UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche "Marco Fanno"

SEQUENTIAL INNOVATIONS WITH UNOBSERVABLE FOLLOW-ON INVESTMENTS

STEFANO COMINO
Università di Trento

FABIO M. MANENTI
Università di Padova

ANTONIO NICOLO'
Università di Padova

June 2007

"MARCO FANNO" WORKING PAPER N.41

Sequential innovations with unobservable follow-on investments*

Stefano Comino† Fabio, M. Manenti‡ Antonio Nicolò §

February 7, 2008

Abstract

We consider a cumulative innovation process in which a follow-on innovator invests in R&D activities that influence both the expected commercial value of the innovation as well as the probability of infringing on the patent of an earlier inventor. We show that, when the second innovator investments are not observable, *ex-ante* licensing agreements are ineffective and the follow-on innovator fails to invest efficiently. Because of this inefficiency, a large patent breadth may be harmful for the first innovator too, and therefore it may be Pareto-dominated. This occurs when a large patent breadth exacerbates the overinvestment or the underinvestment of the follow-on innovator.

J.E.L. codes: K3, L5, O3.

Keywords: sequential innovation, patents, licensing, intellectual property.

*Paper presented at the Diparmimento di Ingegneria Gestionale - Politecnico di Milano, at the 34^{th} EARIE Conference, Valencia 2007, and at the 2^{nd} EPIP Conference, Lund 2007.

†Corresponding author: Dipartimento di Economia, Università di Trento, Via Inama 5, 38100 Trento (Italy). E-mail stefano.comino@economia.unitn.it, tel. +390461882221, fax +390461882222.

‡Dipartimento di Scienze Economiche "M. Fanno", Università di Padova, Via del Santo 33, 35123 Padova (Italy).

§Dipartimento di Scienze Economiche "M. Fanno", Università di Padova, Via del Santo 33, 35123 Padova (Italy).

1 Introduction

In several industries, technical advance does not fit the stylized representation of stand-alone inventions traditionally portrayed by Nordhaus (1969).[1] In semiconductors, biotechnology, aircraft, or computer software technical advance is cumulative and subsequent generations of innovators build on and interact with technologies provided by earlier inventors. In these instances, follow-on innovators "stand on the shoulders of giants" that laid down the foundations of the industry (Scotchmer, 1991).

When innovation is cumulative and it is carried out by subsequent innovators, the patent system has to balance two, potentially conflicting, goals: ensure sufficient rewards to the early innovators, without, at the same time, discouraging follow-on R&D efforts. The contribution to the social welfare of early discoveries is broader than in case of industries with stand-alone inventions. They are valuable not only per se but also because they enable or facilitate valuable derived inventions. This externality calls for broad intellectual property rights to protect early discoveries: in order to align private and social incentives to R&D, early innovators should obtain a significant stake in the revenues generated by the innovations to which they contribute. This can be accomplished by granting the original patent a broad scope so that infringing subsequent innovators will need to negotiate the permission (license) of the patent-holder in order to commercialize their discoveries.

However, rewarding early innovators by means of a strong patent protection might undermine future R&D. Anticipating that early innovators are warranted significant claims on derived inventions, follow-on innovators may have sub-optimal incentives to perform R&D activities. This hold-up problem arises especially when inventors sunk specific investments before negotiating the terms of the licensing agreement with the patent-holder.

Since the surplus on which parties negotiate is represented by the commercial value of the derived innovation and it does not take into account the costs that have already been

[1]According to Merges and Nelson (1990) it is worth distinguishing at least four different industrial patterns of technical advance: discrete (stand-alone) invention model, cumulative technologies, chemical technologies and "science-based" technologies.

sunk (Lemley and Shapiro, 2007; Scotchmer, 1991; Shapiro, 2001),[2] the follow-on inventor is in a weak bargaining position; in this way, the follow-on innovator might not be able to obtain a sufficient return on the R&D investment.[3]

As already explicitly suggested in Scotchmer (1991), parties can mitigate the hold-up problem by employing ex-ante licensing contracts (or prior agreements), that is by negotiating the licensing agreement before the follow-on innovator has sunk the R&D costs. In case of an ex-ante agreement, the surplus over which parties bargain is represented by the commercial value of the derived invention net of the R&D expenditures; in this way, the costs borne by the follow-on innovator are taken into account in the bargaining process.

In a seminal paper, Green and Scotchmer (1995) formalize this idea and show that, assuming that parties negotiate in a context of symmetric information, the feasibility of ex-ante licensing ensures that the follow-on R&D investment always occurs efficiently. In this scenario, the only task of the patent policy is to ensure enough rewards to the early inventor, and this is accomplished by granting her/him a patent with a very large, if not infinite, breadth.

The assumption that ex-ante contracting under symmetric information is feasible has been repeatedly employed in the subsequent theoretical contributions on cumulative innovation (see O'Donoghue et al., 1998 , Scotchmer, 1996 , and Schankerman and Scotchmer, 2002).[4] In two recent papers, Bessen (2004) and Bessen and Maskin (2007) consider the

---

[2]According to Shapiro (2001) the hold-up problem represents a real threat to future innovation in several industries. This problem is exacerbated by the lengthy approval process of Patent Offices with the danger that new products infringe on patents issued after these products were designed. The concern about these so-called "submarine patents" is particularly relevant in the software industry, see Graham and Mowery (2004).

[3]The risk of of slowing down the pace of future innovations is compounded in case of "patent thickets", that is, when several patents read, at the same time, on a given product or technology. Patent thickets are common the IT sector (Lemley and Shapiro, 2007; Siebert and von Graevenitz, 2006). See also Heller and Eisenberg (1998) for a discussion of the consequences of patent thickets in the context of biomedical research.

[4]Matutes et al. (1996) and Chang (1995) do not consider the possibility of ex-ante licenses.

case where the development costs of the improvement are private information of the follow-on innovator. The authors show that ex-ante licensing does not guarantee that all efficient follow-on innovations occur: at the equilibrium, in some cases the second innovator fails to invest efficiently.[5]

In this paper, we present a model based on Green and Scotchmer (1995) and we obtain results much sharper than those in Bessen (2004), and Bessen and Maskin (2007). Under the realistic assumption that when contracting over the licensing terms the early innovator cannot observe whether the follow-on inventor has already undertaken the R&D activity, ex-ante agreements are ineffective. The follow-on inventor benefits from collecting more precise information about its innovation before entering the licensing negotiations with the patent holder and this fact always prevents efficient licensing. In the paper we show that, at the equilibrium, not only undervinvestment but also overinvesment may occur. The level of R&D activity of the follow-on innovator has both a commercial effect, i.e. it increases the expected commercial value of the innovation, and an infringement effect, i.e. it reduces the probability of infringing the first innovator's patent; when the infringement effect prevails, the follow-on inventor may be induced to invest more than the efficient level.

As a consequence of this inefficiency in the follow-on level of investment, the profits of the patent-holder do not necessarily increase with the strength of patent protection. An increase in the breadth of the patent improves the probability of infringement but also alters the incentives to invest of the follow-on innovator. This additional effect reduces the profits

See Gallini and Scotchmer (2002) for a recent review on these issues.

[5]Siebert and von Graevenitz (2006) formalize the choice of ex-post *vs* ex-ante licensing considering the case of n firms simultaneously involved in developing a common technology. In case of ex-post licensing, firms enter in a patent race: by augmenting its number of patents, a firm strengthens its bargaining position during the ensuing licensing negotiations. With ex-ante licensing, defined as agreements "to share future research results prior to R&D investments", firms avoid the event of a patent race. The authors show that the choice between reaching an agreement ex-ante or ex-post depends on the strength of the patent portfolios that firms already have in stock, and on the nature of competition in the product market.

of the patent-holder whenever it exacerbates the inefficiency in the R&D level of the second innovator.

