better position to assess validity and infringement, for a given revenue of the patent holder, social welfare is usually higher when licensing takes

[5] Legal costs would have no impact when licensing takes place upstream but they would yield a higher royalty rate downstream. This effect would reinforce the patent holder's preference for downstream licensing.

place downstream. The reason in this case is that the same revenue can be obtained by charging a low royalty rate downstream that is paid regardless of whether the patent is valid or not and a higher royalty rate that is paid by the upstream firm only if the patent is valid. The former is preferable because the magnitude of the distortion is convex in the royalty rate. Only when the expected royalty rate increases substantially as a result of downstream licensing welfare might decrease.

However, once we endogeneize the investment of the upstream and downstream firm we identify a second effect that operates in the opposite direction. To the extent that the patent holder can extract more surplus from the production of the good when licensing takes place downstream, incentives to innovate are undermined, reducing the endogenous quality of the product.

We characterize the terms of this trade off and show that when the technology is such that the sensitivity of investments to the choice of the level of licensing is neither too high nor too low, upstream licensing is more desirable from the point of view of society even though the patent holder would prefer downstream licensing.

The previous trade-off is quite general and, as we show in section 6, it holds when we relax some of the assumptions of the model. In particular, the results are robust to other informational assumptions including situations where the investment of the upstream and downstream firm are unobservable by the patent holder. Relative to the benchmark case, the only difference is that the extent to which the incentives to invest are undermined is less pronounced. Similarly, in the appendix we show that the results are robust to assuming that the patent holder does not know how the investment of each firm contributes to the value of the final product. In that case we show that, relative to the benchmark, the investment of the producer with stronger bargaining power increases (and that of the producer with weaker bargaining power might decrease). Finally, we also consider the effects of uncertainty in cost, rather than in value.

This paper is related to a growing literature aiming to understand the appropriate level at which licensing should take place in a vertical chain. The most common explanation is the existence of transaction costs (e.g., Langus and Lipatov (2022)). Licensing is more efficient if it takes place in the stage of production that is most concentrated. By negotiating with a small number of firms some costs are avoided. Papers like Padilla and Sääskilahti (2021) also argue that transaction costs are minimized when licensing takes place downstream as this facilitates the monitoring of the number of units sold. In contrast, Ivus et al. (2020) considers a setup like the one we propose in this paper and analyze the effect of presumptive patent exhaustion,[6] such that patent holder can opt out from patent exhaustion and license downstream as well as upstream. In their model, downstream licensing allows for price discrimination but involves additional transaction costs to learn the value of the product. They find that when transaction costs are neither too high nor to low, the patent holder engages in mixed licensing, with individual licensing for high valuation buyers and uniform licensing for low valuation buyers. In contrast with our paper, these authors abstract from the endogenous quality of the product.

Langus and Lipatov (2022) also analyze the appropriate level at which licensing should take place. However, they assume that because of regulation, the patent holder cannot affect the royalty rate. They consider the incentives for firms to invest along the value chain depending on the stage in which an (exogenous) license fee is levied. As they assume ad-valorem royalties, the Royalty Neutrality result does not apply in their setup. In contrast, we show that once we endogeneize the royalty rate the patent holder always prefers to license downstream and, by doing so, to induce a payment that depends on the value of the product in the final product market. For this reason, ad-valorem royalties have the same chilling effects on investment as in our baseline model.[7]

The rest of the paper is organised as follows. Section 2 introduces a simple model where royalty neutrality holds. In section 3 proposes a very simple case that allows us

[6]Patent exhaustion implies that when a technology is licensed upstream, the patent holder relinquishes all rights to the use of the technology downstream and, therefore, it cannot require a license in that stage.

[7]Sinitsyn (2021) analyzes the trade-off between per-unit and ad-valorem royalty rates in a setup where the firm sells a product line, combining the effects discussed here.

to illustrate the main insights of the paper. Section 4 studies the case of ad-valorem royalties and shows that the previous results apply. In section 5 we consider the case in which the downstream manufacturer has imperfect information on the validity of the patent. Section 6 discusses how the results change as we relax some of the assumptions. Section 7 concludes.

2 A Simple Model of Royalty Neutrality

A downstream firm, denoted as D, produces one unit of a final product that requires one unit of a component, which is produced by a unique upstream manufacturer, denoted as U. Firm U incurs a marginal cost of production $c < 1$, while the cost for firm D is normalized to 0.[8] Production also requires the use of a patented technology developed by a patent holder (or licensor), that we denote as firm L.

The product has a value for the downstream firm that is made up of two additive components, $X + \theta$. The component $X \leq c$ is deterministic and results from the investment in quality carried out by the downstream and upstream firms. The second component is random and arises from a uniform distribution, $\theta \sim U[0, 1]$.

The patent holder can charge a royalty rate both upstream $r_U \geq 0$ and downstream $r_D \geq 0$. After observing these rates, the upstream manufacturer and the downstream buyer bargain over the price p at which the component is traded. We assume that the rents from this negotiation are distributed to the upstream and the downstream firm in a proportion γ and $1-\gamma$, respectively. This allocation captures the bargaining power of each of the parties and might reflect the relative significance of competition at different stages of the production process. That is, the more firms could have produced the upstream component (the final product) the lower (higher) will be the value of γ.

The quality of the product X results from firms' investment. We assume that $X \equiv \beta x_U + (1 - \beta)x_D$, where x_U and x_D is the quality improvement brought about by the

[8]The results would be unchanged if the firm D also incurred in a constant marginal cost of production. In section 6.2 we analyze the implications of uncertain c.

investment of the upstream and downstream firm, respectively. The parameter $\beta \in [0,1]$ is the relative importance of upstream investment in innovation. Obtaining an improvement of size x implies the same cost for both firms, $C(x)$, assumed to be increasing, continuous, and convex in x.[9]

Finally, the timing of the model is as follows. In the first stage, quality investments by the upstream and downstream firm, x_U and x_D, are carried out. Then, the valuation θ is drawn. In the third stage, the patent holder determines the royalty rates, (r_U, r_D). In the last stage, the upstream and downstream firms negotiate the price for the component, p.

To solve the model, we start with the determination of the input price in the last stage of the game for a given value of X. The joint surplus of the upstream and downstream firms corresponds to $\theta + X - r_D - r_U - c$. Given the bargaining power of each of the parties we can characterize their profits gross of investment costs as

$$\pi_U(\theta, X, r_U + r_D) = \gamma(\theta + X - c - r_D - r_U), \tag{1}$$

$$\pi_D(\theta, X, r_U + r_D) = (1-\gamma)(\theta + X - c - r_D - r_U). \tag{2}$$

The price that supports this allocation is $p = \gamma(\theta + X - r_D) + (1-\gamma)(r_U + c)$. This implies that a higher upstream and/or a lower downstream royalty rate translate in the negotiation as a higher price for the input.

One of the implications of the previous result is that the final allocation depends only on the sum of the royalty rates that the upstream and downstream firm pay, $R \equiv r_U + r_D$. This result is a specific example of the *Royalty Neutrality Result* that was originally formulated in Layne-Farrar et al. (2014).[10] Hence, and without loss of generality, we only need to characterize the total royalty rate, R.

A transaction takes place whenever $\theta + X - c - R \geq 0$ so that a positive surplus

[9] The main results would go through even if both firms faced a different cost function. The implications of a higher upstream marginal cost of investment are qualitatively similar to having a lower β.

[10] The authors show that this result holds under more general conditions and it applies to general demand functions and market structures.

arises. Since θ is uniformly distributed, the royalty rate that the patent holder will choose emerges from the following profit maximization problem,

$$R^* = \arg\max_R \, (1 + X - c - R)R = \frac{1 + X - c}{2}. \tag{3}$$

As expected, this royalty rate is increasing in the quality of the product resulting from the investment of the upstream and downstream firm.

Finally, notice that the *Royalty Neutrality Result* also implies that the optimal investment choice by the upstream and downstream producer in the first stage of the model is independent of the stage where licensing takes place. In the interest of brevity, we defer the characterization of the equilibrium X to the next section, where this case constitutes a limit result.

To summarize, this benchmark case provides three important implications for the rest of the paper. First, as we just pointed out, the incentives for firms to invest are unaffected by the way in which a fixed total royalty rate is allocated between the upstream and the downstream firm. Second, to the extent that the royalty rate is positive, there will be a deadweight loss. Transactions should occur whenever $\theta \geq c - X$ but in equilibrium they will only take place if $\theta \geq c - X + R^*$. Finally, since R^* is increasing in X this implies that the patent holder appropriates part of the rents from the investment undertaken by the upstream and downstream firms. Together with the deadweight loss previously mentioned, this implies that since both firms do not enjoy all the benefit from their investment, the quality provided will be inefficiently low.

In the rest of the paper we relax some of the assumptions underpinning the previous result. We do that in two stages. In the next section we analyze a very stylized setup where the contractual arrangements that can be reached upstream and downstream are different. In further sections, we use the insights uncovered by our analysis to understand the effects in two relevant cases: the use of ad-valorem royalty rates and the consequences of asymmetries in the information regarding the validity of the patents.

3 Value-Contingent Licensing

The previous analysis assumed that the patent holder could not make the royalty rate dependent on the realized value of θ. Suppose now that while θ is always known to the upstream and downstream firms, it is only known to the patent holder with probability α.[11] This case provides intuition that will become useful in the rest of the paper.

