Investment and Patent Licensing in the Value Chain*

Gerard Llobet
CEMFI and CEPR

Damien Neven
The Graduate Institute Geneva, CEPR, and Compass Lexecon

May 25, 2022

Abstract

We analyze the welfare effects of patent licensing at different stages of the production chain and the level at which a patent holder would choose to license. We consider the incentive to invest in enhancing the value of the final product at the different stages of production. We study the effects of allowing a patent holder to discriminate among different products downstream and/or accounting for differences in information that potential licensees at different stages might have about the validity of the patent. We show that in those circumstances, 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. Whereas the patent holder usually prefers to target downstream firms, society is often better off if upstream firms take the license.

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

*The ideas and opinions in this paper, as well as any errors, are exclusively those of the authors. Financial support from ACEA 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 in which licensing takes place has been controversial in particular with respect to Standard Essential Patents (SEPs) that must be licensed for the product to be sold in the final market.[1] For instance, regarding mobile standards, some holders of essential patents have advocated for a license to be acquired by final good producers, while these firms have often argued for licensing to take place upstream. But does it really matter whether the license is taken by an upstream or a downstream firm?

It is useful to identify circumstances in which it does not. In this respect, 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 maximises the benefit of the parties involved, the way in which the royalty rate is structured does not affect affect social welfare. This means that a patent holder cannot use the royalty structure in a opportunistic way to affect market structure or 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 accordingly; if the upstream firm licenses the technology, a part of the royalty rate will be passed through into the price of the input. If the royalty rate is paid by the producer in the final market, the price of the input will adjust downwards to accommodate the lower willingness to pay of the downstream buyer.[2]

In this paper, we consider two sets of arguably common circumstances under which the previous neutrality does not hold; we allow the patent holder to discriminate among

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

[2]The intuition for this effect is very closely related to the results on tax incidence in public finance.

different products downstream and we allow for differences in information that potential licensees might have about the validity of the patent, depending on the stage where they operate. We show that in those circumstances, a conflict arises between the stage at which patent holders want offer their technology and the stage at which it is optimal from a social perspective for this licensing to take place. In particular, whereas the patent holder usually prefers to license the technology downstream, society is often better off if upstream firms take the license.

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 analyze the effects of licensing the technology upstream or downstream. We assume that 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 possible when licensing takes place downstream. 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.

The previous trade-off implies that for those products for which the value is mostly driven by the investment carried out by firms along the value chain, upstream licensing is more desirable from the point of view of society. This is in contrast with what determines the incentives for the patent holder to license at difference stages. This firm will often inefficiently choose to license downstream since it internalizes all the gains from price discrimination but only some of the losses from the reduction in investment (the rest being incurred by the upstream and downstream firm).

The previous result is quite general and, as we show in section 4, 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 in this case is that the extent to which the incentives to invest are undermined is less pronounced. Similarly, 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).

The paper also explores the presence of asymmetric information regarding the validity of the patent within the vertical chain. In section 5 we consider the realistic case where the upstream firm is more likely to have the knowledge of the technology necessary 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 implemented) and some legal costs would also be incurred. As a result, the patent holder will offer a royalty rate sufficiently low to be accepted.

We show that, abstracting from transaction and legal costs, for a given revenue of the patent holder social welfare is higher when licensing takes place downstream. The reason 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. Due to the convex cost of distortions related to an increase in the production cost, the former option is socially preferred. For the same reason, the patent holder prefers to charge a higher average royalty rate under downstream licensing and obtains higher profits.[3] Under this higher royalty rate, however, the welfare implications of downstream licensing are less clear.

Again, once we endogeneize the investment of the firms along the value chain the results change. As in the price discrimination case discussed above, to the extent that licensing downstream reduces the profits for the upstream and downstream firm, their investment incentives are reduced. This is more so when the legal costs that the downstream firm would bear when going to court increase. In that case, profits for the patent holder are likely to be substantially higher when licensing takes place downstream, creating a bigger divergence with the socially optimal licensing stage.

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

[3]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.