In the paper, we provide two cases where the overall effect of a larger breadth implies less profits for the patent-holder. We show that this can occur because the larger patent breadth exacerbates either the underinvestment or the overinvestment of the follow-on innovator.

These last results are interesting and, in fact, they show that a large patent breadth might be Pareto-dominated: a too strong patent protection of the initial innovator, may harm not only the follow-on inventor, but also the patent-holder. This fact is in clear contrast with the previous literature. Both in Green and Scotchmer (1995) and Bessen (2004) the optimal patent policy has to balance opposing interests: a larger breadth benefits the early innovator to the detriment of the follow-on firm. To the contrary, we show that the interests of the two firms may not necessarily diverge in terms of patent breadth.

Our paper contributes to the current debate about the optimal scope of patents in industries where innovation is cumulative. As Gallini and Scotchmer (2002) put it, several arguments in favor of either weak or strong standards for IPR (intellectual property rights) have been proposed, and the existing literature is inconclusive as to whether broad or narrow patents are better suited to encourage innovations. However "one lesson is clear: the optimal design of IP depends importantly on the ease with which rights holders can contract around conflicts in rights" (Gallini and Scotchmer, 2002 p. 67). In this paper we show that, under reasonable conditions, the possibility to enter into ex-ante agreements fails to ensure efficient follow-on investment. Very broad patents may result in serious underinvestment or overinvestment that goes to the detriment of all the industry participants.

The paper is organized as follows: in Section 2, we outline the model, and in Section 3 we derive the main results of our analysis. Section 4 discusses some of the main hypotheses at the base of our results and provides some possible policy implications.

2 The Model

We consider a cumulative innovation process in which once the first inventor, firm 1, patents its innovation, a second inventor, firm 2, gets an "idea" for an improvement. Firm 2 can get the idea at any point in time after firm 1 has patented its invention. As we describe below, firm 2 undertakes some R&D activity in order to develop its idea into a commercially valuable innovation that might infringe the patent of the first innovator or not; in case of infringement, the second inventor has to obtain a license from firm 1 in order to commercialize its idea.

In what follows, we restrict the analysis to the case in which the overall commercial value of the two innovations resides in the follow-on invention; that is, the early innovation is a research tool that has no commercial value per se.[6] Moreover, we focus on the second inventor's behavior being the first innovation already in place and protected by a patent.

The idea of the second inventor may be more or less promising both in terms of the commercial benefits that it can generate and in terms of the probability of infringing on firm 1's patent. Formally we represent an idea as $\left\{p\left(r\right), c\left(r\right), V^G, V^B, \gamma(b), \beta(b)\right\}$ whose terms are described below. In order to develop the idea, firm 2 undertakes a certain amount of R&D activity, $r \geq 0$ incurring a cost $c(r)$; once the R&D cost has been sunk, with probability $p(r)$ a "good state" of the world occurs, and with probability $1 - p(r)$ a "bad state" of the world occurs. In the analysis we assume that $c'(r) > 0$, $c''(r) > 0$, $p'(r) > 0$, $p''(r) \leq 0$.

In the good state of the world, the innovation, i) has a commercial value V^G, and ii) it does not infringe on the patent of the first innovation with probability $\gamma\left(b\right)$, where $b \in \Re_+$ represents the patent breadth set by Government regulations. In the bad state, the commercaial value is $V^B \leq V^G$, and $\beta\left(b\right)$ is the probability that the follow-on innovation does not infringe the patent of firm 1, with $\beta\left(b\right) \leq \gamma(b)$, for any b. In other words, in the good state the innovation has a larger commercial value and generally a lower probability of infringing upon the first innovator's patent.

[6] Obviously, our arguments apply also to the case in which both innovations have a positive commercial value but the two innovators serve unrelated markets.

It is natural to assume that the probabilities of not infringing decrease with the patent breadth: $\gamma'(b) \leq 0$, and $\beta'(b) \leq 0$.

From this setting it follows that: i) there is a positive relationship between the expected commercial value and the probability of not infringing of the second innovation, and ii) a larger r, i.e. a larger R&D activity, increases both the expected commercial value, $p(r)V^G + (1-p(r))\,V^B$, as well as the probability that the second innovation does not infringe on the patent of the first inventor, $p(r)\gamma(b) + (1-p(r))\,\beta(b)$. All through the paper we assume that an interior solution does exists, that is that the selected amount of firm's 2 R&D r is such that $0 < p(r) < 1$.

2.1 Timing and Information Structure

The timing of the game is as follows:

1. firm 2 gets an idea;

2. firm 2 chooses how much R&D activities to undertake, r, and, afterwards, it observes the realized commercial value V^i, $i = G, B$ and whether its innovation infringes upon the first innovator's patent or not;

3. in case of infringement, firm 2 needs to be licensed by the first innovator. Firm 2 has two options as to when to negotiate the licensing terms: i) it may ask for the licensing agreement before having undertaken any R&D activity; or ii) it may choose r first, and then asking for the licensing agreement after having observed both commercial value, V^i, $i = G, B$, and whether the invention infringes or not. In the former case, we say that the second inventor is looking for an *ex-ante* licensing agreement while in the latter we say that it is looking for an *ex-post* licensing agreement. We assume that in case firms fail to reach an agreement then, in case of infringement, the terms of licensing are set by the Court. We assume that the Court observes V^i and mandates a licensing fee $\bar{L}^i$, $i = G, B$; for the sake of simplcity we assume that $\bar{L}^i \leq V^i$, $i = G, B$.

A crucial aspects of the paper is about the informational structure of the game. In what follows, we assume that the R&D activity is neither verifiable nor observable by the first inventor; in particular, the non-observability of the R&D activity implies that, when contracting over the licensing terms, firm 1 ignores whether firm 2 has already sunk $c(r)$ or not. Moreover, we assume that both the commercial value V^i, $i = G, B$ and the infringement of the patent are verifiable, but only once the second innovation is brought to the market; namely, the second innovator holds this information privately till the moment it markets its invention.

3 Results

We start by defining the equilibrium in the licensing agreement.

3.1 Licensing agreement

We assume that the bargaining during the negotiations of the licensing agreement are as follows:

- firm 2 asks for a licensing agreement and shows the idea $\left\{p\left(r\right), c\left(r\right), V^G, V^B, \gamma(b), \beta(b)\right\}$;

- firm 1 makes a take-it-or-leave.it proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$: namely a payment $\tilde{L}^i$ contingent on infringement and occurence of outcome $i = G, B$;

- if firm 2 rejects the proposal, then the fees are those imposed by the Court.

The following proposition highlights the characteristics of the equilibrium of the licensing negotiations.

Proposition 1. *In equilibrium, all licensing occurs at the fees implemented by the Court.*

Proof. As first consider the choice of firm 2 once it has asked a licensing agreement, and firm 1 has offered a contract $\left(\tilde{L}^G, \tilde{L}^B\right)$, namely a licensing contract that specifies a payment $\tilde{L}^i$ contingent on infringement and on the occurence of V^i, with $i = G, B$.

If firm 2 is ex-post and has observed V^i it accepts the proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ if and only if $\tilde{L}^i \leq \bar{L}^i$. In case firm 2 is ex-ante it will: *i)* accept any proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i \leq \bar{L}^i$

$i = G, B$; *ii)* reject any proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i \geq \bar{L}^i$, $\tilde{L}^j > \bar{L}^j$, with $i = G, B$, $j = G, B$, and $i \neq j$; *iii)* either accept or reject a proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i < \bar{L}^i$, $\tilde{L}^j > \bar{L}^j$, with $i = G, B$, $j = G, B$, and $i \neq j$.