The observability of θ determines the characteristics of the licensing contract that the patent holder can write in the following way. Regarding the downstream market, we assume that when the patent holder observes θ, it can offer a royalty rate $r_D(\theta) \geq 0$ that conditions on it. Otherwise, the downstream royalty rate is constant and denoted as r_D^0. In contrast, we assume that the upstream royalty rate can never depend on θ and it is denoted as r_U. These assumptions capture the asymmetry in the scope for discrimination in upstream and downstream licensing, which is a common feature in licensing, as recognized for instance by the SEPs Expert Group gathered by European Commission.[12] The asymmetry arises because the same component sold by the upstream manufacturer might have different downstream uses. It might thus be difficult to establish a different royalty rate for the same component based on the different uses because this information is not necessarily verifiable, particularly in multistage production processes where the upstream stage is further removed from the final product. Even if different royalty rates could be ascertained, differences across end uses may not be enforceable because of arbitrage that the patent holder cannot constraint due to legal constraint stemming from exhaustion (such that it looses control over the product using the technology after payment of the

[11] In section 6.1 we show that the results are qualitatively unchanged if the patent holder and the upstream firm have the same imperfect information.

[12] See, Group of Experts on Licensing and Valuation of Standard Essential Patents (Part II, section 3). https://ec.europa.eu/docsroom/documents/45217

royalties).[13] We assume that the manufacturers make investments without knowing if in the case of downstream licensing the patent holder will be able to observe θ and enforce contingent royalties. Whether the value of the product is observed by the patent holder or not is determined at the same time the valuation θ is drawn.

In this section we discuss the implications of this model of contingent royalties in terms of the stage in which licensing takes place and of the incentives for firms to invest.

3.1 Incentives to License Downstream

When the value θ is observable by the patent holder, this firm can extract all surplus from production by choosing a total royalty rate

$$r_D(\theta) + r_U = \theta + X - c.$$

Since royalty rates are assumed to be non-negative, it is always a weakly dominant strategy to choose $r_U = 0$ and $r_D(\theta) = \theta + X - c$. As a result, the price is $p = \theta + X - r_D(\theta) = c$. The distribution of bargaining power in this case is irrelevant. Notice that production will take place whenever $\theta \geq c - X$, yielding the efficient outcome.

The comparison of this case with the situation where θ is non-observable yields the following insights. First, royalty neutrality does not apply and, for a given value of X, the patent holder will prefer to license downstream, as it enables the extraction of the entire surplus. Second, the observability of θ eliminates the dead-weight loss from production. Third, when θ is observable, an increase in X translates into an equivalent increase in the royalty rate. As a consequence, the fact that θ is observable allows the patent holder to appropriate all the rents from the investment undertaken by upstream and downstream firms.

[13] This consideration is central in the analysis Ivus et al. (2020), which distinguishes between a regime of absolute exhaustion and a regime of presumptive exhaustion. In the former regime, there is no discrimination across uses and a single price for the component. In the later regime, a patent holder can set a royalty rate downstream that discriminates according to the value of the innovation, $r_D(\theta)$. Doing so implies a transaction cost. As a result, only for those uses for which θ is high a downstream royalty rate will be established. When the downstream value is low, a constant upstream royalty rate r_U will be offered.

3.2 Endogenous Investment Choice

In the first stage of the game upstream and downstream firms carry out their investments simultaneously. The expression for the profits of these two firms can be written as follows

$$\begin{aligned}\Pi_U(x_U, x_D, \alpha) =& (1-\alpha)\int_{R^*+c-X}^{1} \pi_U(\theta, \beta x_U + (1-\beta)x_D, R^*)d\theta - C(x_U),\\ \Pi_D(x_D, x_U, \alpha) =& (1-\alpha)\int_{R^*+c-X}^{1} \pi_D(\theta, \beta x_U + (1-\beta)x_D, R^*)d\theta - C(x_D).\end{aligned}$$

Since all the profits accrue to the patent holder when θ is observed, firms only obtain a revenue from the investment with probability $1-\alpha$. For a value θ the total return is allocated according to the firm's bargaining power, which is described in the profit functions (1) and (2). The case in section 2 corresponds to $\alpha = 0$.

Replacing the total royalty rate chosen by the patent holder in (3), profits for the upstream and the downstream firm can be written as

$$\Pi_U(x_U, x_D, \alpha) = (1-\alpha)\gamma\frac{(1+\beta x_U + (1-\beta)x_D - c)^2}{8} - C(x_U), \tag{4}$$

$$\Pi_D(x_D, x_U, \alpha) = (1-\alpha)(1-\gamma)\frac{(1+\beta x_U + (1-\beta)x_D - c)^2}{8} - C(x_D). \tag{5}$$

Notice that the previous functions are not necessarily concave in x_U and x_D, respectively. When this is not the case the optimal choice of investment might not be finite. The next assumption rules out this possibility and allows us to focus on situations where the investment levels are uniquely determined using the first order conditions. It also rules out a corner solution where the product is supplied regardless of the value of θ.[14]

Assumption 1. *The cost function $C(x)$ is sufficiently convex so that the profits of firm $i = U, D$ are always concave in x_i. Furthermore, $C(x)$ guarantees that the equilibrium total quality, X^*, is always lower than c.*

This assumption together with the previous expressions allow us to characterize the equilibrium investment level of each firm.

[14]The implications of relaxing this assumption are discussed at the end of the section 3.4.

Proposition 1. *The investment of the upstream and downstream firms are strategic complements. The equilibrium levels, x_U^* and x_D^*, are decreasing in α.*

The strategic complementarity between the decision of both firms implies that an increase in the investment of one of the firms raises the value of the investment of the other firm. This result is due to the fact that the profit functions of both upstream and downstream firms are convex in X, since a higher investment expands both the value of the product and the probability that it might be produced in equilibrium — i.e., for lower realizations of θ. This effect, in turn, increases the profitability of the investment of the other firm. This result has relevant implications for the effect of α on the quality of the product. A higher α discourages investment by directly reducing the return from the innovation. The complementarity between x_D and x_U exacerbates this effect, as a lower investment of one of the firms indirectly reduces the incentives for the other firm to invest as well.

The previous complementarity between both investments also implies that changes in β and γ would have, in principle, ambiguous effects on the total investment. For example, a higher bargaining power by the upstream producer fosters the investment of this firm at the expense of lowering the investment of the downstream producer which, in turn, depresses the incentives of the upstream firm to invest in the first place. This means that the consequences of different values of γ on the overall quality will greatly depend on the elasticity of the investment of each firm with respect to its bargaining power.[15]

3.3 Investments and welfare

The previous results also allow us to conclude that the profits of both the upstream and downstream firm are decreasing in α. For the case of the downstream firm, for example,

[15] This effect is akin to what occurs in the context of the Theory of the Firm where, in bilateral negotiations, the allocation of residual rights to one of the parties reduces the incentives of the other to invest (see Hart (1995)).

we can see that

$$\frac{d\Pi_D}{d\alpha}(x_D^*, x_U^*, \alpha) = -(1-\gamma)\frac{(1+X^*-c)^2}{8} + (1-\alpha)(1-\gamma)\beta\frac{1+X^*-c}{4}\frac{dx_U^*}{d\alpha} < 0,$$

so that the negative direct effect of the first term is enhanced by the decrease in the complementary investment of the other firm.

Using the equilibrium level of investment, we can now characterize the profits of the patent holder as

$$\Pi_L(\alpha) = \alpha \int_{c-X^*}^{1} (\theta + X^* - c)d\theta + (1-\alpha)\frac{(1+X^*-c)^2}{4} = \frac{1+\alpha}{4}(1+X^*-c)^2.$$

The previous expression identifies two sources of revenue for the patent holder. First, when θ is observable, the patent holder extracts all the surplus from the transaction. Second, when the patent holder cannot condition on θ, standard monopoly profits are realized. The total effect of α results from the combination of these two sources as

$$\Pi_L'(\alpha) = \frac{1}{4}(1+X^*-c)^2 + \frac{1+\alpha}{2}(1+X^*-c)\frac{dX^*}{d\alpha}. \tag{6}$$

From this expression we can observe that changes in α affect the profits of the patent holder in two ways that operate in opposite directions. The first term corresponds to a direct effect that captures the increase in profits derived from price discrimination, which allows the patent holder to extract all rents when θ is observable and enables the sale of the product when θ is low. The second term corresponds to the indirect effect, which indicates that an increase in α, by making the hold-up problem more acute, undermines the incentives for the upstream and downstream firm to invest, reducing the overall value of the product. To the extent that the royalty rates are increasing in X this second effect is detrimental to the patent holder's profits.

The next result characterizes how firm profits and social welfare are affected by changes in α.

Proposition 2. *The profits of the upstream and downstream firms are always decreasing in α. When $\frac{dX^*}{d\alpha}$ is sufficiently negative social welfare and the profits of all firms are decreasing in α. When $\frac{dX^*}{d\alpha}$ takes an intermediate value, patent holder profits are increasing*

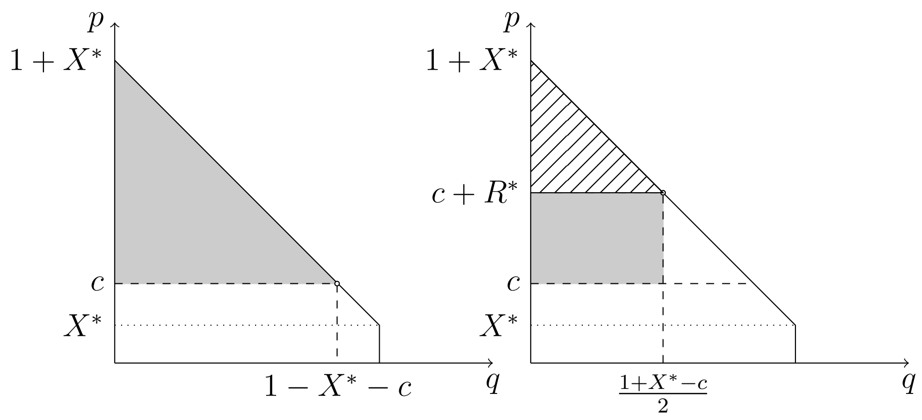

Figure 1: The grey area indicates the patent holder's profits when θ is observable (left) and when it is not (right) for a given value of X. The dashed area indicates the sum of the profits of the upstream and downstream firm.

but social welfare is decreasing in α. *When* $\frac{dX^*}{d\alpha}$ *is sufficiently close to* 0 *social welfare and patent holder profits are increasing in* α.