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 analyse the effect of presumptive patent exhaustion, such that patent holder can opt out from patent exhaustion and licence 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 transactions cost 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 however 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) licence fee is levied. Contrary to our paper, they assume that royalty rates are ad-valorem (e.g. proportional to the price of the product) either downstream or in both stages. In that case, the Royalty Neutrality result does not hold. Although ad-valorem royalty rates are typically more efficient (Llobet and Padilla, 2016), it is often the case that the contracts establish per-unit royalty rates as assumed in this paper. In section 4.4 we allow for ad-valorem royalties in our setup. We show that they induce a payment that depends on the value of the product in the final market and, for this reason, they have the same chilling effects on investment than in our baseline model.[4]

The rest of the paper is organised as follows. Sections 2 and 3 present the model and characterize the equilibrium. Section 4 discusses how the results change as we relax some of the assumptions. Section 5 studies the effects of licensing when there is private

[4]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.

information on the validity of the patent. Section 6 concludes.

2 The Model

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.[5] The final product also makes use of a patented technology developed by a patent holder (or licensor), that we denote as firm L.

The value of the product for the downstream firm has two additive components, $X+\theta$. The component X 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 endogenous quality component is known to all firms. In contrast, we assume that θ is always known to the upstream and downstream firms but it is only known to the patent holder with probability α. Alternatively, the parameter α can be interpreted as the probability that the exogenous value of the product is contractible.[6]

The patent holder can charge a royalty rate both upstream $r_U \geq 0$ and downstream $r_D \geq 0$. The observability of θ determines the characteristics of these royalty rates. 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. This asymmetry is a reflection of various common features of vertical contracting. First, the same component sold by the upstream manufacturer might have different downstream uses. This implies that it is difficult to establish a different royalty rate for the same component based on the differ-

[5]The results would be unchanged if the firm D also incurred in a constant marginal cost of production. In section 4.3 we analyze the implications of private information on c.

[6]In section 4.1 we show that the results are qualitatively unchanged if the patent holder and the upstream firm have the same imperfect information.

ent uses because this information is not necessarily verifiable, particularly in multistage production processes where the upstream stage is further removed from the final product or because of the possibility of arbitrage.[7]

The upstream manufacturer and the downstream buyer bargain over the price at which the component is traded, p. In this negotiation we assume that the rents 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 is the result from the investment of the upstream and downstream firm. 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 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 $C(x)$, increasing, continuous, and convex in x.[8]

Finally, the timing of the model is as follows. In the first step, quality investments by the upstream and downstream firm, x_U and x_D, are carried out. Then, the valuation θ is drawn and observed by the patent holder with probability α. In the third stage the patent holder determines the royalty rates, (r_U, $r_D(\theta)$) or (r_U,r_D^0), depending on whether θ is observable or not. In the last stage, the manufacturer and the downstream firm negotiate the price for the component, p.

In this model, the patent holder will thus appropriate part of the rents from the investment undertaken by the upstream and downstream firms. Even when θ is not

[7]The same assumption is made in Ivus et al. (2020). In that case a patent holder can set a royalty rate downstream that discriminates on 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.

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

observed, the patent holder will determine a royalty that increases with the investment undertaken by the downstream and upstream firms. However, when θ is observed, the patent holder can extract the entire rent downstream so that an increase in α exacerbates the extent to which the patent holder can appropriate rents from the innovation.[9]

3 Equilibrium of the Model

We characterize the equilibrium in two steps. First, we discuss the price and royalty rate that emerge in the last two stages of the game. We later discuss how these variables determine the quality investments that arise in equilibrium.

3.1 Royalty-Rate Determination

As usual, we start with the determination of the input price in the last stage of the game for a given total quality X. When θ is observable to the patent holder, which occurs with probability α, this firm can extract all the 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 $p = \theta + X - r_D(\theta) = c$. Negotiation power in this case is irrelevant.

When θ is not observable to the patent holder, which occurs with probability $1-\alpha$, the joint surplus of the upstream and downstream firm corresponds to $\theta + X - r_D^0 - r_U - c$. Based on the bargaining power of each of the firms and the zero outside option if no production takes place, we can determine the equilibrium price as $p = \gamma(\theta + X - r_D^0) +$

[9]Enforcement of FRAND rates might reduce the extent to which rents can be extracted when θ is observed. This depends on the specifics of the enforcement procedure (see Langus et al. (2013)) In this respect, α can be interpreted as the probability that royalties can be made contingent on the value of the product and enforced.