Consider equilibria where firm 2 asks for a licensing agreement only ex-post

Consider firm 1. Given that firm 2 is ex-post, then offering any $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^G < \tilde{L}^G$ and/or $\tilde{L}^B < \tilde{L}^B$ is dominated by $\left(\tilde{L}^G, \tilde{L}^B\right) = \left(\bar{L}^G, \bar{L}^B\right)$; moreover, any proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i \geq \bar{L}^i$, $i = G, B$ yields to firm 1 the same pay-off as $\left(\bar{L}^G, \bar{L}^B\right)$, since it obtains $\bar{L}^i$, $i = G, B$. Therefore it is optimal to offer $\left(\tilde{L}^G, \tilde{L}^B\right) = \left(\bar{L}^G, \bar{L}^B\right)$. Given the proposal of firm 1, firm 2 is indifferent between going ex-ante or ex-post and between accepting or rejecting $\left(\tilde{L}^G, \tilde{L}^B\right) = \left(\bar{L}^G, \bar{L}^B\right)$; thus going ex-post and accepting the proposal is optimal. Therefore we have defined an equilibrium where: firm 2 asks for a licensing agreement only ex-post and, after observing V^i, it accepts the proposal of firm 1 if and only if $\tilde{L}^i \leq \bar{L}^i$, with $i = G, B$; firm 1 offers $\left(\tilde{L}^G, \tilde{L}^B\right) = \left(\bar{L}^G, \bar{L}^B\right)$.

Consider equilibria where firm 2 asks for a licensing agreement only ex-ante

Suppose that, with probability 1 firm 2 chooses to ask for a licensing agreement ex-ante. The contract $\left(\tilde{L}^G, \tilde{L}^B\right)$ that firm 1 proposes solves the following maximization problem:

$$\max_{\tilde{L}^G, \tilde{L}^B} p(\tilde{r})\, \tilde{L}^G (1 - \gamma(b)) + (1 - p(\tilde{r}))\, \tilde{L}^B (1 - \beta(b))$$

subject to

$$p(\tilde{r}) \left(V^G - \tilde{L}^G (1 - \gamma(b))\right) + (1 - p(\tilde{r})) \left(V^B - \tilde{L}^B (1 - \beta(b))\right) - c(\tilde{r}) \geq$$
$$p(\bar{r}) \left(V^G - \bar{L}^G (1 - \gamma(b))\right) + (1 - p(\bar{r})) \left(V^B - \bar{L}^B (1 - \beta(b))\right) - c(\bar{r}),$$

where $\tilde{r}$ is the investment level that firm 2 chooses once it has accepted the proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$, while $\bar{r}$ is the investment level it chooses after rejecting the proposal, given that in case of infringement it will pay the licensing fees imposed by the Court $\left(\bar{L}^G, \bar{L}^B\right)$.

After some manipulations it is possible to show that the following are the Kuhn-Tucker condition of the maximization problem:

$$p(\tilde{r})(1 - \gamma(b)) + p'(\tilde{r}) \tfrac{\partial \tilde{r}}{\partial \tilde{L}^G} \left[\tilde{L}^G (1 - \gamma(b)) - \tilde{L}^B (1 - \beta(b))\right] = \lambda p(\tilde{r})(1 - \gamma(b))$$
$$(1 - p(\tilde{r}))(1 - \beta(b)) + p'(\tilde{r}) \tfrac{\partial \tilde{r}}{\partial \tilde{L}^B} \left[\tilde{L}^G (1 - \gamma(b)) - \tilde{L}^B (1 - \beta(b))\right] = \lambda (1 - p(\tilde{r}))(1 - \beta(b))$$

$\lambda \geq 0$, $\lambda \frac{\partial \pounds}{\partial \lambda} = 0$

This system is satisfied for $\lambda = 1$ and $\left(\tilde{L}^{G*}, \tilde{L}^{B*}\right)$ satifying the following two conditions: $\tilde{L}^{G*}(1-\gamma(b)) = \tilde{L}^{B*}(1-\beta(b))$, and the participation constraint is binding (i.e. firm 2 is indifferent between accepting or rejecting the offer); the first condition induces firm 2 to invest efficiently, while the second condition just implies that firm 1 extract all the surplus of the relationship.

As first note that if $\left(\bar{L}^{B}, \bar{L}^{G}\right)$ is such that $\bar{L}^{B}(1-\beta(b)) = \bar{L}^{G}(1-\gamma(b))$, then the following is an equilibrium: firm 2 asks for an agreement ex-ante; firm 1 proposes $\left(\bar{L}^{B}, \bar{L}^{G}\right)$; firm 2 accepts the proposal and invests $\bar{r}$.

Consider now the case where $\bar{L}^{B}(1-\beta(b)) \neq \bar{L}^{G}(1-\gamma(b))$; in what follows we show that in this case there cannot be an equilibrium where firm 2 asks for a licensing agreement ex-ante with probability 1. Note first that when $\bar{L}^{B}(1-\beta(b)) \neq \bar{L}^{G}(1-\gamma(b))$ it cannot be that either $\tilde{L}^{G*} = \bar{L}^{G}$ and/or $\tilde{L}^{G*} = \bar{L}^{B}$. To the contrary suppose $\tilde{L}^{i*} = \bar{L}^{i}$, then: *i)* if $\tilde{L}^{j*} = \bar{L}^{j}$, then condition $\tilde{L}^{G*}(1-\gamma(b)) = \tilde{L}^{B*}(1-\beta(b))$ is not satisfied; *ii)* if $\tilde{L}^{j*} < \bar{L}^{j}$, then the participation constraint is not binding and firm 1 is not maximizing its profits; *iii)* if $\tilde{L}^{j*} > \bar{L}^{j}$, then the participation constraint is violated and firm 2 rejects the proposal (where $i = G, B$, $j = G, B$ and $i \neq j$). Similarly, it cannot be $\left(\tilde{L}^{G*}, \tilde{L}^{B*}\right)$ such that: $\tilde{L}^{i*} < \bar{L}^{i}$, with $i = G, B$ since the participation constraint would not be binding; nor it cannot be $\tilde{L}^{i*} > \bar{L}^{i}$, with $i = G, B$, since firm 2 would reject such proposal. Therefore, the contract $\left(\tilde{L}^{G*}, \tilde{L}^{B*}\right)$ that firm 1 proposes satisfies the conditions $\tilde{L}^{i*} < \bar{L}^{i}$, and $\tilde{L}^{j*} > \bar{L}^{j}$, with $i = G, B$, $j = G, B$ and $i \neq j$. However, given the proposal of firm 1, it is not optimal for firm 2 to ask the licensing agreement ex-ante. By making the investment first, firm 2 reduces the licensing fees it pays: after observing V^{i}, it can select the minimum licensing fee between the one proposed by firm 1 and the one implemented by the Court.

Consider equilibria where firm 2 asks for a licensing agreement ex-ante with probability $\alpha \in (0,1)$ and ex-post with probability $(1-\alpha)$

In order to have firm 2 to be willing to randomize between asking the licensing agreement ex-ante or ex-post, firm 1's proposal has to be such that $\tilde{L}^{i} \leq \bar{L}^{i}$ $i = G, B$. However, in what follows we show that it is never optimal for firm 1 to offer a contract $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i \leq \bar{L}^i$, and $\tilde{L}^j < \bar{L}^j$, with $i = G, B$, $j = G, B$ and $i \neq j$. This fact implies that two things might occur: *i)* either in equibrium firm 1 finds it optimal to offer $\left(\tilde{L}^G, \tilde{L}^B\right) = \left(\bar{L}^B, \bar{L}^G\right)$ (from the arguments above, this occurs for instance when $\left(\bar{L}^B, \bar{L}^G\right)$ are such that $\bar{L}^B\left(1-\beta\left(b\right)\right) = \bar{L}^G\left(1-\gamma\left(b\right)\right)$); *ii)* there cannot be an equilibrium where firm 2 randomizes between asking for a licensing agreement ex-ante or ex-post.