Figure 1 allows us to interpret the proposition as the result of the trade-off between a *market-expansion effect* and an *investment effect.* The figure on the left characterizes the case where θ is observable to the licensor. In that case, the quantity produced is efficient and all surplus is captured by the patent holder. In contrast, the figure on the right illustrates the case where θ is not known. The optimal royalty for the patent holder, R^*, implies that when $\theta \in [c, c+R^*)$ the good is not produced. This effect generates a dead-weight loss that is decreasing in α. For a given quality, the more the patent holder can condition on the realization of θ, as measured by a higher value of α, the higher the production and social welfare. This is the market-expansion effect.

The investment effect is driven by the returns that the upstream and the downstream firm can appropriate from increases in quality. When θ is observable these returns are 0, as shown in the figure, implying that firms have no incentives to invest. As a result, the equilibrium quality of the product, X^*, is driven by the profits that the upstream and downstream producer can obtain when θ is not known by the patent holder. A higher α decreases the returns from the investment which affects the value of the product in both states of the world.

The effect of α on the patent holder and the upstream and downstream firm can also be used to discuss the alignment of their incentives with those of society at large. It is useful to start by noting that social welfare can be decomposed as

$$W(\alpha) = \Pi_L(\alpha) + \Pi_U(\alpha) + \Pi_D(\alpha).$$

Since $\Pi'_U(\alpha) + \Pi'_D(\alpha) < 0$ it has to be that $W'(\alpha) < \Pi'_L(\alpha)$. In other words, whenever profits of the patent holder are decreasing in α, so will be social welfare.

This result implies that, as characterized in the previous proposition, there are three regions depending on the relevance of the investment by the upstream and downstream producer. In situations where this investment (and the associated quality of the product) is not very sensitive to the scope for contracting on θ, understood as a high value of α, then price discrimination will be in the interest of the patent holder and it will also be socially worthwhile ($W'(\alpha)$ and $\Pi'_L(\alpha)$ are both positive). This is the standard result that has been emphasized in the literature (Layne-Farrar et al., 2014). At the other extreme, when the investments (and associated quality of the product) are very sensitive to the scope for contracting on θ, the patent holder is not interested in discriminating prices and, as long as it can commit not to do so, it would be in its interest (and that of society) to preserve some of the firm rents ($W'(\alpha)$ and $\Pi'_L(\alpha)$ are both negative). When the sensitivity of investments (and associated quality of the product) takes an intermediate value, however, the interest of the patent holder and society diverge ($W'(\alpha)$ is negative and $\Pi'_L(\alpha)$ is positive). The market-expansion effect from increased price discrimination is not enough to overcome the loss in quality that it brings about. In contrast, the profits of the patent holder increase from price discrimination beyond the market-expansion effect, as it allows to extract more rents from the upstream and downstream firms. This means that it benefits from a higher value of α.

The following example, where only the investment of the downstream producer matters, illustrates the previous forces.

Example 1 (Downstream Investment Only)**.** *Assume* $\beta = 0$ *and also, in order for the bargaining power to be allocated efficiently, that* $\gamma = 0$*. Notice that this case also describes a perfectly competitive upstream market.*

As there will never be upstream investment, $X = x_D$*. Using* (5)*, we can characterize the effect of* α *on* x_D^* *as*

$$\frac{\partial x_D^*}{\partial \alpha} = \frac{\frac{1+x_D^*-c}{4}}{\frac{1-\alpha}{4} - C''(x_D^*)} < 0,$$

where concavity of the profit function and an interior result requires $C''(x_D^*) > \frac{1-\alpha}{4}$*.*

Equation (6) *allows us to show that the profits of the patent holder are increasing in* α *when*

$$\frac{\partial x_D^*}{\partial \alpha} > -\frac{1 + x_D^* - c}{2(1+\alpha)}.$$

Social welfare can be computed as $W(\alpha) = \frac{3+\alpha}{8}(1 + x_D^* - c)^2 - C(x_D^*)$ *and we have that* $W'(\alpha) < 0$ *if*

$$\frac{\partial x_D^*}{\partial \alpha} < -\frac{1 + x_D^* - c}{4(1+\alpha)}.$$

This implies that when the effect of α *on the magnitude of the innovation takes an intermediate value,*

$$\frac{\partial x_D^*}{\partial \alpha} \in \left(-\frac{1 + x_D^* - c}{2(1+\alpha)}, -\frac{1 + x_D^* - c}{4(1+\alpha)}\right],$$

the profits of the patent holder are increasing in α *while the effect on social welfare is negative.*

3.4 Socially-Optimal Licensing Level

The previous result has implications regarding the choice of the appropriate level at which licensing should take place and whether the choice could be left to the patent holder or not. When the investments (and associated innovation) is sufficiently sensitive to the scope for discrimination, social welfare would increase locally if α were reduced. It is also possible that in those circumstances, social welfare could be enhanced if price discrimination was prevented altogether. This can be achieved in the context of our model by imposing that licensing only takes place upstream. Indeed, if the patent holder is constrained to set the royalty rate upstream, whether θ is observable or not is irrelevant as the licensing contract cannot be used to engage in price discrimination. The social welfare obtained if upstream licensing is imposed is also, for any given value of the parameters, the social welfare that is achieved if one imposes $\alpha = 0$.

When it is socially optimal to prevent discrimination, two cases arise regarding the incentives of the patent holder. First, it may very well be that the patent holder would also obtain higher profits without discrimination. This is more likely to occur when the investment is highly sensitive to price discrimination so that the profits of the patent holder decrease locally as α increases. In those circumstances, the patent holder would like to commit not to discriminate and if a regulator establishes that licensing should take place upstream, the patent holder would not object. Second, the patent holder may be better off with discrimination. This is more likely to arise, as discussed above, for intermediate values of the sensitivity of investments with respect to discrimination (so that locally, the profit of the patent holder increases with α). In those circumstances, if the choice of the level of licensing is left to the patent holder, it will take place downstream even though it would be socially optimal to license at the upstream level.

Whether there is a conflict between the choice of the patent holder and the choice that maximizes social welfare depends on the parameters and, in particular, on the shape of the cost function. To illustrate this discussion, let's return to Example 1, where we assumed that $\gamma = \beta = 0$. Furthermore, suppose that the cost function is quadratic $C(x) = \frac{k}{2}x^2$. In that case, the equilibrium innovation chosen by the downstream producer corresponds to

$$x_D^* = \frac{(1-\alpha)(1-c)}{4k + \alpha - 1},$$

for $k > \frac{1}{4}$. We can verify that $\frac{\partial^2 x_D^*}{\partial k \partial \alpha} > 0$. Since $\frac{\partial x_D^*}{\partial \alpha} < 0$, this means that investment becomes less sensitive to α as k increases. Following the previous discussion, we can then show that for low values of k, $k < \frac{3}{4}$, both patent holder profits and social welfare increase when licensing takes place upstream and price discrimination is not possible. At the

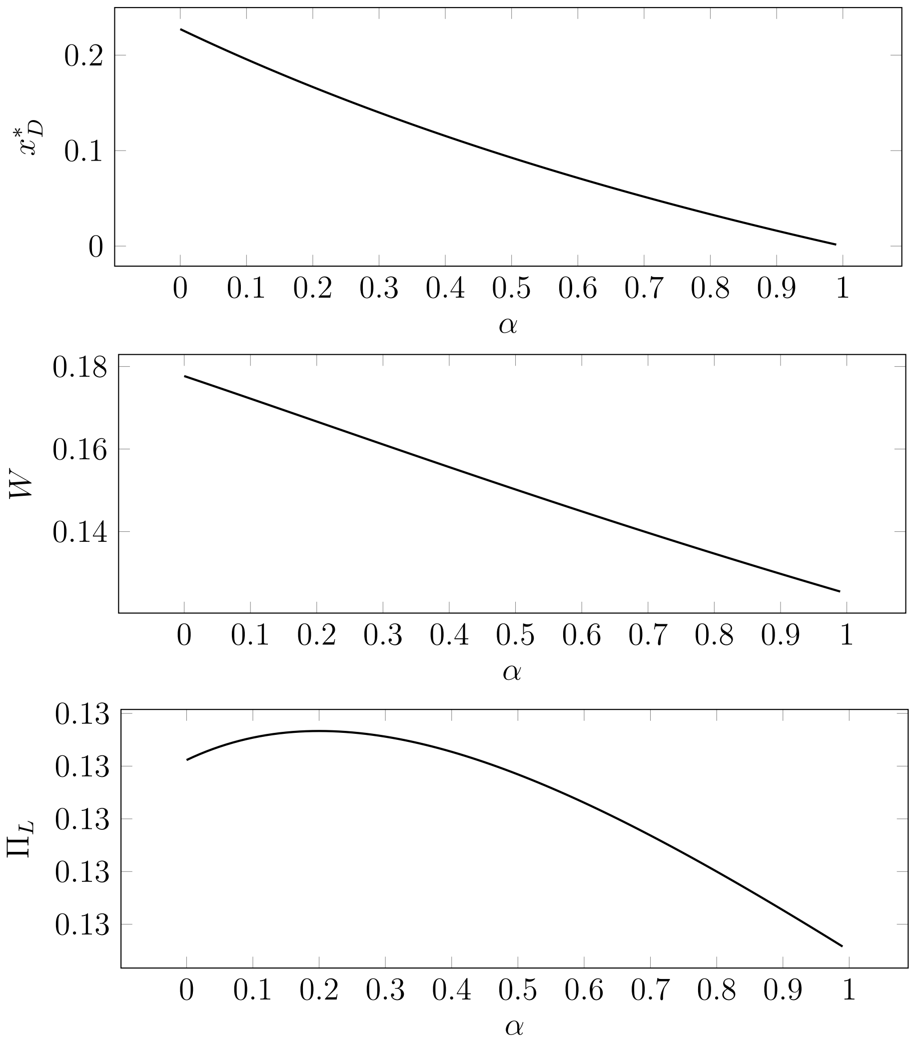

Figure 2: Equilibrium investment of the downstream firm, total welfare, and profits of the patent holder when $C(x) = \frac{k}{2}x^2$ and with parameter values $\gamma = \beta = 0$, $c = \frac{1}{2}$, and $k = \frac{4}{5}$. Under this parameterization, $x_U^* = 0$.

other extreme, when $k > \frac{5}{4}$, innovation responds very little to k and price discrimination increases both patent holder profits and social welfare. In the intermediate region, when $k \in \left[\frac{3}{4}, \frac{5}{4}\right]$, $W(\alpha)$ is strictly decreasing in α and $\Pi_L(\alpha)$ is maximized at $\alpha = 4k - 3 > 0$. Thus, there is a range of value of α for which social welfare is maximized when upstream licensing is imposed (so that welfare with $\alpha = 0$ is obtained) but for which the patent holder would benefit from discrimination. Figure 2 illustrates this case.

