

Sequential innovations with unobservable follow-on investments^{*}

Stefano Comino[‡] Fabio, M. Manenti[§] Antonio Nicolò[¶]

September 16, 2007

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. When the second innovator investments are not observable, licensing of the first innovation never occurs efficiently, and, at the equilibrium, the follow-on innovator either underinvests or overinvests. We show that a large patent breadth may be harmful for the first innovator too, and therefore Pareto-dominated; as long as the undervestment problem becomes more pronounced, the value generated by the follow-on innovator reduces, and so do the licensing revenues of the first inventor.

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

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

^{*}Paper presented at the Dipartimento di ingegneria gestionale - Politecnico di Milano, and at the 34th EARIE Conference, Valencia 2007.

[†]Copyright of the paper resides with the authors. Submission of a paper grants permission to the EPIP-2007 Conference Scientific and Organizing Committees to include it in the conference material and to place it on relevant websites.

[‡]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 with 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 subsequent inventors sunk specific investment before negotiating the terms of the license agreement with the initial patent-holder. The follow-on inventor is in a weak bargaining position in that the surplus on which parties negotiate is represented by the commercial value of the derived innovation and does not take

¹According 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.

into account the costs that have already been sunk (Lemley and Shapiro, 2007; Scotchmer, 1991; Shapiro, 2001);² in this way, the follow-on innovator might not be able to recoup the whole R&D investment. The threat that very broad intellectual property rights may slow down the pace of future innovation is compounded in case of “patent thickets”, that is, when several patents read, at the same time, on a given product or technology.³ In these instances, the downstream innovator needs to negotiate licensing agreements with every single patent-holder. Besides exacerbating the hold-up, patent thickets may also cause the “complements problem”.⁴ When negotiating the licensing agreement with the downstream innovator, each patent-holder imposes a negative externality on the other patent-holders: by requiring a large licensing fee it reduces what the other holders may collect. This fact may increase the overall burden of licensing fees that the downstream innovator bears up to a point where subsequent discoveries may be threatened.

As explicitly suggested already in Scotchmer (1991), parties can curb the hold-up problem described above 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; that is, 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 the analysis of the optimal

²According 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).

³Patent thickets are common in the IT sector where many different components of a technology are protected by one or several patents (Lemley and Shapiro, 2007; Siebert and von Graevenitz, 2006).

⁴The analysis of the classic “complements problem” goes back to Cournot, 1838, and Shapiro (2001) presents an application to the case of cumulative innovation. Heller and Eisenberg (1998) discuss the consequences of multiple blocking patents in the context of biomedical research, and warn against a potential “tragedy of the anti-commons”: in the presence of multiple patent-holders the resources may be underused.

patent policy in case of cumulative/sequential innovation using a two-stage game. In the first stage, the original innovator chooses whether to invest in order to develop her idea, and, in case she does so, in the second stage, a second innovator has the opportunity to improve upon the original discovery. In case the improvement infringes on the patent of the original discovery, in order to commercialize his invention the second innovator has to obtain a license from the original patent-holder. Assuming that parties negotiate in a context of symmetric information, the authors show that ex-ante licensing contracts, i.e. signed before the follow-on innovator has sunk its R&D costs, ensure that the improvement is realized whenever efficient. In this scenario, the main task of the patent policy is to ensure enough rewards to the original inventor, and this is accomplished by granting her a very strong patent protection; in particular, patent breadth, which determines how profits are actually shared among the two inventors, should be very large, if not infinite.

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).⁵ In a recent paper, Bessen (2004) considers the case where the development costs of the improvement are private information of the follow-on innovator. Bessen shows 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.⁶

In this paper, we present a model based on Green and Scotchmer (1995) and we obtain results much sharper than those in Bessen (2004). Under the realistic assumption that when

⁵Matutes et al. (1996) and Chang (1995) do not consider the possibility of ex-ante licenses. See Gallini and Scotchmer (2002) for a recent review on these issues

⁶Siebert 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: augmenting the number of patents that a firm possesses strengthen 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 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.

contracting over the licensing terms the early innovator cannot observe whether the follow-on inventor has already undertaken the R&D activity, investment to develop the derived invention is always inefficient. In particular, the mere inability of the early innovator to tell whether the follow-on inventor is “truly ex-ante” always prevents efficiency. The intuition for this result is simple. If the early innovator were sure that the follow-on inventor is approaching her ex-ante, then she would be willing to offer her technology at an efficient lump-sum licensing fee. However, in this case, the follow-on inventor benefits from asking the licensing agreement ex-post: he pays the fee only in case of infringement, and, on top of that, he can take advantage of the possibility of choosing the most favorable licensing terms by selecting between the proposal of the early innovator and the fees implemented by the courts. Given this, the early innovator is better-off conforming to the fees implemented by the courts and not proposing to license her technology at an efficient lump-sum fee.

We show that, at the equilibrium, both underinvestment and overinvestment 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 invests more than the efficient level.

Interestingly, the inefficiency in the R&D investment of the second innovator has important consequences in terms of patent policy. While in Green and Scotchmer (1995) patent breadth only affects the division of profits among initial and follow-on innovators, in this paper it also affects the incentives to invest in R&D of the latter. In particular, in the paper we provide three simplified examples whereby we show that the first innovator can be harmed when the breadth of the patent protecting its invention is too broad. As a consequence, we prove that, there are circumstances where a large patent breadth is Pareto-dominated, that is both innovators are better-off with a limited breadth. This result 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 do 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 that goes to the detriment of all the industry participants.