Denote with $\alpha \in (0,1)$ $(1-\alpha$, respectively) the probability that firm 2 has asked the licensing agreement ex-ante (ex-post). Suppose that firm 1 proposes $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i \leq \bar{L}^i$, and $\tilde{L}^j < \bar{L}^j$, with $i = G, B$, $j = G, B$ and $i \neq j$. Note that such proposal is accepted by firm 2 both in case it is ex-ante and also in case it is ex-post. The expected licensing revenues of firm 1 are:

$$\alpha\left(p\left(\tilde{r}\right)\left(1-\gamma\left(b\right)\right)\tilde{L}^G + \left(1-p\left(\tilde{r}\right)\right)\tilde{L}^B\right) +$$
$$(1-\alpha)\left(prob(good)\min\left\{\tilde{L}^G, \bar{L}^G\right\} + \left(1-prob(good)\right)\min\left\{\tilde{L}^B, \bar{L}^B\right\}\right),$$

where, $\tilde{r}$ is the investment level that an ex-ante firm 2 that accepts $\left(\tilde{L}^G, \tilde{L}^B\right)$ will choose; $prob\,(good)$ $\left(\left(1-prob(good)\right)\right.$, respectively) is the probability that firm 2 has asked an agreement ex-post after observing that there is infringement and that V^G (V^B respectively) has occurred.

a) Consider a proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^B = \bar{L}^B$ and $\tilde{L}^G < \bar{L}^G$. If firm 1 were sure that firm 2 is ex-post then it would prefer proposing a contract $\left(\tilde{L}^{G\prime}, \tilde{L}^{B\prime}\right)$ with $\tilde{L}^{G\prime} > \tilde{L}^G$ and $\tilde{L}^{B\prime} \geq \tilde{L}^B$. Therefore, to prove that proposing $\tilde{L}^B = \bar{L}^B$ and $\tilde{L}^G < \bar{L}^G$ cannot be optimal for firm 1 it is sufficient to check that it prefers to increase either $\tilde{L}^B$ and/or $\tilde{L}^G$ when facing an ex-ante type. Before entering in the analysis, note that firm 2 that accepts the contract $\left(\tilde{L}^G, \tilde{L}^B\right)$ before investing then will solve:

$$\max_r \quad p(r)\left[\gamma\left(b\right)V^G + \left(1-\gamma\left(b\right)\right)\left(V^G - \tilde{L}^G\right)\right] +$$
$$\left(1-p(r)\right)\left[\beta\left(b\right)V^B + \left(1-\beta\left(b\right)\right)\left(V^B - \tilde{L}^B\right)\right] - c\left(r\right).$$

Taking the derivative with respect to r, it is easy to show that the amount of R&D activity chosen by firm 2 satisfies the following condition:

$$p'(\tilde{r})\left[V^G - V^B\right] - p'(\tilde{r})\left[(1-\gamma(b))\tilde{L}^G - (1-\beta(b))L^B\right] = c'\left(\tilde{L}^G\right) \quad (1)$$

- – Let's differentiate implicitly the first order condition (1) with respect to $\tilde{L}^G$ and with repect to $\tilde{L}^B$: $\frac{\partial \tilde{r}}{\partial \tilde{L}^G} = \frac{p'(\tilde{r})(1-\gamma(b))}{p''(\tilde{r})\left[V^G - V^B - (1-\gamma(b))\tilde{L}^G + (1-\beta(b))\tilde{L}^B\right] - c''(\tilde{r})}$, and $\frac{\partial \tilde{r}}{\partial \tilde{L}^B} = \frac{-p'(\tilde{r})(1-\beta(b))}{p''(\tilde{r})\left[V^G - V^B - (1-\gamma(b))\tilde{L}^G + (1-\beta(b))\tilde{L}^B\right] - c''(\tilde{r})}$. The second order condition implies that the first implicit derivative is negative while the second one is positive. If $(1-\gamma(b))\tilde{L}^G \geq (1-\beta(b))\tilde{L}^B$, firm 1 is better offering $\left(\tilde{L}^G, \tilde{L}^B + \varepsilon\right)$, so that: the ex-ante type still accepts the proposal (at the initial contract the participation constraint is not binding), increases $\tilde{r}$,so that firm 1 increases the probability of obtaining $(1-\gamma(b))\tilde{L}^G$ rather than $(1-\beta(b))\tilde{L}^B$, and in case of the bad state of the world it obtains $(1-\beta(b))\left(\tilde{L}^B + \varepsilon\right)$, rather than just $(1-\beta(b))\tilde{L}^B$. If $(1-\gamma(b))\tilde{L}^G < (1-\beta(b))\tilde{L}^B$, firm 1 is better-off by offering $\left(\tilde{L}^G + \varepsilon, \tilde{L}^B\right)$, so that: the ex-ante type still accepts the offer, decreases $\tilde{r}$ and firm 1 increases the probability of obtaining $(1-\gamma(b))\tilde{L}^B$ rather than $(1-\beta(b))\tilde{L}^G$, and in case of the good state of the world it obtains $(1-\gamma(b))\left(\tilde{L}^G + \varepsilon\right)$ rather than just $(1-\gamma(b))\tilde{L}^G$. The previous arguments imply that any proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^B = \bar{L}^B$ and $\tilde{L}^G < \bar{L}^G$ is not optimal for firm 1 when it faces an ex-ante type: it prefers to increase the fees it asks in at least one of the two states of the world.

b) Consider a proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^G = \bar{L}^G$ and $\tilde{L}^B < \bar{L}^B$. If firm 1 were sure that firm 2 is ex-post, then it would prefer proposing a contract $\left(\tilde{L}^{G\prime}, \tilde{L}^{B\prime}\right)$ such that $\tilde{L}^{G\prime} \geq \tilde{L}^G$ and $\tilde{L}^{B\prime} > \tilde{L}^B$. Therefore, to prove that $\tilde{L}^G = \bar{L}^G$ and $\tilde{L}^B < \bar{L}^B$cannot an optimal it is sufficient to check that firm 1 prefers to increase either $\tilde{L}^B$ and/or $\tilde{L}^G$ when facing an ex-ante type.

   – as before we know that $\frac{\partial r}{\partial \tilde{L}^G} < 0$ and $\frac{\partial r}{\partial \tilde{L}^B} > 0$. Therefore, if $(1-\gamma(b))\tilde{L}^G \geq (1-\beta(b))\tilde{L}^B$, firm 1 is better offering $\left(\tilde{L}^G, \tilde{L}^B + \varepsilon\right)$, so that: the ex-ante type still accepts the offer (at the initial contract the participation constraint is not binding), increases $\tilde{r}$ and firm 1 increases the probability of obtaining $(1-\gamma(b))\,\tilde{L}^G$ rather than $(1-\beta(b))\,\tilde{L}^B$, and in case of the bad state of the world it obtains $(1-\beta(b))\left(\tilde{L}^B+\varepsilon\right)$. In case, $(1-\gamma(b))\,\tilde{L}^G < (1-\beta(b))\,\tilde{L}^B$, firm 1 is better offering $\left(\tilde{L}^G+\varepsilon,\tilde{L}^B\right)$, so that: the ex-ante type still accepts the offer, it decreases $\tilde{r}$ and firm 1 increases the probability of obtaining $(1-\gamma(b))\,\tilde{L}^B$ rather than $(1-\beta(b))\,\tilde{L}^G$, and in case of the good state of the world it obtains $(1-\gamma(b))\left(\tilde{L}^G+\varepsilon\right)$. The previous arguments imply that any proposal $\left(\tilde{L}^G,\tilde{L}^B\right)$ such that $\tilde{L}^G=\bar{L}^G$ and $\tilde{L}^B<\bar{L}^B$ is not optimal for firm 1 when it faces an ex-ante type: it prefers to increase the fees it asks in at least one of the two states of the world.

c) Similar argument apply for the case $\tilde{L}^B<\bar{L}^B$ and $\tilde{L}^G<\bar{L}^G$ and are omitted for the sake of brevity.