Before concluding this section, it is worth pointing out that our assumption that $X \leq c$ downplays the conflict between the stage in which licensing should take place from a social stand point and the optimal choice of the patent holder. In particular, this assumption implies that, for a given level of investment, value contingent licensing always enhances social welfare, as it increases total production and eliminates the dead-weight loss. In situations where the optimal level of X is sufficiently larger than c even when a patent holder could not condition on θ the downstream firm would still license the technology for all values of the final product. As a result, even though the patent holder might still benefit from downstream licensing, social welfare would always be maximized using non-contingent upstream licensing.

4 Ad-Valorem Royalties

Per-unit royalties — understood as a payment for each unit sold — are becoming the norm in some industries. Nevertheless, royalty rates that are expressed as a percentage of the price of the product — denoted as ad-valorem royalties — are still relevant. In this section, we show that under ad-valorem royalties the patent holder will earn more profits by licensing downstream. As in the previous section, this result arises because downstream licensing allows the patent holder to extract rents more effectively.[16]

We return to the case where θ is not observable to the patent holder. We characterize ad-valorem royalties as a percentage s_U and s_D of the revenue that the patent holder extracts from the upstream and/or the downstream firm, respectively. This means that the total surplus over which firms bargain can be written as

$$\pi_U(\theta, X, s_D, s_U) + \pi_D(\theta, X, s_D, s_U) = (1 - s_D)(\theta + X) - s_U p - c,$$

where

$$\pi_U(\theta, X, s_D, s_U) = (1 - s_U)p - c,$$

$$\pi_D(\theta, X, s_D, s_U) = (1 - s_D)(\theta + X) - p,$$

[16] The difference between both kinds of royalties has been the object of recent interest in the literature. Llobet and Padilla (2016) consider a context where the value of the innovation is always known and there is no hold-up risk. The authors show that ad-valorem royalties tend to imply lower final prices because they decrease the distortions caused by double-marginalization and mitigate the royalty-stacking problem that arises when several patent holders have complementary patents that are necessary to sell a final product.

are the counterparts of (1) and (2) under ad-valorem royalties. This total surplus can be interpreted as follows. The downstream firm keeps a proportion $1 - s_D$ of the final value of the product, $\theta + X$. Similarly, the upstream firm must pay to the patent holder a proportion s_U of the endogenous price, p, at which it transacts the input with the downstream producer.

Given that the upstream firm has a bargaining power γ, the price resulting from Nash Bargaining can be characterized as

$$\max_{p} \ \pi_U(\theta, X, s_D, s_U)^{\gamma} \pi_D(\theta, X, s_D, s_U)^{1-\gamma}$$

or

$$p^*(\theta) = \gamma(1 - s_D)(\theta + X) + (1 - \gamma)\frac{c}{1 - s_U}. \tag{7}$$

As in the case of per-unit royalties, this price is decreasing in the downstream royalty rate and increasing in the upstream rate. The product will be sold as long as $(1-s_D)(\theta+X) - s_U p^*(\theta) - c \geq 0$ or

$$\theta \geq \theta^* \equiv \frac{c}{(1 - s_D)(1 - s_U)} - X. \tag{8}$$

Notice that this threshold value is independent of γ. The reason is that since the negotiation is efficient, the valuations of the product that are served depend on the aggregate surplus that they generate and not on how the rents are allocated between the upstream and the downstream producer. We can compute the profit maximizing combination of ad-valorem royalty rates as the result of

$$\Pi_L = \max_{s_U, s_D} \int_{\theta^*}^{1} \left[s_U p^*(\theta) + s_D(\theta + X)\right] d\theta. \tag{9}$$

The next result provides a characterization of the optimal combination of royalty rates.

Proposition 3. *Under ad-valorem royalties the patent holder always chooses $s_U^* = 0$ when $\gamma < 1$. The optimal downstream royalty s_D^* is decreasing in c, increasing in X, and independent of γ. When $\gamma = 1$ the profits of the patent holder depend only on $(1 - s_D)(1 - s_U)$.*

This proposition indicates that royalty neutrality does not hold for any $\gamma < 1$. Furthermore, the patent holder always prefers to charge only a downstream royalty. The intuition is that ad-valorem royalties allow the patent holder to obtain a revenue that depends on the value of the product even when this royalty rate cannot be directly adjusted to different values of θ. When this royalty is applied upstream, the revenue depends on p^* is obtained from the value $\theta + X$ but this relationship is mediated by the bargaining power of each party. In contrast, when the royalty rate applies downstream the revenue of the patent holder can be directly associated to the value of the product in the final market. This situation allows for a more effective surplus extraction that enhances the profits of the licensor. For precisely this reason, ad-valorem royalties either upstream or downstream are always superior for the licensor to per-unit royalty rates that are not value contingent.[17]

A higher marginal cost reduces the net value of the product (i.e., it increases θ^*) and it raises the distortions generated by the downstream royalty rate. It is, therefore, optimal to decrease s_D. A higher value of the product has the opposite effect. Finally, the royalty rate is independent of γ. The reason is that, as pointed out before, θ^* is independent of γ.

Interestingly, royalty neutrality is recovered when $\gamma = 1$. In that case, the input price becomes proportional to the value of the product, $p^* = (1 - s_D)(\theta + X)$, allowing the patent holder to extract a constant proportion of the value of the product. Replacing this price in the profit function of the patent holder in (9) we obtain that for $\gamma = 1$

$$\Pi_L = \max_{s_U, s_D} \int_{\theta^*}^{1} \left[s_U(1 - s_D) + s_D \right] (\theta + X) d\theta = \int_{\theta^*}^{1} \left[1 - (1 - s_U)(1 - s_D) \right] (\theta + X) d\theta,$$

which only depends on the $(1 - s_D)(1 - s_U)$.

The results regarding investment derived in the previous section applies mutatis mutandis to the case of ad-valorem licensing. Indeed, as the previous result illustrates, imposing ad-valorem royalties downstream — as opposed to the upstream producer —

[17] In section 6.2 we show that upstream royalties might be optimal when costs are uncertain.

allows the patent holder to better discriminate. This means that the same trade-offs uncovered in the previous section between better rent extraction and ex ante incentives to invest also arise in this case.

5 Uncertain Validity of the Patent

The benchmark model assumed that the patent to be licensed was valid. However, in practice there is often uncertainty about this validity, particularly in the case of SEP holders that own a small portfolio of weak patents.[18]

Potential licensees are heterogeneous in their capacity to ascertain whether a patent is valid or not. It is likely that larger and more research-oriented firms have better knowledge of the underlying technology. Similarly, an upstream firm, being closer to the technology itself, is more likely to possess the necessary knowledge.[19] This is in contrast with the downstream producer, who might buy a component that already embeds the technology and adapt it to its own needs.[20] In this section, we focus on the latter dimension of heterogeneity and we analyze its implications.[21]

To make the previous notion operational, we now extend the benchmark model to accommodate the possibility that the patent is invalid. We assume that the validity status is known to the upstream firm. However, the downstream producer only knows that the patent is valid with probability $\phi \in (0, 1)$. This setup allows us to study how asymmetric information affects the incentives for the patent holder to offer a royalty rate

[18]Potential licensees might also be uncertain about whether the patent has been infringed and/or it is essential to the standard. The implications are very similar to the case of uncertain validity and the results in this section can be understood as the combination of all these sources of private information.

[19]See Group of Experts on Licensing and Valuation of Standard Essential Patents (Part II, section 3), for evidence in support of this view.

[20]In industries with multiple layers, the technology is more likely to be part of the core business for firms producing a component up the chain. In the downstream market, multiple components are integrated in the same final product, with less precise knowledge of each specific technology.

[21]Our model is only meant to capture the existence of an asymmetry as a stylised fact. This is not to deny that there might be situations where a small number of technology-savvy downstream producers buy their input from an atomized upstream layer. In that case, the downstream informational advantage might lead to opposite results. One would also expect downstream firms to improve their knowledge over time by investing in specific assets, for instance by hiring IP experts from upstream component manufacturers.

upstream or downstream. To make this case consistent with the rest of the paper, we also assume that the realization of θ occurs after the validity of the patent is determined.