$(1-\gamma)r_U$. In that case, profits from the transaction of the component become

$$\pi_U(\theta, X, r_U + r_D^0) = \gamma(\theta + X - c - r_D^0 - r_U), \tag{1}$$

$$\pi_D(\theta, X, r_U + r_D^0) = (1-\gamma)(\theta + X - c - r_D^0 - r_U), \tag{2}$$

for the upstream and downstream firm, respectively. As in the case where θ was observable, it is important to note that the final allocation depends only on the sum of the royalty rate that the upstream and downstream firm pay. This result is a specific example of the *Royalty Neutrality Result* that was formulated in Layne-Farrar et al. (2014). Hence, and without loss of generality, we will only refer to the total royalty rate in this case as $R \equiv r_U + r_D^0$.

When θ is not observable, there is a transaction whenever $\theta + X - c - R \geq 0$ so that a positive surplus arises. Since θ is uniformly distributed, the royalty rate that the patent holder will choose emerges from the following profit maximization problem,

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

As expected, this royalty rate is increasing in the quality of the product resulting from the investment of the upstream and downstream firm. However, as opposed to the case where θ was observable, an increase in X does not translate into an equivalent increase in the royalty rate. As a consequence, the fact that the patent holder cannot observe θ reduces the extent it can appropriate rents from the investment undertaken by upstream and downstream firms.

3.2 Quality Investments

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

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

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) and it is driven by each firm's bargaining power γ and $1-\gamma$.

Replacing the total royalty rate chosen by the patent holder, R^*, profits for the manufacturer and the downstream firm can be written as

$$\Pi_U(x_U, x_D, \alpha) =(1-\alpha)\gamma\frac{(1+\beta x_U + (1-\beta)x_D - c)^2}{8} - C(x_U), \tag{3}$$
$$\Pi_D(x_D, x_U, \alpha) =(1-\alpha)(1-\gamma)\frac{(1+\beta x_U + (1-\beta)x_D - c)^2}{8} - C(x_D). \tag{4}$$

Notice that the previous functions are not necessarily concave on 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 case and allows us to focus on situations where the investment levels are uniquely characterized using the first order conditions. It also rules out a corner solution where the product is supplied regardless of the value of θ.

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

The previous expressions allow us to characterize the equilibrium investment level of each firm.

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.*

An important feature of the quality investment of both firms is that they are strategic complements. That is, an increase in the investment of one of the firms increases 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 as a result — i.e. for lower realizations of θ. This effect, in turn, increases the profitability of the investment of the other firm. This result has interesting implications for the effect of α on the quality of the product. A higher α discourages investment by reducing the return from the innovation. The complementarity between x_D and x_U implies that this effect is exacerbated, as a lower investment of one of the firms indirectly reduces the incentives of the other firm to invest as well.

The previous complementarity between both investments also implies that changes in β and γ would have, in general, ambiguous effects on total investment. For example, a higher bargaining power by the upstream producer fosters investment by 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, in general, 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.[10]

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, 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 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

[10]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)).

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 obtained. The total effect of α results from the combination of these two sources of profits as

$$\Pi_L'(\alpha) = \frac{1}{4}(1+X^*-c)^2 + \frac{1+\alpha}{4}(1+X^*-c)\frac{dX^*}{d\alpha}. \tag{5}$$

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 shows that an increase in α, by making the hold-up problem more relevant, 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, however, patent holder profits are increasing and social welfare is decreasing 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 known. In that case, the quantity produced is efficient and all surplus