∎

The intuition of the result is as follows; in case firm 1 is willing to propose a contract different from the one impemented by the Court, namely some $\left(\tilde{L}^G,\tilde{L}^B\right)\neq\left(\bar{L}^G,\bar{L}^B\right)$, firm 2 approaches the first inventor ex-post in order to select the most favorable licensing fee: after observing the outcome, in case of infringement, it can choose whether to pay the fee proposed by firm 1 or that implemented by the Court; however, in case firm 2 is ex-post the first inventor has no incentive to offer any $\left(\tilde{L}^G,\tilde{L}^B\right)\neq\left(\bar{L}^G,\bar{L}^B\right)$.

The implication of the result is important. The licensing contract cannot improve upon the fees implemented by the Court and this implies that it cannot be used to ameliorate the R&D incentives of the second innovator.

3.2 R&D investment and licensing revenues

We want to investigate the optimal R&D choice of firm 2 and how this decision is affected by the patent breadth. From Proposition 1 we know that licensing occurs at the fees implemented by the Court, namely $\left(\bar{L}^G, \bar{L}^B\right)$. In what follows, in order to simplify the analysis we we will assume that the Court sets the fees proportional to the commercial value of the innovation: $\bar{L}^i = (1-\rho)V^i$. The parameter ρ determines how the value of the innovation is shared across inventors and it might be related both to the bargaining power of the two firms, and to the extent to which Government and Courts favor a more or less pronounced "pro-patent" environment.

Before determining the optimal amount of R&D activities performed by firm 2, it is useful to define the efficient level of R&D, that is the value of r that maximizes the joint profits of the two firms; formally, the efficient level of R&D, r^{eff}, is given by:

$$r^{eff} = argmax_r \left\{ p(r)V^G + (1-p(r))\, V^B - c\,(r) \right\}.$$

The first order condition is simply:

$$p'(r^{eff}) \left[V^G - V^B\right] = c' \left(r^{eff}\right). \tag{2}$$

Let us now consider the investment level that the follow-on innovator actually chooses. Whenever its invention does not infringe the patent of the early innovator, firm 2 benefits of the full expected commercial value it generates; this occurs with probability $\gamma(b)$ in the good state and with probability $\beta(b)$ in the bad state. In case of infringement firm 2 gets only a share ρ of the commercial value. Therefore, firm 2 chooses r in order to maximize the following expression:

$$\max_r \quad p(r) \left[\gamma\,(b)\, V^G + (1-\gamma\,(b))\, \rho V^G\right] + (1-p(r)) \left[\beta\,(b)\, V^B + (1-\beta\,(b))\, \rho V^B\right] - c\,(r)\,.$$

Taking the derivative with respect to r, it is easy to show that the amount of R&D activity chosen by firm 2 satisfies the following condition:

$$p'(r^*) \left[\underbrace{\left(V^G - V^B\right)(\gamma\,(b) + (1-\gamma\,(b))\, \rho)}_{commercial\ effect} + \underbrace{(\gamma\,(b) - \beta\,(b))\, V^B\, (1-\rho)}_{infringement\ effect} \right] = c'\,(r^*)\,. \tag{3}$$

This expression has a clear interpretation. A larger level of R&D activity increases the probability of the good state of the world, and decreases that of the bad state. This fact has two effects on firm 2's expected profits. On the one side, given the probability of infringement, the expected commercial value of the innovation is larger (commercial effect). On the other side, since $\gamma(b) \geq \beta(b)$, by making the good state of the world more likely, firm 2 reduces the probability of infringing upon firm 1's innovation thus benefitting from the lower expected licensing fees, $(1-\rho)V^B$ (infringement effect). These two effects are clearly highlighted in the left hand side of expression (3). The commercial effect is represented by the first term into the square brackets and it is proportional to $V^G - V^B$. The second term into the square brackets is proportional to the reduction in the probability of not infringing, $\gamma(b) - \beta(b)$, and it represents the infringement effect.

Comparing expressions (2) and (3) the following result holds:[7]

Proposition 2. *Whenever* $(1-\gamma(b))\, V^G \geq (1-\beta(b))\, V^B$*, firm 2 underinvests and it overinvests otherwise.*

According to the above proposition there is underinvestment when V^G is large relative to V^B and when $\gamma(b)$ is close to $\beta(b)$. In other words, firm 2 tends to underinvest when the commercial effect of the R&D activity is large compared to the infringement effect; for instance, if $\gamma(b) = \beta(b)$ the infringement effect disappears, and firm 2 invests less than r^{eff} since it does not obtain the full commercial value of its innovation. Conversely, there is overinvestment when the infringement effect dominates the commercial one; for instance, when $V^G = V^B$ the commercial effect vanishes and firm 2 overinvests.

Note that for each idea $\{p(r), c(r), V^G, V^B, \gamma(b), \beta(b)\}$, there is a level $b > 0$ such that the commercial and the infringement effects balance each other, and firm 2 is induced to invest efficiently. Nevertheless, since the patent breadth is set by the Government before the idea is extracted, then the probability that the selected b induces the efficient R&D activity is null. Therefore, Propositions 1 and 2 make clear that the inability of firm 1 to observe whether firm 2 has already undertaken its R&D activities or not prevents that the licensing of the first innovation to occur efficiently.

[7]The Proof of this and of the following propositions are in the Appendix.

3.2.1 Patent breadth and R&D investment

We are now in the position to analyse the effect of a larger breadth on the investment incentives of the follow-on innovator. Expression (3) implicitly defines the optimal investment level as a function of the patent breadth, $r^*(b)$. Using the implicit function theorem we can calculate the following expression:

$$\frac{\partial r^*}{\partial b} = \frac{p'(r)\left[\left(V^G - V^B\right)\gamma'(b)(1-\rho) + \left(\gamma'(b) - \beta'(b)\right)V^B\left(1-\rho\right)\right]}{-p''(r)\left[\left(V^G - V^B\right)\left(\gamma(b) + \left(1-\gamma(b)\right)\rho\right) + \left(\gamma(b) - \beta(b)\right)V^B(1-\rho)\right] + c''(r)}. \quad (4)$$

The sign of this expression coincides with that of the numerator; more specifically, the numerator represents the marginal variation of the commercial and the infringement effect (the first and the second term into the square brackets, respectively). Consider the commercial effect; an increase in b reduces the commercial value that firm 2 appropriates, and this induces the follow-on innovator to invest less in R&D activities: at the margin the commercial effect decreases with r. The impact of a larger b on the infringement effect is in general indeterminate. When $\gamma'(b) - \beta'(b) > 0$, then as b gets larger the difference between $\gamma(b)$ and $\beta(b)$ also increases; therefore, according to the infringement effect, firm 2 is induced to invest more. When $\gamma'(b) - \beta'(b) < 0$ the opposite occurs.

3.2.2 Firm 1's licensing revenues

Under the assumption that the first innovation is a research tool, firm 1's profits coincide with the licensing revenues it gets from the follow-on innovator. It is interesting to evaluate how these revenues change with the breadth of the patent.

From Proposition 1 follows that firm 1 obtains a share $(1-\rho)$ of the commercial value of the second innovation whenever there is infringement; formally, firm 1's licensing revenues are given by:

$$\Pi_1(b) = (1-\rho)\left[p(r^*)\left(1-\gamma(b)\right)V^G + \left(1-p(r^*)\right)\left(1-\beta(b)\right)V^B\right]$$

With probability $p(r^*)$ the good state of the world occurs; in this case the follow-on innovation infringes on firm 1's patent with probability $(1-\gamma(b))$. Similarly, with probability

$(1-p(r^*))$ the bad state occurs and there is infringement with probability $(1-\beta(b))$. In both cases firm 1 obtains a share $(1-\rho)$ of the commercial value.