If licensing takes place upstream only, the potential licensee will obviously refuse to pay a positive royalty rate when the patent is invalid. When the patent is valid, however, the result coincides with the benchmark model. That is, the downstream firm will acquire the input if $\theta + X - c - r_U > 0$ and this means that the royalty rate that maximizes profits for the patent holder can be computed as

$$r_U^* = \arg\max_{r_U} \phi r_U(1 + X - c - r_U) = \frac{1 + X - c}{2}.$$

Suppose now that the patent is licensed downstream. If the uninformed downstream firm accepts the license, its profits for a given realization of θ would become

$$\pi_D(\theta, X, r_D) = (1 - \gamma)(\theta + X - c - r_D).$$

If instead, the firm refuses to license a patent that turns out to be valid, it will not be allowed to produce. Furthermore, the firm might incur an additional cost $M \geq 0$ (e.g., legal fees). In that case, the expected profits of the downstream producer can be written as

$$\hat{\pi}_D = -\phi M + (1 - \phi)(1 - \gamma) \int_{c-X}^{1} (\theta + X - c) d\theta.$$

As a result, the downstream firm will accept any royalty rate $r_D \leq \tilde{r}_D$, where the latter is defined as

$$\int_{c+\tilde{r}_D-X}^{1} \pi_D(\theta, X, \tilde{r}_D) d\theta = \hat{\pi}_D. \tag{10}$$

Notice that $\tilde{r}_D$ is increasing in ϕ and M. That is, the more likely it is that the patent is valid or the higher the cost of not obtaining a license it if it turns out to be valid, the higher the royalty rate that the downstream firm will be willing to accept.

Under downstream licensing, the patent holder will choose the royalty rate that results from

$$r_D^* = \arg\max_{r_D \leq \tilde{r}_D} r_D(1 + X - c - r_D) = \min\left\{\tilde{r}_D, \frac{1 + X - c}{2}\right\}.$$

It is clear that if $r_D^* = \frac{1+X-c}{2}$, so that $r_D^* = r_U^*$ for example because M is sufficiently high, the patent holder will be better off licensing the patent downstream, as the royalty payment would be received even when the patent is invalid. The next result extends the previous intuition and shows that the patent holder is always better off under downstream licensing.

Proposition 4. *For a given X, downstream licensing maximizes profits for the patent holder for all values of ϕ and M.*

When $r_D^* = \tilde{r}_D$ the patent holder faces a trade-off. Licensing downstream implies a higher probability of obtaining licensing revenues but a lower royalty rate than what would have been obtained with upstream licensing. The first effect dominates because a low royalty rate implies a lower expected distortion in the quantity sold than a much higher royalty rate that is paid with probability ϕ. In other words, licensing downstream provides a more efficient way to extract surplus from production and this benefits the patent holder.

The previous result hinges on the assumption that by going to court the downstream producer risks not being able to produce, enticing the firm to accept a higher r_D. Alternatively, we could think of situations where the patent holder could request an injunction that, while the legal process is resolved, prevents production and reduces the revenues in a proportion $\delta \leq 1$. In case the court sides with the downstream producer the patent holder would pay a compensation for the foregone profits. Otherwise, a new royalty rate would be determined $\hat{r}_D$. This alternative specification would result in profits

$$\hat{\pi}_D = -\phi M + (1-\delta)\phi(1-\gamma)\int_{c-X-\hat{r}_D}^{1}(\theta + X - \hat{r}_D - c)d\theta + (1-\phi)(1-\gamma)\int_{c-X}^{1}(\theta + X - c)d\theta.$$

These profits would approximate the case discussed above when δ is sufficiently high or if the new negotiated (or court-mandated) royalty rate, $\hat{r}_D$, can depend on θ and contribute to extracting more surplus from production.

In this section, we endogenize the level of investments by manufacturers and show that as in the model with value-contingent licensing the choice of the patent holder will be inefficient from a social perspective. However, this case also yields some new insights.

5.1 Investment and uncertainty on validity

We now consider the optimal investment when there is uncertainty on validity and infringement. From Proposition 4 we know that under downstream licensing the patent holder will always obtain higher profits. When this increase comes at the expense of the profits of the producers in both stages, downstream licensing will generate a negative *investment effect*.

The next result captures a countervailing force that we denote the *royalty-allocation effect*. For a given total revenue for the patent holder, a royalty rate downstream spreads out the burden over the two states of the world, when it is valid and when it is not. As a result, distortions are reduced and social welfare raises.

Proposition 5. *Consider a downstream royalty r_D so that the patent holder obtains the same revenue as in the case where it licenses upstream at a rate r_U^*. Under downstream licensing welfare increases and so do the investment incentives.*

The royalty rate set by the patent holder increases firm costs and, as a result, it generates a distortion since products with a low θ become unprofitable. This distortion is convex in the royalty rate. This means that, for the same expected revenue for the patent holder, it is always preferable that the same royalty rate is paid in all states of the world, even those in which the patent is invalid. As this can only occur when the licensee does not know the validity of the patent, this effect favors downstream licensing.

The combination of the two forces means that, in some circumstances, upstream licensing will be socially optimal: on the one hand, as shown by Proposition 5, downstream licensing is attractive from a social perspective to the extent that it reduces distortions. Specifically, investments and the profits of upstream and downstream firms (and, hence, social welfare) would be higher under the downstream royalty that would keep the patent holder indifferent with its optimal upstream royalty. On the other hand, the patent holder would optimally charge a royalty downstream higher than the one that would keep it indifferent. This reduces investment and welfare. When the investment is sufficiently sensitive to the rents of upstream an downstream firms, the second effect dominates, so that upstream licensing is preferred from a social perspective.

The fact that under downstream licensing the patent holder can extract rents more effectively, means that there are circumstances in which its incentives will not be aligned with social welfare. The patent holder might prefer downstream licensing when upstream licensing is socially optimal. Using the intuition from the previous sections, this might be the case when the investment has an intermediate sensitivity with respect to the rents accrued by the manufacturers.

Figure 3 illustrates the previous discussion using the same specification we formulated in Example 1. In that case we focused on the situation where $\beta = \gamma = 0$ so that all bargaining power was in the hands of the downstream firm and only its investment mattered. We characterize the equilibrium royalty rate and profits of the downstream firm and the patent holder (the upstream firm makes zero profits given $\gamma = 0$) for different probabilities that the patent is valid with probability, ϕ. We assume no legal costs, $M = 0$ and fix X. As expected, when the patent is licensed upstream, the royalty rate does not depend on the validity of the patent, although it is paid only when it is valid. In contrast, under downstream licensing the royalty rate increases in the probability that the patent is valid. When ϕ is sufficiently high the constraint determined by $\tilde{r}_D$ is not binding and the patent holder chooses the monopoly rate $r_D^* = \frac{1+X-c}{2}$.

Regarding firm profits, when $\phi = 1$ royalty neutrality applies and both cases are equivalent. At the other extreme, when $\phi = 0$, profits are also identical for all firms, since the upstream producer would pay the monopoly rate with probability 0 and the downstream firm would always prefer to take its chances in court than to pay a positive royalty rate.[22] For intermediate values, and consistent with Proposition 4, the patent holder is better off under downstream licensing, since the lower distortions it generates allows the firm to increase the royalty payments. For the same reason, the downstream producer is worse off.

In terms of welfare, under downstream licensing these lower profits undermine the incentives to innovate. This negative effect must be balanced out with the effect on social welfare of downstream licensing for a given value of X as observed in the last panel of the figure. As expected, downstream licensing is superior when $\tilde{r}_D$ is sufficiently lower than $r_U^* = \frac{1+X-c}{2}$ as in this parameter range the investment effect is limited. However, when the constraint is not binding and $r_D^* = \frac{1+X-c}{2}$ welfare is unambiguously higher under upstream licensing. In this range, there is no benefit from reducing distortions by spreading royalty payments through a lower rate across all states of the world. This implies that when ϕ is sufficiently high (but less than one) there is no trade-off and upstream licensing is always superior, regardless of the importance of firm investment.

Consider now an increase in M. In that case, the patent holder will increase the royalty rate, r_D^*, under downstream licensing but not when licensing takes place upstream. This implies an increase in the profits that the patent holder obtains from downstream licensing (relative to upstream licensing) and increases the extent to which downstream licensing undermines investment and social welfare.

The combination of the two effects uncovered here allows us, more generally, to draw conclusions similar to those discussed above regarding the ability to price discriminate. In that case, the socially optimal regime considered a trade-off between the market-expansion and the investment effect and it implied that the patent holder tended to choose a licensing stage which, while expanding the market, also undermined the incentives for firms to invest, reducing social welfare. Here, we have shown that asymmetric information provides a similar trade-off, such that downstream licensing, while often

[22] Remember that given the parametric assumptions of this example, the upstream firm always makes 0 profits

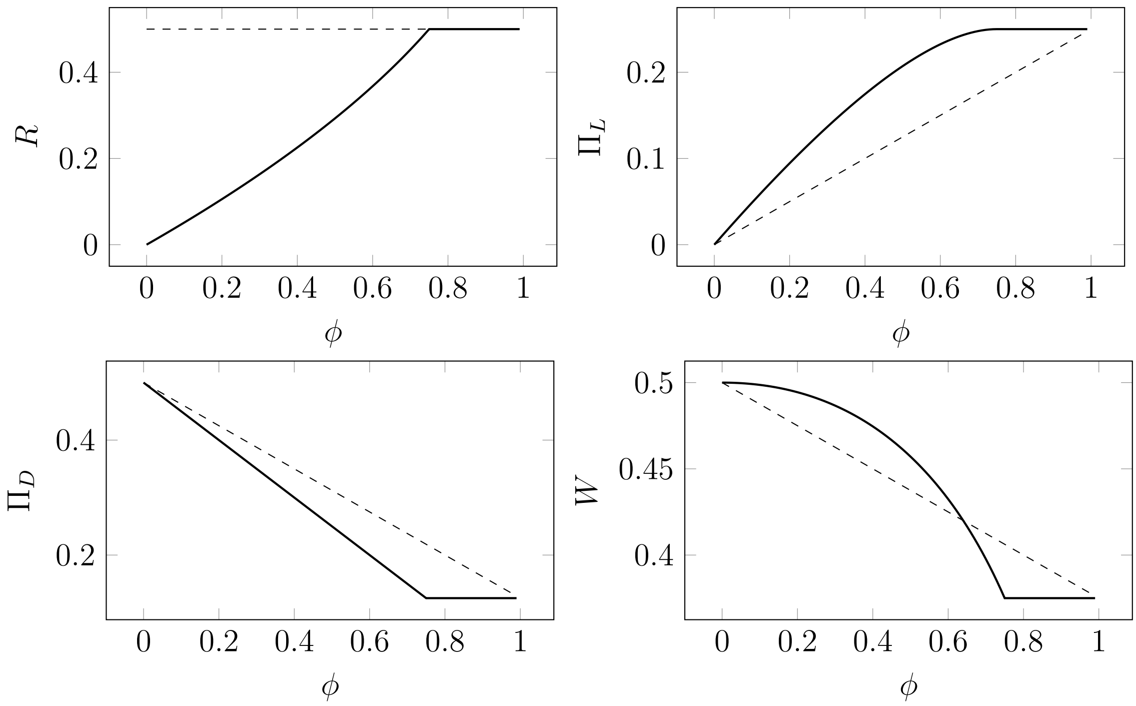

Figure 3: For a given X, equilibrium total royalty rate, patent holder profits, downstream profits, and social welfare for different values of ϕ under downstream licensing (solid line) and upstream licensing (dashed line). The parameterization assumes $M = 0$, $X = c = 0.5$, and $\gamma = \beta = 0$. For this reason, the upstream producer makes zero profits in either case and social welfare, gross of investment cost, is $W = \Pi_L + \Pi_D$.

reducing distortions, also undermines the incentive to invest.