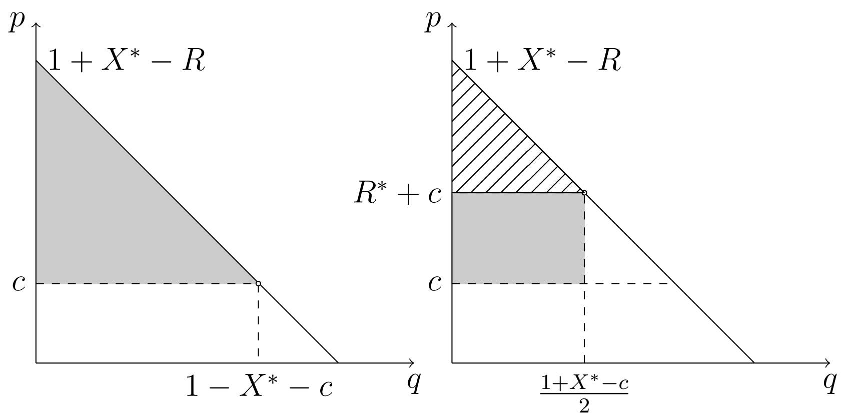

Figure 1: The grey area indicates the patent holder's profits when θ is observable (left) and when it is not (right) for a given value of X. The dashed area indicates the sum of the profits of the upstream and downstream firm.

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 invest which affects the value of the product in both states of the world.

The effect of α on the patent holder, 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 written 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 there are essentially three regions depending on the relevance of the investment by the upstream and downstream producer. In situations where this investment (and associated quality of the product) is not very sensitive to the scope for discrimination, 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 (both $W'(\alpha)$ and $\Pi_L'(\alpha)$ are 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 discrimination, 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 (both $W'(\alpha)$ and $\Pi_L'(\alpha)$ are 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 discrimination is not enough to overcome the loss in quality that it brings about in this case. 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 next example illustrates the previous forces by describing a extreme example where only the investment of the downstream producer matters.

Example 1 (Downstream Investment Only)**.** *Consider the case* $\beta = 0$ *and, in order for the bargaining power to be allocated efficiently, assume that* $\gamma = 0$. *Notice that this case also coincides with perfect competition upstream.*

As there will never be upstream investment, we equate the effect of α *on total investment to the effect on* x_D, $X = x_D$. *Using previous calculations we can characterize the* *effect of α 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 $C''(x_D^*) > \frac{1-\alpha}{4}$.

Using (5), we can show that the profits of the patent holder are decreasing in α when

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

Social welfare can be computed as

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

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,

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

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

3.3 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.

While in cases where θ is not known, licensing upstream and downstream has no impact on firm profits and incentives to innovate, this is not the case when θ is known. The possibility to license downstream — and maybe also upstream — implies that the patent holder can adjust the royalty rate to the value of the final product. As we have seen, this effect creates a trade-off. On the one hand, it exacerbates the hold-up problem, as it reduces the rents of the upstream and downstream firm from the sale of high quality products. On the other, it enables the efficient provision of lower value goods.

As discussed in the previous section, where 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 discrimination was prevented altogether. This can be achieved in the context of our model by imposing that licensing takes place at the upstream level. 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 that $\alpha = 0$.

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 profit without discrimination. This is more likely to arise when the investment is highly sensitive to price discrimination so that the profit of the patent holder decreases locally as α increases. In those circumstances, the patent holder wishes that he could commit not to discriminate and if a regulator imposes the absence of discrimination, the patent holder will not object. Still, if the choice is left to the patent holder, he may find it difficult to commit not to discriminate. 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, he would choose to license 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 maximises 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 have assumed that $\gamma = 0$ and $\beta = 0$. Furthermore, suppose that the cost function is quadratic $C(x) = \frac{k}{2}x^2$. In that case, the equilibrium innovation size 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 $\frac{\partial^2 x_D^*}{\partial k \partial \alpha} > 0$. As $\frac{\partial x_D^*}{\partial \alpha} < 0$ 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 licensor profits and social welfare increase when licensing takes place upstream and price discrimination is not possible. At the other extreme, when $k > \frac{5}{4}$, innovation responds very little to k and price discrimination increases both licensor profits and social welfare. In the intermediate region, when $k \in \left[\frac{3}{4}, \frac{5}{4}\right]$ then $W(\alpha)$ is strictly decreasing in α and $\Pi_L(\alpha)$ is maximized at $\alpha = 4k - 3 > 0$. There is thus a range of value of α for which that 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.

4 Robustness and Extensions

In this section we discuss how the results change under alternative assumptions.