The impact of a variation in patent protection on the first innovator's profits is obtained by simple differentiation of $\Pi_1(b)$:

$$\begin{aligned}\frac{\partial \Pi_1(b)}{\partial b} &= (1-\rho)\Bigg(\underbrace{-\left(p(r^*)\gamma'(b)V^G+(1-p(r^*))\beta'(b)V^B\right)}_{direct\ effect}+\\ &\quad +\underbrace{\frac{\partial p(r^*)}{\partial r}\frac{\partial r^*}{\partial b}\left((1-\gamma(b))V^G-(1-\beta(b))V^B\right)}_{indirect\ effect}\Bigg).\end{aligned}$$

As shown above, the effect of a larger b on Π_1 depends on the combination of two effects, a direct and an indirect one. On the one hand, given the investment level chosen by firm 2, the revenues of the first innovator get larger due to the increased probability of infringement; this is the direct effect of an enlarged patent breadth and it has a positive sign. On the other hand, an increase of b alters the investment incentives of firm 2: this is the indirect effect, namely the effect mediated by the change in r^*; since the change in the R&D investment affects both the expected commercial value of the second innovation and the probability of infringement,[8] the indirect effect can be decomposed into two further effects.

In general, the indirect effect has an indeterminate sign and it can either increase or decrease the licensing revenues of firm 1. However, it is worth noticing that the following result holds.

Proposition 3. *The indirect effect is negative whenever a larger patent breadth increases the inefficiency of the R&D investment of firm 2.*

[8]Note that to disentangle the double effect of the change in r^* on the expected commercial value and on the probability of infringement, then the indirect effect should be rewritten as $\frac{\partial p(r^*)}{\partial r}\frac{\partial r^*}{\partial b}\left((1-\gamma(b))\left(V^G-V^B\right)+(\gamma(b)-\beta(b))V^B\right)$.

As proved in Proposition 3, the indirect effect reduces firm 1's revenues in two cases: i) when firm 2 is underinvesting and the increase in the patent breadth induces a further reduction of the investment; ii) when firm 2 is overinvesting and a larger b causes the investment to increase even further.

Obviously, the indirect effect is positive and it goes in the same direction of the direct one whenever a larger b induces a reduction in the inefficiency of firm 2 R&D investment.

3.2.3 Simplified cases

In this section we provide two simplified scenarios whereby we show that the first innovator can indeed be harmed by a too broad and strong patent protection. Formally, we present two cases where the indirect effect is negative and of sufficient magnitude so that it more than compensate the direct effect. In both cases we assume that:

(A.1) $p(r) = r$, and $c(r) = \frac{r^2}{2}$;

(A.2) $\beta'(b) < 0$, and $\lim_{b \to \infty} \beta(b) = x$, with x sufficiently close to zero.

Case 1: no infringement effect ($\gamma(b) = \beta(b)$). In what follows we assume the probabilities of non-infringement are the same in the two states of the world, namely $\gamma(b) = \beta(b)$, for any b. In this case the infringement effect shown in expressions (3) and (4) disappears and, firm 2 investment decision is driven only by the commercial effect.

Absent the infringement effect, from the previous analysis it follows that:

- firm 2 invests less than r^{eff} (see Proposition 2);
- r^* decreases with b (see expression (4)), and therefore firm 2's underinvestment becomes more severe a,s the patent breadth increases;
- the indirect effect is negative (see Proposition 3).

In this simplified setting the following result holds:

Proposition 4. *Assume that (A.1), (A.2) hold and that $\gamma(b) = \beta(b)$ for all b, if $\rho < \frac{1}{2}$, then for any V^G such that $\forall V^G > V^B + \frac{\sqrt{V^B(1-2\rho)}}{(1-2\rho)}$ there exists a $\hat{b}$ such that firm 1's profits are decreasing for any $b > \hat{b}$.*

The above Proposition shows that firm 1 may be harmed when its invention is protected by a patent with a too broad scope. The intuition for this result is the following: a large b exacerbates the underinvestment problem, and this further reduces the value generated by the second innovator, part of which goes to the benefit of firm 1 through the licensing fees. Therefore, firm 1 may be harmed by an increase in the patent breadth because of the smaller commercial value generated by the second innovation. Indeed, as shown in the proof, the (positive) direct effect on firm 1's profits shrinks as b gets larger, given that r^* decreases with the patent breadth; therefore, for large values of b the (negative) indirect effect prevails.

Case 2: no commercial effect and strong infringement effect $V^G = V^B$ and $\gamma(b) = 1$. When $V^G = V^B$, the commercial effect shown in expressions (3) and (4) vanishes and the likelihood of infringement is determined only by firm 2's R&D investment. On top of that, for the sake of simplicity, we assume that infringement never occurs in the good state of the world, formally $\gamma(b) = 1$ for any b.

Absent the commercial effect, from the previous analysis it follows that:

- $r^{eff} = 0$, and firm 2 overinvests (see Proposition 2);

- the marginal infringement effect is positive ($\gamma'(b) - \beta'(b) = -\beta'(b) > 0$) and r^* increases with b so that firm 2's overinvestment becomes more severe as the patent breadth increases (see expression (4));

- the indirect effect is negative (see Proposition 3).

In this simplified setting the following result holds:

Proposition 5. *Assume that (A.1), (A.2) hold and that $V^G = V^B$ and $\gamma(b) = 1$, then for all V^G such that $V^G > \frac{1}{2(1-\rho)}$ there exists a $\hat{b}$ such that firm 1's profits are decreasing for any $b > \hat{b}$.*

In this case, a larger patent breadth may harm firm 1 given that it might reduce the probability of infringement without altering the expected commercial value of the second innovation: as b grows, firm 2 increases its R&D and this has the effect to reduce the probability of infringement with no effect on the value it creates.

In particular the above remark shows that there exists a threshold level for b such that firm 1 profits decrease as the patent breadth grows larger than the threshold. Indeed, the direct effect $\left(-\beta'(b)\left(1-r^*\right)(1-\rho)V^G\right)$ decreases with b, given that the patent breadth increases r^*. Therefore, for large values of b the indirect effect more than compensate the direct one and firm 1's profits decrease with the patent protection.

3.3 Firm 1 moves as first

Proposition 1 makes it clear that all licensing occurs at the fees implemented by the Court. This implies that the licensing contract is not useful in order to improve the R&D incentives of the second innovator, and that, in general, the investment of firm 2 is inefficient. What prevents firm 1 to propose a "more efficient" contract is that if it does so firm 2 comes ex-post, and, at this stage, the first inventor prefers not to offer a contract $\left(\tilde{L}^G, \tilde{L}^B\right) \neq \left(\bar{L}^G, \bar{L}^B\right)$: if firm 2 is ex-post there is no investment to be incentivated and threfore firm 1 is better-off just offering $\left(\bar{L}^G, \bar{L}^B\right)$.

Consider now the following scenario: at time $t = 0$, just after innovating, firm 1 has the possibility to commit to a licensing scheme e.g. it can announce through the internet the conditions under which it is willing to license its invention. Note that at $t = 0$, firm 1 knows for sure that firm 2 has not got its idea, and therefore it has not invested yet; this means that firm 1 could find it profitable to propose $\left(\tilde{L}^G, \tilde{L}^B\right) \neq \left(\bar{L}^G, \bar{L}^B\right)$ in order to give more appropriate R&D incentives to firm 2.