6 Robustness and Extensions

In this section we discuss extensions of our framework. Specifically, for the model of value-contingent licensing, we consider the possibility that the upstream manufacturer may have the same information than the patent holder. In appendix A.1, we also consider the possibility that asymmetric information arises regarding the cost of the upstream manufacturer. In another extension we explore the case where the marginal cost of production is uncertain.

6.1 Inefficient Price Negotiation

A maintained assumption throughout the paper has been that θ was always observable both by the upstream supplier and the downstream producer. This assumption implied that there was never an inefficiency in the negotiation of the input price. We now relax this assumption and consider the case where the upstream supplier always has the same information as the patent holder. We do so in the context of the model of value-contingent licensing. That is, suppose that the upstream firm also observes the exogenous component of the value of the product, θ, with probability $1-\alpha$. For simplicity, we assume that the upstream firm has all the bargaining power, $\gamma = 1$. Relative to the benchmark model, the outcome only changes when θ is not observed by the patent holder (and, thus, the upstream firm).

Given a combination (r_U^0, r_D^0), the fact that $\gamma = 1$ is equivalent to assuming that the upstream firm chooses the price to maximize

$$\max_p \, (p - r_U^0 - c)(1 + X - p - r_D^0).$$

As a result, the monopoly price offered by the upstream supplier corresponds to $p^*(r_U^0, r_D^0) = \frac{1+X+r_U^0-r_D^0+c}{2}$.

This price is internalized by the patent holder who now chooses a combination of royalties to maximize profits as follows:

$$\max_{r_U^0, r_D^0} \, (r_U^0 + r_D^0)\left(\frac{1 + X - r_U^0 - r_D^0 - c}{2}\right),$$

resulting in an equilibrium total royalty rate $R^* = r_U^{0*} + r_D^{0*} = \frac{1+X-c}{2}$. This means that the Royalty Neutrality result holds in this environment when $\alpha = 0$.

In this case we can compute the profits of the upstream and downstream firms as

$$\begin{aligned}
\pi_U(X) =& (p^* - r_U^{0*} - c)\frac{1 + X - c}{4} = \frac{(1 + X - c)^2}{16}, \\
\pi_D(X) =& \int_{p+r_D^{0*}-X} (\theta + X - p^* - r_D^{0*})d\theta = \frac{(1 + X - c)^2}{32}.
\end{aligned}$$

As expected, the equilibrium outcome exhibits, for a given X, a double-marginalization distortion that reduces total profits compared to (4) and (5).

Notice, though, that the effects of double marginalization are not straight-forward. For a given value of X the deadweight loss is higher in this case, so that the option of preventing price discrimination becomes relatively less attractive from a social perspective.

At the same time, the decrease in the profits of the upstream and downstream producer undermines the overall incentives for firms to invest.

In addition, when θ is observed the upstream supplier can extract all the surplus from the transaction and discourage the downstream firm from investing. This is a kind of hold-up is similar to what we discussed in the case of the patent holder. When β is low, so that the investment of the downstream producer is particularly relevant, the fact that the upstream supplier cannot observe θ might provide commitment value, increasing total quality and social welfare.

6.2 Uncertain Cost of the Upstream Manufacturer

Consider now a variation of the model where $\theta \leq 1$ is known to all firms but the cost of the upstream manufacturer is uncertain, with $c \sim U[0,1]$. As in the basic model in section 3 we assume that c is observable to the patent holder with probability α but it is always known to the upstream supplier and the downstream producer.

When the patent holder does not observe c the total royalty rate is chosen as a result of

$$R^* = \arg\max_R \, (\theta + X - R)R = \frac{\theta + X}{2}.$$

In contrast, when c is observable to the patent holder, it is natural to argue that conditioning on the upstream cost of production is easier through the royalty rate that applies to that firm. In the limit case, this means that the patent holder would extract all the surplus by choosing a royalty rate upstream $r_U(c) = \theta + X - c$.

The previous discussion shows that the results in this case are the mirror image of those when θ was uncertain. That is, when the relevant dimension of uncertainty is the production cost, upstream licensing reduces the static dead-weight loss but, by doing so, it can undermine the incentives for firms to invest. Notice, however, that this result would not change the implications in terms of the misalignment of preferences between the patent holder and society as a whole.

As in the benchmark model, the previous result can also be used to illustrate the effect of cost uncertainty on the usage of ad-valorem royalties. In particular, when the patent holder uses upstream and downstream royalty rates s_U and s_D, respectively, production will take place when

$$c^* \leq (1 - s_D)(1 - s_U)(\theta + X),$$

where this expression is the counterpart of (8). The problem of the patent holder then becomes

$$\Pi_L = \max_{s_U, s_D} \int_0^{c^*} [s_U p^*(c) + s_D(\theta + X)] \, dc, \tag{11}$$

where $p^*(c)$ is defined in (7) after we have replaced the dependency of θ by c.

The next result characterizes the optimal royalty rates and equilibrium production.

Proposition 6. *Suppose that $\theta < 1$ is constant and $c \sim U[0, 1]$. Under ad-valorem royalties, the patent holder chooses $s_D^* = 0$ and $s_U^* = \frac{1}{2}$ when $\gamma < 1$. Production takes place if $c \geq \frac{\theta + X}{2}$.*

In the benchmark model downstream licensing allowed the patent holder to relate the payment directly to the value of the innovation. Under upstream licensing this value was imperfectly captured by the price and was mediated by the bargaining power of each of the firms. In the case of cost uncertainty, the result is more extreme. Upstream licensing works in a similar way, and it is still true that the payment amount depends on the realization of the uncertainty (in this case on costs), mediated by the bargaining power of each party. Downstream licensing, however, implies a constant payment $s_D(\theta + X)$ independent of the cost of the product.

The comparison of the two sources of uncertainty leads to another interesting insight. Ad-valorem royalties are a more powerful instrument to discriminate prices when there are differences in the value. The reason is that royalty rates can never condition directly on the cost.

Finally, notice that the optimal ad-valorem royalty rate is independent of the value of the product, $\theta + X$. This result, however, is an artifact of the uniform distribution assumed for c and that does not necessarily translate to other distributions. It is also worth pointing out that as in the case where the value of the innovation is uncertain, royalty neutrality is recovered when $\gamma = 1$ since, in that case, the input price is $p = \theta + X$ and conditioning on this price or the final value is equivalent.

7 Concluding Remarks

In this paper we have analyzed the effect of licensing in different stages of the production process on the investment that producers carry out. We have shown that, under some realistic assumptions, downstream licensing tends to benefit the patent holder. In contrast, upstream licensing tends to foster investment. In industries where investment by upstream and downstream firms is sufficiently relevant, there will be a conflict between the decision of the patent holder and that of society as a whole.

The kind of circumstances that we have considered in this paper are likely to be very relevant in practice, particularly in industries where the technology arises from an standardization process. In that case, upstream components resulting from this technology are embedded in a variety of final products. Downstream licensing is likely to allow price discrimination in ways that reduce the returns from production. At the same time, downstream producers are also more likely to have very limited knowledge of the underlying technology and their investment will typically focus on its use and the integration with their own technology. This limited information will negatively affect their ability to negotiate a license.

An important caveat of this paper is that we have abstracted from the incentives of the patent holder to develop the technology in the first place. This is obviously a critical element of these industries. Our results ought to be understood as how the same technology should be licensed for different uses depending on the investment required for its integration by the firms that produce the good.

References

Dalton, Hugh, *Principles in Public Finance*, Routledge and Kegan Paul Ltd, 3rd edition, 1936.

Hart, Oliver, *Firms, Contracts and Financial Structure*, Oxford, 1995.

Ivus, Olena, Edwin L.-C. Lai and Ted Sichelman, "An economic model of patent exhaustion," *Journal of Economics & Management Strategy*, October 2020, *29*(4), pp. 816–833.

Langus, Gregor and Vilen Lipatov, "Efficient Level of SEPs Licensing," February 2022, CESifo Working Paper.

Layne-Farrar, Anne, Gerard Llobet and Jorge Padilla, "Patent Licensing in Vertically Disaggregated Industries: The Royalty Allocation Neutrality Principle," *Communications and Strategies*, 2014, *95*(3), pp. 61–84.

Llobet, Gerard and Jorge Padilla, "The Optimal Scope of the Royalty Base in Patent Licensing," *Journal of Law and Economics*, 2016, *59*(1), pp. 45–73.

Padilla, Jorge and Pekka Sääskilahti, "Revisiting the Bizarre SEP Level of Licensing Antitrust Controversy," *Journal of European Competition Law & Practice*, 06 2021, *12*(9), pp. 680–687.