4.1 Private Information on the Exogenous Component of the Product Value

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. 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

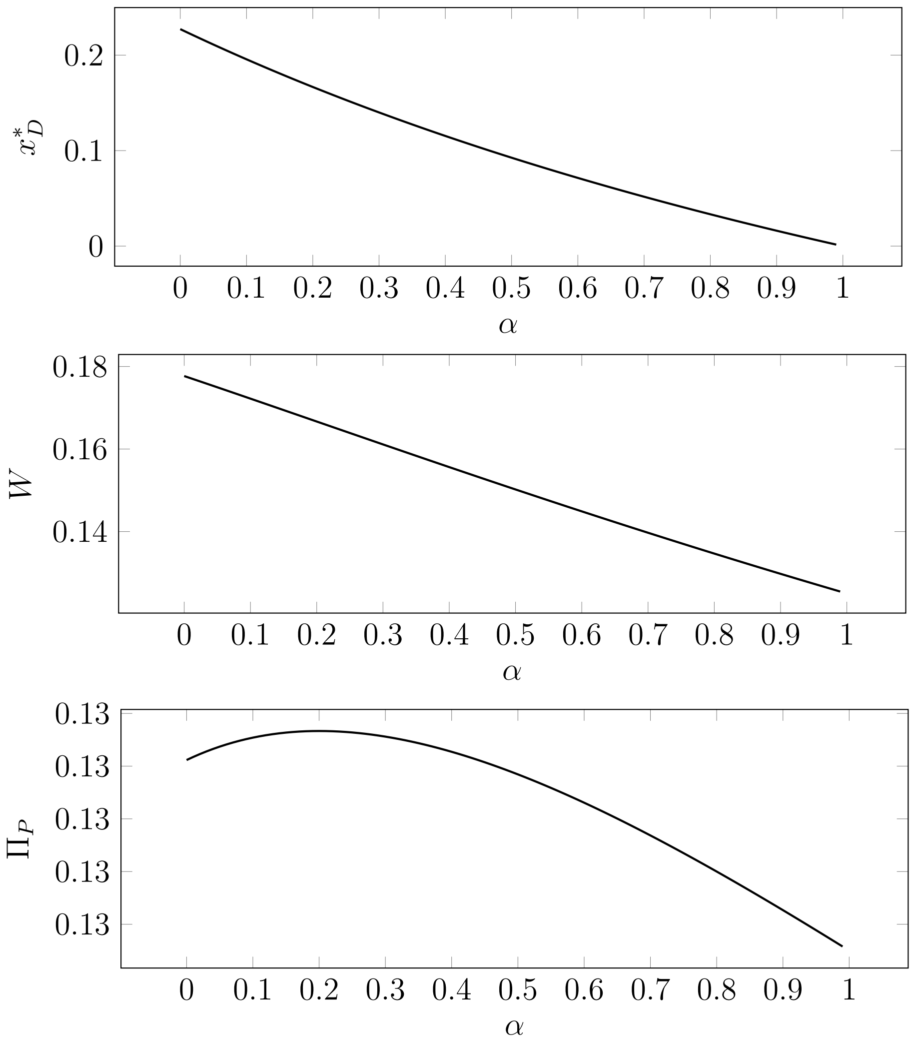

Figure 2: Equilibrium investment of the downstream firm, total welfare, and profits of the patent holder when $C(x) = \frac{k}{2}x^2$ and with parameter values $\gamma = \beta = 0$, $c = \frac{1}{2}$, and $k = \frac{4}{5}$. Under this parameterization, $x_U^* = 0$.

to the benchmark model, the only outcome that changes relates to the situation in which 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 internalised 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 $R^* = r_U^{0*} + r_D^{0*} = \frac{1+X-c}{2}$. As before, royalty neutrality holds in this environment.

In the case where θ is only observed by the downstream producer the previous royalty rate and price imply that the profits become

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

As expected, the equilibrium outcome exhibits, for a given X, a double-marginalization distortion that reduces total profits compared to the situation in (3) and (4).

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 firm undermines the overall incentives for firms to invest.

In addition, in cases where γ is sufficiently close to 1 when θ is observed the upstream supplier can extract almost all the surplus from the transaction, and provide no incentives for the downstream firm to invest. This is a kind of hold-up 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 actually increase total quality and social welfare.

4.2 Private Information on the Endogenous Quality of the Product

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 case where the parameter β is also private information and, therefore, the patent holder cannot anticipate with certainty the equilibrium value of X.