Of course there are different practical reasons why in reality firm 1 does not make such a proposal:

- firm 1 is unable to describe all the possible future applications/ideas of its invention. If it could then it would be able to carry out these applications on its own;

- making the proposal at time $t = 0$, firm 1 ignores the specific idea of the second inventor, $\left\{V^G, V^B, \gamma, \beta, c(r), p(r)\right\}$. This means that it can offer an "average" contract, and this might not be profitable if there is high heterogeneity for instance in the cost or probability functions of the ideas.

Even though it is not the aim of this paper to determine under which circumstances it is optimal for firm 1 to commit to a licensing scheme just after obtaining its innovation, we show that also in a very extreme scenario whereby firm 1 knows the idea of the second innovator, the licensing contract cannot be efficient.

More specifically, consider the extreme scenario where firm 1 knows that there is only one possible idea: $\left\{V^G, V^B, \gamma, \beta, c(r), p(r)\right\}$. The following Proposition shows that also in this case there cannot be efficient licensing

Proposition 6. *Efficiency in licensing agreement is not restored even in the case where firm 1 has the possibility to commit to a licensing scheme at $t = 0$.*

Proof. Consider the case where $\bar{L}^B(1-\beta(b)) \neq \bar{L}^G(1-\gamma(b))$. Note that whatever the proposal of firm 1, firm 2 has the possibility of choosing an ex-post strategy in order to pay the minimum between $\tilde{L}^i$ and $\bar{L}^i$; therefore, we can restrict to the case where firm 1 announces a proposal $\left(\tilde{L}^G, \tilde{L}^B\right)$ such that $\tilde{L}^i \leq \bar{L}^i$, $i = G, B$. In order to induce an efficient investment $\left(\tilde{L}^G, \tilde{L}^B\right)$ has to be such that $\tilde{L}^B(1-\beta(b)) = \tilde{L}^G(1-\gamma(b))$; in this case, firm 1 obtains an expected licensing revenue $\tilde{L}^B(1-\beta(b)) = \tilde{L}^G(1-\gamma(b))$. Note that, from the previous arguments the following two inequalities are verified $\tilde{L}^B(1-\beta(b)) \leq \bar{L}^B(1-\beta(b))$, $\tilde{L}^G(1-\beta(b)) \leq \bar{L}^G(1-\beta(b))$, and in the case where $\bar{L}^B(1-\beta(b)) \neq \bar{L}^G(1-\gamma(b))$, at least one holds with the strict inequality sign. However, this fact implies that firm 1 would be better-off making no offer and obtaining $p(\bar{r})\bar{L}^G(1-\gamma(b)) + (1-p(\bar{r}))\bar{L}^B(1-\beta(b)) >$

$\tilde{L}^B(1-\beta(b)) = \tilde{L}^G(1-\gamma(b))$, provided that $p(\bar{r}) \in (0,1)$, where $\bar{r}$ is the investment level of firm 2 when it pays a licensing fee $\left(\bar{L}^G, \bar{L}^B\right)$. ∎

The intuition of the result is as follows. Once firm 1 has committed to a licensing scheme, firm 2 has always the option of going ex-post. This means that, in case of infringment firm 2 does not pay more than $\bar{L}^i$, $i = G, B$. By offering the efficient contract firm 2 obtains to little and would prefer not to offer any contract.

4 Final Remarks and Policy Implications

Recent theoretical contributions on the economics of sequential innovation have stressed the merits of ex-ante licensing in order to curb the risk that future inventions are held-up by patents protecting existing technologies. Notably, Green and Scotchmer (1995) show that, in a context of symmetric information, there is no hold-up in case the follow-on innovator has the possibility to negotiate the licensing agreement with the patent-holder before incurring the R&D costs.

In this paper we have extended the model of Green and Scotchmer (1995) and we have shown that ex-ante agreements are ineffective in case the first innovator is unable to observe whether the follow-on inventor has already undertaken its R&D activities. The follow-on inventor is better-off by collecting more precise information about its innovation before negotiating the licensing agreement and this fact prevents efficient licensing. In the paper we have assumed that the second inventor observes the commercial value of its innovation just after having sunk the R&D investment; however, intuitively, the arguments behind Proposition 1 hold also in case the second inventor, once incurred the R&D expenditures, simply obtains a signal about the commercial value of its innovation which is more precise than the initial idea.[9]

[9]Note that in the paper we also assume that, after choosing r, firm 2 observes whether there is infringement or not. This assumption does not distort the result of Proposition 1 given that parties, at any time, can write a licensing contract stipulating that the fee is payable contingent on infringement.

An important consequence of the inefficient licensing is that the profits of the first inventor do not necessarily increase with the breadth of the patent that protects its innovation. Indeed, even though an enlarged patent protection increases the likelihood of infringement it also alters the incentives to invest of the follow-on inventor. This latter, indirect, effect of a larger patent breadth might go to the detriment of the first inventor; we have shown that this occurs whenever the increase in the patent protection exacerbates the inefficiency in the R&D of the second innovator.

In the paper we provide two simplified scenarios whereby we show that the first inventor is actually damaged by a too broad patent protection. This fact might occur because a large patent breadth may either i) reduce the incentives to invest of the second innovator and thus lower the commercial value of its innovation, part of which goes to the benefit of the first innovator through licensing (Proposition 4), or ii) it may induce the second innovator to overinvest in order to reduce the probability of infringement (Proposition 5).

Throughout the paper, we have assumed that the first and the second innovations/inventors are not rivals; nonetheless, it would be possible to show that, qualitatively, our main results hold also when the first innovation has a commercial value and the follow-on invention steals the business of the initial inventor. A result similar to that of Proposition 4 may also occur when the follow-on innovation/product "increases substantially the willingness to pay of consumers" e.g. because it markets a product of superior quality or, in case of industries with network externalities, because it enlarges the size of the installed base of products. In this scenario, a large patent breadth reduces the incentives to invest of the second innovator and the early inventor may be harmed exactly for the same reason as in Proposition 4.[10]

Similarly, in case the R&D investment affects only the probability of infringement but not the quality of the product with which firm 1 competes, a large patent breadth may be harmful for the early inventor because of the same reason as in Case 2.

[10]It is well known that in markets characterized by network effects, an incumbent provider may prefer to face a compatible rival rather then acting monopolistically; the presence of the rival enlarges the installed base of adopters and this may go to the benefit of the incumbent through the larger network effects (Economides, 1996).

Policy Implications. Our paper contributes to the ongoing discussion on the effects of (strong) intellectual property rights in industries where innovation is sequential. As we have made it clear, whenever the follow-on inventor has the possibility to conceal some relevant information about its invention, the availability of ex-ante contracts is ineffective and licensing agreements are inefficient. This result rests on the assumption that both the amount of the R&D investment of the second inventor, and when such investment is undertaken are not observable. This assumption seems plausible, especially if one thinks to industries such as software, hardware and more broadly to high-tech sectors, where large part of R&D is made of intellectual activities aimed at knowledge creation; the very nature of these activities is clearly intangible and therefore of more difficult measurement.

Even though a careful analysis of the optimal patent breadth goes beyond the scope of this paper, it is interesting to note that the inefficiency at the licensing stage implies that the patent breadth does not only influence the division of profits among different innovators, but also the overall amount. In particular, we have shown that the inefficiency at the licensing stage might be so substantial that a large patent breadth may actually be Pareto-dominated in that it may harm the first inventor too. This implies that it might be in the mutual interest of the two innovators to have a lighter patent protection.

Interestingly, we show that at the equilibrium not only underinvestment but also overinvestment might occur. There is overinvestment when the R&D efforts reduce the probability of infringement without substantially affecting the value of the innovation; this amounts to saying that the follow-on innovator spends resources in order to design around existing products. This is obviously inefficient since it represents a mere duplication of efforts. The software industry represents a clear example of an industry where this is frequently the case: very often, commercial firms prefer to re-write from scratch modules or lines of codes instead of using the already existing ones just to avoid patent infringement. In this case, a clear overinvestment occurs: the duplication of the lines of code does not add value but it decreases the probability of infringement.