Sinitsyn, Maxim, "Licensing Mechanisms for Product Lines," 2021, mimeo.

A Appendices

A.1 Private Information on the Endogenous Quality

In the benchmark model the patent holder could observe X and set the royalty rate accordingly. This had important implications as the upstream and the downstream producer could anticipate that a higher X would result in a higher royalty rate, undermining the incentives to invest. In this section we relax the previous assumption and we explore the situation where the investment is not observable to the patent holder and holdup is mitigated. To understand its implications we distinguish two cases. In the first, we assume that the rest of the parameters of the model are known. In the second, we go further and we discuss the situation where the parameter β is also private information and, therefore, the patent holder cannot anticipate with certainty the equilibrium value of X. We do so in the value-contingent context discussed in section 3, where the patent holder can make the licensing contract downstream depend on θ with probability α.

A.1.1 The Value of X is Private Information

When X is private information, in choosing the royalty rate, the patent holder entertains the belief that the quality level obtained is $\hat{X}$. Given this belief, the optimal royalty rate can be characterized in the same way as in the benchmark model. This implies that when θ is known, the patent holder can extract all the surplus and charge a total royalty rate $r_U(\theta) + r_D = \theta + \hat{X} - c$. When θ is not observable, the total royalty rate becomes $R^*(\hat{X}) = \frac{1+\hat{X}-c}{2}$. Notice that in this section we are making explicit the dependence of the royalty rate on the belief about the quality.

Following the same arguments used in the value-contingent licensing model, we can characterize the investment choice of the upstream and downstream producer, respectively, as a result of their profit maximization problem

$$\max_{x_U} (1-\alpha)\gamma \frac{(1+\beta x_U + (1-\beta)x_D - R^*(\hat{X}) - c)^2}{2} - C(x_U), \tag{12}$$

$$\max_{x_D} (1-\alpha)(1-\gamma) \frac{(1+\beta x_U + (1-\beta)x_D - R^*(\hat{X}) - c)^2}{2} - C(x_D). \tag{13}$$

In equilibrium it must be that $\hat{X} = \tilde{X}$ where $\tilde{X}$ is the quality level attained as a result of the firm's investment. The next result shows that when X is not observable to the patent holder, firms will tend to choose a higher equilibrium investment than when X is known.

Proposition 7. *When X is not observable to the patent holder, the equilibrium investment of all firms increases compared to the benchmark model. Investments decisions are strategic complements. The equilibrium royalty rate when θ is unobservable increases.*

The fact that in the benchmark model the royalty rate increases in X implies that even in the case where θ is not known to the patent holder, there is some degree of hold up. The upstream and downstream firms reduce their investment in the anticipation that part of the returns from a higher X will be extracted by the patent holder. This effect is mitigated when X is not known. However, since innovation increases in that case the royalty rate that the patent holder charges also increases.

A.1.2 The Value of β is Private Information

We now discuss the case where the parameter determining the weight of the upstream and downstream levels of investment, β, is unknown to the patent holder. Notice first that, due to the neutrality result that holds both in the case where θ is known and when it is not, observing X is enough for the patent holder to choose the royalty rate in the benchmark model. Hence, to make the private information on β meaningful, we also assume that X is private information.

To simplify the discussion, we analyze a particular case where $\beta = 0$ with probability $\frac{1}{2}$ and $\beta = 1$ otherwise. For consistency with the rest of the model, we assume that β is realized before the investment is carried out. This means that we will denote the optimal investment of each firm as a function of the realized β, $x_i^*(\beta)$ for $i = U, D$. Finally, we consider the case where $\gamma < \frac{1}{2}$ so that in the benchmark model we would have that $x_D^*(0) > x_U^*(1)$ and $x_D^*(1) = x_U^*(0) = 0$.[23]

When the patent holder does not observe θ the total royalty rate R will be set to maximize expected profits

$$R^*(\hat{x}_U, \hat{x}_D) = \arg\max_R \ \frac{1}{2}(1 + \hat{x}_D - c - R)R + \frac{1}{2}(1 + \hat{x}_U - c - R)R = \frac{1 + \frac{\hat{x}_D + \hat{x}_U}{2} - c}{2}.$$

where $\hat{x}_D$ and $\hat{x}_U$ are the expected effort choices of the upstream and downstream firm when $\beta = 0$ and $\beta = 1$, respectively. Due to the uncertainty, the royalty rate adjusts to the average expected quality.

We can now characterize the profit function of the downstream firm — the profit function of the upstream firm would be symmetric — as

$$\max_{x_D} \ (1-\alpha)\frac{1-\gamma}{2}\left[\int_{c+R^*-x_D}^{1} (\theta + x_D - c - R^*(\hat{x}_U, \hat{x}_D))d\theta \right.$$
$$\left. + \int_{c+R^*-x_U}^{1} (\theta + x_U - c - R^*(\hat{x}_U, \hat{x}_D))d\theta\right] - C(x_D).$$

This expression includes two revenue terms that depend on the realization of β. The first term captures the expected revenue from the product when only the downstream investment x_D is useful, while in the second term only x_U matters. Notice that this last term does not affect the incentives for the downstream producer to innovate since it does not depend on x_D.

The first order condition of the previous problem becomes,

$$(1-\alpha)(1-\gamma)\frac{1 - x_D^* - R^*(\hat{x}_U, \hat{x}_D) - c}{2} - C'(x_D^*) = 0,$$

which is relevant if $C(x)$ is sufficiently convex (i.e., under Assumption 1). A similar expression determines the optimal investment of the upstream producer, where $1 - \gamma$ is

[23]Notice that, due to the assumption that both firms have the same cost function, $\gamma > \frac{1}{2}$ would lead to $x_D^*(0) < x_U^*(1)$. The case $\gamma = \frac{1}{2}$ is uninteresting since it would imply that there is no uncertainty on the total quality produced as a function of β.

replaced by γ. Since $\gamma < \frac{1}{2}$ we have that, as expected, $\tilde{x}_D(0) > \tilde{x}_U(1)$, where we identify the equilibrium choice with a tilde to distinguish it from the benchmark case and, as before, the parenthesis indicates the realization of β.

Importantly, and compared to the previous case, the investment decisions of both firms here are independent. This is due to a combination of two assumptions. First, the extreme values of β imply that both firms do not invest at the same time in equilibrium. Second, as R^* depends only on the expected investments, a change in x_D (in x_U) does not affect the royalty payment of the upstream (downstream) firm.

The comparison with the case where X and β are known can be broken down in two parts: the effect on the innovation and the distortions that private information brings about. The next result indicates that the downstream firm will always invest more as a result. The implications for the upstream firm, however, are less clear-cut.

Proposition 8. *Compared to the benchmark case, when X and β are not observable by the patent holder, the downstream producer will always increase investment. In contrast, the upstream producer will only increase investment if γ is sufficiently close to $\frac{1}{2}$.*

To interpret the previous result, it is useful to point out once more that the incentives for firms to invest arise in circumstances when θ is private information. In that case, since $\gamma < \frac{1}{2}$, we have that the equilibrium quality level when $\beta = 1$ is smaller than when $\beta = 0$, $\tilde{x}_D(0) > \tilde{x}_U(1)$. The combination of private information on X and β has two main implications. First, the royalty rate set by the patent holder is not affected by the actual choice of investment of the upstream and downstream producer. This effect was studied in section A.1.1 and it implies higher incentives for firms to invest. Second, the fact that the patent holder cannot choose a royalty rate that adjusts to the realized X implies that when only the upstream investment matters, the payment will be too high. The royalty rate is determined as a function of expected investment choice and the investment of the upstream firms when $\beta = 1$ is less than investment of the downstream firm when $\beta = 0$. This effect reduces incentives for the upstream firm to invest (but increases the incentive to invest of the downstream firm). While the first effect is general, the second is likely to be small when γ is close to $\frac{1}{2}$ as in that case the choice of the royalty rate would be close to optimal in both states of the world.

Regarding the output distortions that arise as a result of private information, it is important to point out that they will be higher when X and β are unknown. When θ is private information, we know that the impossibility to adjust the royalty rate to the case when the upstream investment matters makes the double-marginalization problem more severe and, as a result, it increases the dead-weight loss. The royalty rate is comparatively lower when only the downstream investment matters which leads to a lower distortion in this case. However, the convexity of the social welfare function with respect to the price implies that the first effect dominates.

We now turn to the royalty rate that the patent holder chooses when θ is known. A big difference in this case is that since $\tilde{x}_D(0) > \tilde{x}_U(1)$, for a given realization of θ the product has a different value depending on whether $\beta = 0$ or $\beta = 1$. This means that the patent holder has two options. First, it can set a royalty rate to cater the whole market regardless of the realization of β, $r^D(\theta) = \theta + \tilde{x}_U(1) - c$. Alternatively, it can set a higher royalty rate so that consumers buy only when the upstream innovation is relevant, $r^D(\theta) = \theta + \tilde{x}_D(0) - c$. The comparison of the two cases is such that serving the market regardless of the realization of β is optimal if

$$\theta \geq c + \tilde{x}_D(0) - 2\tilde{x}_U(1).$$

That is, for low values of θ it is optimal for the patent holder to focus only on the case where $\beta = 0$. All the market is served otherwise. When γ is small the difference between $\tilde{x}_D(0) - \tilde{x}_U(1)$ is likely to be high. As a result the social loss that arises when the patent holder decides to sell only when $\beta = 0$ will arise for a large range of values of θ.

A.2 Proofs

This section includes the proof of all the results.