4.2.1 The Value of X is Private Information

When X is private information, in choosing the royalty rate, the patent holder has the belief that the quality level obtained is $\hat{X}$. Given those beliefs, 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 benchmark model, we can characterize the investment choice of the upstream and downstream producer, respectively, as a result of their profit maximization problem

$$\max_{x_U} (1-\alpha)\gamma \frac{(1+\beta x_U + (1-\beta)x_D - R^*(\hat{X}) - c)^2}{2} - C(x_U), \tag{6}$$

$$\max_{x_D} (1-\alpha)(1-\gamma) \frac{(1+\beta x_U + (1-\beta)x_D - R^*(\hat{X}) - c)^2}{2} - C(x_D). \tag{7}$$

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 3. *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.

4.2.2 The Value of β is Private Information

We now discuss the case where the parameter determining the relevance 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 $x_D^*(1) = x_U^*(0) = 0$.[11]

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

[11] 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.

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 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 4. *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 $\beta = 0$, $\tilde{x}_D(0) > \tilde{x}_U(1)$. 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 4.2.1 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 for 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 worse 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 the consumers buy only when the upstream innovation is relevant, $r^D(\theta) = \theta + \tilde{x}_D(0) - c$. The comparison of the two case 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.

4.3 Heterogeneous Cost of the Upstream Manufacturer

Suppose now that $\theta \leq 1$ is known to all firms but the cost of the upstream manufacturer is heterogeneous, with $c \sim U[0, 1]$. 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 c is observable, the patent holder chooses the total royalty rate as a result of

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

To simplify the exposition we study the case where $\beta = \gamma = 0$ so that $x_U^* = 0$. Then, the downstream firm solves

$$\max_{x_D} \, (1 - \alpha) \int_0^{\theta + x_D - R^*} (\theta + x_D - R^* - c)dc - C(x_D) = (1 - \alpha)\frac{(\theta + x_D)^2}{8} - C(x_D).$$

It is easy to observe that the results in this case are the mirror image to those discussed in Example 1 where we also focused on the case where only the downstream investment mattered. For this reason, the implications in this case would the opposite to those discussed before. When the relevant dimension of heterogeneity is the production cost of the upstream supplier, downstream licensing, to the extent that it cannot condition on the cost of the input, would increase the incentives to innovate. 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.

4.4 Ad-Valorem Royalties

In some industries, per-unit royalties — understood as a payment for each unit sold — are becoming the norm. Nevertheless, royalty rates that are a percentage of the price of the product — denoted as ad-valorem royalties — are still relevant. Here we show that they might undermine further the incentives to innovate.

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 due to two different channels. First, they decrease the distortions caused by double-marginalization. Second, they mitigate the royalty-stacking problem that arises when several patent holders have complementary patents that are necessary to sell a final product. The authors also show that ad-valorem royalties tend to benefit the patent holder and reduce the profits of the final-product manufacturer.

In this section we revisit the discussion between per-unit and ad-valorem royalties in a vertical chain where hold up might be a concern. We focus on the benchmark model where we assume that $\beta = \gamma = 0$ so that only the investment of the downstream firm matters. We also focus on the case $\alpha = 0$ so that there is no scope for choosing different royalty rates depending on θ.

Under the previous assumptions, the extreme kind of hold-up that we highlighted in the benchmark model would never arise under per-unit royalties. The patent holder would always choose a total royalty rate $R^* = \frac{1+x_D-c}{2}$. From Example 1 the equilibrium quality level arises in an interior solution as the result of

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

Under ad-valorem royalties we can compute the profits of the downstream firm as a result of production as

$$\pi_D(x_D) = (1-s_D)(\theta + x_D - c).$$

Notice that in this case, the patent holder would always maximize its own profits by using $s_D^* = 1$ (assuming that it only applies a royalty rate downstream). By doing so there is production whenever it is efficient, if $\theta + x_D - c \geq 0$, and all the returns from the investment accrue to the patent holder.

The comparison of both cases identifies a new trade-off between dead-weight loss and the incentives for firms to innovate. Under ad-valorem royalties production takes place whenever $\theta + x_D \geq c$. To the extent that $s_D^* = 1$ all profits accrue to the patent holder and no investment is carried out. In contrast, as observed throughout the paper, per-unit royalties do not attain the optimal level of innovation but they provide more incentives for the upstream supplier and the downstream producer to innovate. As a result, in those instances where the incentives to innovate of the downstream producer are particularly relevant, per unit royalties are likely to be socially preferable in spite of the higher distortion in the provision of the good that they engender.