References

Bessen, J. (2004). Holdup and licensing of cumulative innovations with private information. *Economic Letters*, 82:321–326.

Bessen, J. and Maskin, E. (2007). Sequential innovation, patents and imitation. Forthcoming in the RAND Journal of Economics.

Chang, H. (1995). Patent scope, antitrust policy, and cumulative innovation. *RAND Journal of Economics*, 26:34–57.

Economides, N. (1996). Network externalities, complementarities, and invitations to enter. *European Journal of Political Economy*, 12:211–233.

Gallini, N. and Scotchmer, S. (2002). *Intellectual property: when is it the best incentive mechanism?* in Adam Jaffe, Joshua Lerner and Scott Stern, eds (2002) Innovation policy and the economy, Vol. 2, MIT Press, pp. 51-78.

Graham, S. J. and Mowery, D. (2004). Submarines in software? Continuations in US software patenting in the 1980s and 1990s. *Economics of Innovation and New Technology*, 13(5):443–456.

Green, J. R. and Scotchmer, S. (1995). On the division of profit in sequential innovation. *RAND Journal of Economics*, 26(1):20–33.

Heller, M. and Eisenberg, R. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science*, 280:698–701.

Lemley, M. A. and Shapiro, C. (2007). Patent holdup and royalty stacking. *Texas law review*, 85:1190–2049.

Matutes, C., Regibeau, P., and Rockett, K. (1996). Optimal patent protection and the diffusion of innovation. *RAND Journal of Economics*, 27:60–83.

Merges, R. P. and Nelson, R. (1990). On the complex economics of patent scope. *Columbia Law Review*, 90(4):839–916.

Nordhaus, W. (1969). *Invention, growth and welfare*. Cambridge, Mass: MIT Press.

O'Donoghue, T., Scotchmer, S., and Thisse, J. (1998). Patent breadth, patent life, and the pace of technological improvement. *Journal of Economics and Management Strategy*, 7:1–32.

Schankerman, M. and Scotchmer, S. (2001). Damages and injunctions in the protection of intellectual property. *RAND Journal of Economics*, 32:199–220.

Scotchmer, S. (1991). Standing on the shoulders of giants: cumulative research and the patent law. *Journal of Economic Perspectives*, 5(1):29–41.

Scotchmer, S. (1996). Protecting Early Innovators: Should Second-Generation Products be Patentable? *RAND Journal of Economics*, 27:322–331.

Shapiro, C. (2001). *Navigating the patent thicket: cross licensing, patent pools and standard setting*. in Jaffe et al. eds (2001) Innovation policy and the economy.

Siebert, R. and von Graevenitz, G. (2006). Jostling for advantage: licensing and entry into patent races. CEPR - Discussion Paper Series - 5753, July 2006.

5 Appendix

Proof of Proposition 2

Note that $V^G - V^B \geq \left(V^G - V^B\right)\left(\gamma\left(b\right) + \left(1-\gamma\left(b\right)\right)\rho\right) + \left(\gamma\left(b\right) - \beta\left(b\right)\right)V^B\left(1-\rho\right)$ if and only if $\left(1-\gamma\left(b\right)\right)V^G \geq \left(1-\beta\left(b\right)\right)V^B$; in this case, $r^{eff} \geq r^*$ from the assumptions $p''(r) \leq 0$ and $c''(r) > 0$. Similarly, $\left(1-\gamma\left(b\right)\right)V^G < \left(1-\beta\left(b\right)\right)V^B$ implies $r^{eff} < r^*$.

Proof of Proposition 3

The indirect effect is given by $\frac{\partial p(r^*)}{\partial r}\frac{\partial r^*}{\partial b}\left(\left(1-\gamma(b)\right)V^G-\left(1-\beta(b)\right)V^B\right)$ and it is negative either when $\left(\left(1-\gamma(b)\right)V^G>\left(1-\beta(b)\right)V^B\right)$ and $\frac{\partial r^*}{\partial b}<0$ or when $\left(\left(1-\gamma(b)\right)V^G<\left(1-\beta(b)\right)V^B\right)$ and $\frac{\partial r^*}{\partial b}>0$.

Proof of Proposition 4

Straightforward calculations yield $r^*=\left(V^G-V^B\right)\left(\beta(b)+\left(1-\beta(b)\right)\rho\right)$, and $\frac{\partial\Pi_1(b)}{\partial b}=-\beta'(b)\left(1-\rho\right)\left(\left(V^G+V^B\right)r^*-V^B\right)+\frac{\partial r^*}{\partial b}\left[\left(1-\beta(b)\right)\left(1-\rho\right)\left(V^G-V^B\right)\right]$. After some manipulations it is possible to show that $\frac{\partial\Pi_1(b)}{\partial b}<0$ if and only if $\omega\left(V^G\right)^2-\omega V^G+\omega\left(V^B\right)^2-V^B>0$, where $\omega=1+2\left(\rho\beta(b)-\beta(b)-\rho\right)$. The previous expression has real roots if and only if $\omega>0$, that is if and only if $\rho<\frac{1-2\beta(b)}{2(1-\beta(b))}$; note that $\frac{1-2\beta(b)}{2(1-\beta(b))}$ decreases with β (and then increases with b) and $\lim_{b\to\infty}\frac{1-2\beta(b)}{2(1-\beta(b))}=\frac{1}{2}$, and therefore provided that $\rho<\frac{1}{2}$, $\omega>0$ for b large enough.

The real roots of the expression $\omega\left(V^G\right)^2-\omega V^G+\omega\left(V^B\right)^2-V^B$ are $\left(V^G\right)_{1,2}=V^B\pm\frac{\sqrt{\omega V^B}}{\omega}$ and only the positive root is relevant, given that $V^G\geq V^B$ by assumption; therefore, $\frac{\partial\Pi_1(b)}{\partial b}<0$ for any $V^G>V^B+\frac{\sqrt{\omega V^B}}{\omega}$. Note that $V^B+\frac{\sqrt{\omega V^B}}{\omega}$ increases with β (and then decreases with b) and $\lim_{b\to\infty}V^B+\frac{\sqrt{\omega V^B}}{\omega}=V^B+\frac{\sqrt{(1-2\rho)V^B}}{1-2\rho}$ and therefore for any $V^G>V^B+\frac{\sqrt{V^B(1-2\rho)}}{(1-2\rho)}$ it is possible to find a threshold level for b such that $\frac{\partial\Pi_1(b)}{\partial b}<0$ when b grows above the threshold.

Proof of Proposition 5

Straightforward calculations yield $r^*=V^G\left(1-\beta(b)\right)\left(1-\rho\right)$, and $\frac{\partial\Pi_1(b)}{\partial b}=-\beta'(b)\left(1-r^*\right)\left(1-\rho\right)V^G-\frac{\partial r^*}{\partial b}\left(1-\beta(b)\right)\left(1-\rho\right)V^G$. After some manipulations it is possible to show that $\frac{\partial\Pi_1(b)}{\partial b}<0$ provided that $V^G>\frac{1}{2(1-\beta(b))(1-\rho)}$; note that $\frac{1}{2(1-\beta(b))(1-\rho)}$ increases with β (and then decreases with b), and $\lim_{b\to\infty}\frac{1}{2(1-\beta(b))(1-\rho)}=\frac{1}{2(1-\rho)}$. Therefore, $\forall V^G>\frac{1}{2(1-\rho)}$ there exists a $\hat{b}$ such that $\frac{\partial\Pi_1(b)}{\partial b}<0$ for any $b>\hat{b}$.