Proof of Proposition 1: The optimal level of investment of the upstream manufacturer and downstream producer, x_U^* and x_D^*, can be obtained as

$$(1-\alpha)\gamma\beta\frac{1+X^*-c}{4} - C'(x_U^*) = 0,$$
$$(1-\alpha)(1-\gamma)(1-\beta)\frac{1+X^*-c}{4} - C'(x_D^*) = 0,$$

where $X^* = \beta x_U^* + (1-\beta)x_D^*$. An interior solution requires $C''(x_U^*) > \frac{(1-\alpha)\gamma\beta^2}{4}$ and $C''(x_D^*) > \frac{(1-\alpha)(1-\gamma)(1-\beta)^2}{4}$, respectively. Notice that the cross-derivative of the profit function of firm $i \in \{D, U\}$ with respect to x_j^* for $j \neq i$ is positive, indicating that the functions are supermodular and the investments are strategic complements. This result together with the fact that

$$\frac{\partial x_U^*}{\partial \alpha} = \frac{\gamma\beta\frac{1+X^*-c}{4}}{\frac{(1-\alpha)\gamma\beta^2}{4} - C''(x_U^*)} < 0,$$
$$\frac{\partial x_D^*}{\partial \alpha} = \frac{(1-\gamma)(1-\beta)\frac{1+X^*-c}{4}}{\frac{(1-\alpha)(1-\gamma)(1-\beta)^2}{4} - C''(x_D^*)} < 0,$$

allow us to conclude that the investment of both firms is decreasing in α. □

Proof of Proposition 2: The negative effect of α on profits is described in the text.

Social welfare can be written as

$$\begin{aligned} W(\alpha) &= \int_{c-X^*}^{1} (\theta + X^* - c)d\theta - (1-\alpha)\frac{(1+X^*-c)^2}{8} - C(x_U^*) - C(x_D^*) \\ &= \frac{3+\alpha}{8}(1+X^*-c)^2 - C(x_U^*) - C(x_D^*). \end{aligned}$$

The derivative with respect to α, once we apply the first order condition that determines x_U^* and x_D^* becomes,

$$W'(\alpha) = \frac{(1+X^*-c)^2}{8} - \{(3+\alpha) - (1-\alpha)\left[\gamma\beta + (1-\gamma)(1-\beta)\right]\}\frac{1+X^*-c}{4}\frac{dX^*}{d\alpha}.$$

This derivative is negative if

$$\frac{dX^*}{d\alpha} < -\frac{1+X^*-c}{2\{(3+\alpha) - (1-\alpha)\left[\gamma\beta + (1-\gamma)(1-\beta)\right]\}}.$$

This condition is satisfied whenever $\Pi_L'(\alpha) < 0$ which, using (6), occurs when

$$\frac{dX^*}{d\alpha} < -\frac{1+X^*-c}{2(1+\alpha)}.$$

By the same token, when $W'(\alpha) > 0$, which occurs when $\frac{dX^*}{d\alpha}$ is sufficiently close to 0, $\Pi_L'(\alpha) > 0$.

$\square$

Proof of Proposition 3: The case with $\gamma = 1$ is analyzed in the main text. Suppose that $\gamma < 1$. Given the expression of θ^* we can write the upstream royalty as

$$s_U \equiv \sigma(s_D, \theta^*) = 1 - \frac{c}{(1-s_D)(\theta^* + X)}.$$

This allows us to rewrite the profits function of the patent holder in (9) as

$$\max_{s_D, \theta^*} \int_{\theta^*}^{1} \left[\sigma(s_D, \theta^*) p^*(\theta) + s_D(\theta + X)\right] d\theta.$$

This expression can be rewritten as

$$\max_{s_D,\theta^*} \int_{\theta^*}^{1} \left[(1-s_D)(\gamma(\theta+X) + (1-\gamma)(\theta^*+X)) - c\left(\gamma\frac{\theta+X}{\theta^*+X} + (1-\gamma)\right) + s_D(\theta+X)\right] d\theta.$$

We can compute the derivative of this expression with respect to s_D as

$$\frac{1}{2}(1-\gamma)(1-\theta^*)^2 \geq 0,$$

meaning that when $\gamma < 1$ for any θ^* profits are maximized by setting the maximum s_D which, in turn, implies that $s_U^* = 0$.

Given $s_U^* = 0$ we can now rewrite the profit of the patent holder only as a function of s_D as

$$\max_{s_D} \int_{\frac{c}{1-s_D} - X}^{1} s_D(\theta + X) d\theta,$$

with first-order condition

$$\int_{\frac{c}{1-s_D} - X}^{1} (\theta + X) d\theta - \frac{s_D c^2}{(1-s_D)^3} = 0.$$

Since the left-hand side of this expression is always decreasing in c and increasing in X we have that the profit function is submodular in c and supermodular in X. $\square$

Proof of Proposition 4: As profits for the patent holder when licensing downstream increase in M it is enough to show the result for $M = 0$.

Using (10), we can solve for $\tilde{r}_D = (1+X-c)\left(1-(1-\phi)^{\frac{1}{2}}\right)$. Notice that $\tilde{r}_D \leq \frac{1+X-c}{2}$ if $\phi \leq \frac{3}{4}$. That is, a necessary condition for licensing upstream to be optimal is that $\phi < \frac{3}{4}$. Replacing $\tilde{r}_D$ in the profit function of the patent holder, we obtain

$$\Pi_L^D = \tilde{r}_D(1+X-c-\tilde{r}_D) = (1+X-c)^2(1-\phi)^{\frac{1}{2}}\left(1-(1-\phi)^{\frac{1}{2}}\right),$$

where Π_L^D stands for the downstream licensing profits. These profits are higher than those that arise from upstream licensing, $\Pi_L^U = \phi\frac{(1+X-c)^2}{4}$ if $\phi \leq \frac{8}{9}$, proving the result. □

Proof of Proposition 5: Consider the optimal upstream royalty rate r_U^* and a royalty rate downstream, r_D, so that the patent holder indifferent. That is,

$$\phi r_U^*(1+X-c-r_U^*) = r_D(1+X-c-r_D).$$

Notice that this implies that $r_U^* > r_D$ (we ignore the case where r_D is above the monopoly royalty rate) and, therefore, $1+X-c-r_U^* < 1+X-c-r_D$. This, in turn, requires $\phi r_U^* > r_D$.

We now compare the joint profits for the upstream and downstream.

- When the royalty rate is charged upstream,

$$V_U \equiv \Pi_D + \Pi_U = (1-\phi)\int_{c-X}^{1}(\theta+X-c)d\theta + \phi\int_{c-X-r_U^*}^{1}(\theta+X-c-r_U^*)d\theta.$$

- When the royalty rate is charged downstream,

$$V_D \equiv \Pi_D + \Pi_U = \int_{c-X-r_D}^{1}(\theta+X-c-r_D)d\theta.$$

$V_D > V_U$ if and only if $r_D < \phi r_U$. □

Proof of Proposition 6: Following the same strategy as in the proof of Proposition 3, we start by rewriting the downstream royalty rate as

$$s_D \equiv \sigma(s_U, c^*) = 1 - \frac{c^*}{(1-s_U)(\theta+X)}.$$

Replacing in the profit function of the patent holder we obtain

$$\max_{s_U, c^*} \int_0^{c^*} [s_U p^*(c) + \sigma(s_U, c^*)(\theta + X)]\, dc.$$

Using the expression for $p^*(c)$ and $\sigma(s_U, c^*)$, these profits can be rewritten as

$$\max_{s_U, c^*} \int_0^{c^*} \left[\theta + X + \frac{(1 - s_U\gamma)c^* - (1-\gamma)c}{1 - s_U}\right] dc.$$

The derivative with respect to s_U yields the expression

$$\frac{(1-\gamma)c^*}{2(1-s_U)^2} \geq 0.$$

This means that when $\gamma < 1$ for any c^* the upstream royalty should be as high as possible and, therefore, $s_D^* = 0$. As a result, we can rewrite the profit function as

$$\max_{s_U} \int_0^{(1-s_U)(\theta+X)} s_U \left[\gamma(\theta + X) + (1-\gamma)\frac{c}{1-s_U}\right] dc.$$

The derivative with respect to s_U implies $s_U^* = \frac{1}{2}$. Applying the optimal royalty rate to c^* means that production takes place when $c \geq \frac{\theta+X}{2}$. □

Proof of Proposition 7: The first order condition characterizing the optimal investment of the upstream innovator, $\tilde{x}_U$ and $\tilde{x}_D$, can be written as

$$(1-\alpha)\gamma\beta\frac{1+\tilde{X}-c}{2} - C'(\tilde{x}_U) = 0,$$
$$(1-\alpha)(1-\gamma)(1-\beta)\frac{1+\tilde{X}-c}{2} - C'(\tilde{x}_D) = 0,$$

where we have already imposed the equilibrium outcome that $\hat{X} = \tilde{X} = \beta\tilde{x}_U + (1-\beta)\tilde{x}_D$. As the first term in both expressions is higher than the counterparts in the benchmark case, obtained in the proof of Proposition 1 we obtain the desired result. It is immediate that the investment of each firm is increasing in the investment of the other one.

Finally, notice that since the rule determining the optimal royalty rate for the patent holder is the same as when X is known, higher investment translates into a higher royalty rate. □

Proof of Proposition 8: The first order condition that characterizes the investment of the upstream and downstream producer can be obtained as

$$(1-\alpha)\frac{1-\gamma}{8}\left(2-2c+3\tilde{x}_D(0)-\tilde{x}_U(1)\right)-C'(\tilde{x}_D)=0,$$
$$(1-\alpha)\frac{1-\gamma}{8}\left(2-2c+3\tilde{x}_U(1)-\tilde{x}_D(0)\right)-C'(\tilde{x}_U)=0.$$

The comparison with the first order conditions in the proof of Proposition 1 and the fact that $\tilde{x}_D(0) \geq \tilde{x}_U(1)$ allows us to conclude that $\tilde{x}_D(0) > x_D^*$. The comparison regarding the investment of the downstream producer is, in general, ambiguous. Notice, however that if γ is sufficiently close to $\frac{1}{2}$ we have that $\tilde{x}_D(0) \geq \tilde{x}_U(1) > x_D^* \geq x_U^*$. $\square$