5 Uncertain Validity of the Patent

The benchmark model assumes that the patent owned by the patent holder is 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.[12] For the potential

[12]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.

licensee to determine whether it is valid or not and whether it should pay a royalty, knowledge of the underlying technology is necessary. An upstream firm, being closer to the technology itself, is more likely to possess this knowledge. This is contrast with the downstream producer, who might buy a component that already embeds the technology and adapt it to its own needs.[13]

To make the previous notion operational, we now extend the benchmark model to consider the possibility that the patent is invalid. We assume that the validity status is known by 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 upstream or downstream. To isolate the effects of this case, we assume that $\alpha = 0$ so that without private information royalty neutrality applies and the patent holder would be indifferent between both options. Furthermore, 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

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

[13] 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.

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 in 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 a royalty rate $r_D < \tilde{r}_D$ where the latter is defined as

$$\int_{c+\tilde{r}_D - X}^{1} \pi_D(\theta, X, \tilde{r}_D)d\theta = \hat{\pi}_D. \tag{8}$$

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 licensing 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 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 5. *For a given X, licensing downstream 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 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 with probability ϕ. In other words, licensing downstream provides a more efficient way to extract surplus from production and this benefits the patent holder.

We can now move to the previous stage of the game and analyze the effects of upstream and downstream licensing on the incentives for firms to innovate. To do so, it is useful to separate the implications of private information in two parts. First, there is a *royalty-allocation effect*. For the same total revenue for the patent holder, a royalty rate paid even if the patent is invalid spreads out the burden over the two states of the world. Second, there is an *investment effect*. When the total revenue of the patent holder increases, the profits of the upstream and downstream firms are reduced and so are the incentives to invest.

While the second effect is immediate from Proposition 5, the royalty-allocation effect might be less obvious. The next result shows that, if patent holder profits were unchanged, upstream and downstream profits would increase under downstream licensing.

Proposition 6. *Consider a downstream royalty r_D so that the patent holder obtains the same revenue by licensing as in the case where it licenses upstream at a rate r_U^*. Under downstream licensing total profits increase and so do the investment incentives.*

The royalty rate set by the patent holder increases firm costs and, as a result, it generates a distortions since products with a low θ become unprofitable. This distortion is convex in the royalty rate. This means that, for the same revenue for the patent holder, it is always preferred to spread out the cost across 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 favours downstream licensing.

Hence, the patent holder will always prefer downstream licensing. In some circumstances, upstream licensing will however be socially optimal: as shown by Proposition 6, 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 for the downstream royalty that would keep the patent holder indifferent with his optimal upstream royalty. However, the patent holder will optimally charge a higher royalty downstream that the royalty that would keep him indifferent. This reduces investment and welfare. When the investment is very sensitive to the rents of upstream an downstream firms, the second effect dominates, so that upstream licensing is preferred from a social perspective. As a result, even though there are instances in which the choice of the patent holder is the same as what social welfare would dictate, there are also circumstances in which the incentives diverge. The patent holder will prefer downstream licensing when upstream licensing is socially optimal.

Figure 3 illustrates the previous discussion using the same configuration as in Example 1, where only downstream investment matters and there are no legal costs, $M = 0$. 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 licensor 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.[14] For intermediate values, and consistent with Proposition 5, the licensor 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 with the effect on social

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

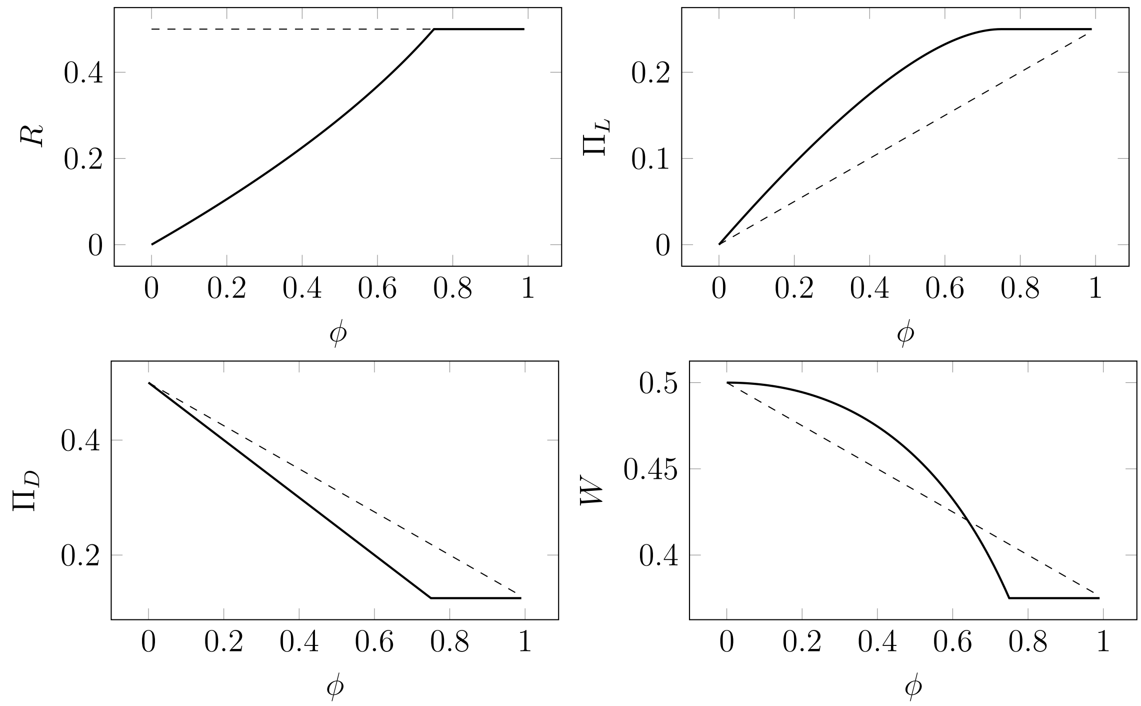

Figure 3: For a given X, equilibrium total royalty rate, licensor profits, downstream profits, and social welfare for different values of ϕ under downstream licensing (solid line) and upstream licensing (dashed line). The parameterization assumes $M = 0$, $X = c = 0.5$, and $\gamma = \beta = 0$. For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

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 $r_U^* = \frac{1+X-c}{2}$, creating a trade-off. Notice, however, that when the constraint is not binding and $r_D^* = \frac{1+X-c}{2}$, welfare is higher under upstream licensing as in this case, double marginalization only occurs with probability ϕ. This implies that when ϕ is sufficiently high (but less than one) there is no trade-off and upstream licensing is always superior.

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 profit 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.

More in general, the combination of the two effects allows us to draw conclusions similar to those in the benchmark model. In that case, the socially optimal regime implied 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 and reduce social welfare. Here, we have shown that asymmetric information provides a similar trade-off, such that downstream licensing, while reducing distortions, also undermines the incentive to invest.

6 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 its own technology. This limited information will negative 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

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.

Langus, Gregor, Vilen Lipatov and Damien Neven, "Standard-Essential Patents: Who Is Really Holding up (and When)?,," *Journal of Competition Law and Economics*, 2013, *9*(2), pp. 253–284.

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 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 both firms with respect to x_U^* and x_D^* 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

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

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

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

This derivative is negative if

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

This condition is satisfied whenever $\Pi_L'(\alpha) < 0$ which, using (5), occurs when

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

□

Proof of Proposition 3: The first order condition characterizing the optimal investment of the upstream innovator can be written as

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

where we have already imposed the equilibrium outcome that $\hat{X} = X^*$. 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 licensor is the same as when X is known, higher investment translates into a higher royalty rate. □

Proof of Proposition 4: 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 $\tilde{x}_D(0) \geq \tilde{x}_U(1) > x_D^* \geq x_U^*$. □

Proof of Proposition 5: As profits for the patent holder when licensing downstream increase in M it is enough to show the result for $M = 0$.

Using (8), 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

$$\Pi_L^D = \phi(1-\gamma)\frac{(1+X-c)^2}{2},$$

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 6: 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 $r_D < \phi r_U$. □