The paper is organized as follows: in Section 2 we present the outline of the paper, we derive the main results and discuss the policy implications of our analysis. In Section 3 we check the robustness of our results to possible generalisations of the model. Finally, in Section 4 we conclude.

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. 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*.

The focus of the paper is on the second inventor’s behavior being the first innovation already in place and protected by a patent.

The idea that the second inventor gets 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 a quadruple $\{F^G(v), F^B(v), \gamma(b), \beta(b)\}$ whose

terms are described below. In order to develop the idea, firm 2 has to undertake a certain amount of R&D activity, $r \geq 0$ incurring a cost $c(r)$; once the R&D cost has been sunk, then with probability r a “good state” of the world occurs, and with probability $(1 - r)$ a “bad state” of the world occurs. In the former case the innovation, i) has a commercial value $v \geq 0$ distributed according to $F^G(v)$ and with an expected value of V^G , and ii) it does not infringe the patent of the first innovation with probability $\gamma(b)$, where $b \in \mathfrak{R}_+$ represents the patent breadth set by Government regulations. In the bad state, v is distributed according to $F^B(v)$ with an expected value of $V^B \leq V^G$, and $\beta(b)$ is the probability that the follow-on innovation does not infringe the patent of firm 1, with $\beta(b) \leq \gamma(b)$, $\forall b$. In other words in the good state of the world both the expected commercial value and the probability of not infringing of the follow-on innovation are larger than in the bad state. It should be noted that these facts imply that: i) there is a positive relationship between the expected commercial value and the probability of not infringing of the second innovation;⁷ ii) a larger r , i.e. a larger R&D activity, increases both the expected commercial value, $rV^G + (1 - r)V^B$, as well as the probability that the second innovation does not infringe on the patent of the first inventor, $r\gamma(b) + (1 - r)\beta(b)$.

Probabilities of not infringing decrease with the patent breadth: $\gamma'(b) \leq 0$, and $\beta'(b) \leq 0$. Moreover, we assume that $\lim_{b \rightarrow \infty} \gamma(b) = \lim_{b \rightarrow \infty} \beta(b) = x$, that is when the Government sets the breadth at its maximum level then the second invention does not infringe the patent of firm 1 with probability $x \in [0, 1)$. Note that $x > 0$ implies that, even in the case of maximum patent protection, follow-on innovators still have a chance not to infringe.⁸ Finally, we assume that $c(r)$ is increasing and convex.

Timing and information structure of the game

In case the follow-on innovation infringes on the patent protecting the early innovation, firm

⁷When $\gamma(b) = \beta(b)$, $\forall b$, the probability of not infringing and the commercial value are uncorrelated.

⁸Patent breadth is set based on the current state of the art. Therefore, it is reasonable to assume that, even in the case of infinite breadth, in the future follow-on innovators still have a positive probability of not infringing.

2 needs a license from firm 1 in order to market the invention.

The timing of the game is as follows :

1. once the first inventor patents its innovation, the second inventor gets an idea, formally a quadruple $\{F^G(v), F^B(v), \gamma(b), \beta(b)\}$;
2. firm 2 faces an alternative: *i*) going to the first inventor asking for a license before having undertaken any R&D activity; or *ii*) asking for a license after having sunk the cost of R&D, thus having observed both commercial value, v , 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.

The details of the contracting phase between first and second inventor are specified below when we discuss how parties bargain over the licensing agreement.

In what follows we will 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. Finally, we assume that both the commercial value v 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.

2.1 Licensing agreement

We assume that the bargaining stage to determine the licensing terms is as follows:

- firm 2 approaches firm 1 asking for a licensing agreement;
- firm 1 proposes a fee or a menu of fees at which it is willing to license its innovation, and firm 2 can either select one of the proposed fees or reject any proposal. In case

firm 2 rejects any offer, then, when there is infringement, the court imposes a licensing fee $L(v) \equiv (1 - \rho)v$, with $\rho \in [0, 1]$. 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 back a more or less pronounced “pro-patent” environment.

Firm 1 can propose a fixed fee and/or a fee that varies with the commercial value v ; moreover, the payment can be contingent on infringement or it might be lump-sum, and payable independently of whether there is infringement or not. Note that, in case of infringement, firm 2 has always the option to resort (ex-post) to the court and obtain the license at the fee $L(v)$; therefore, we can assume that firm 1’s proposal always includes the licensing fee $L(v)$ payable contingent on infringement.

The following proposition shows that at the equilibrium the first innovation is always licensed at the fee $L(v)$ payable contingent on infringement.

Proposition 1. *All licensing occurs at the fee $L(v) = (1 - \rho)v$ payable contingent to the infringement.*

Proof. As reported above, firm 1 always proposes the licensing fee $L(v) = (1 - \rho)v$ payable contingent upon infringement. In what follows, we show that at the equilibrium a lump-sum fee \bar{L} is not proposed by firm 1 or, if it is, it is never accepted by firm 2.

Suppose, on the contrary, that at the equilibrium firm 1 offers \bar{L} with some positive probability. We need to consider three cases:

- i) when $\bar{L} > L(v)$ for every possible realization of v , then firm 2 never accepts the fee \bar{L} ;
- ii) when $\bar{L} \geq L(v)$ for some realizations of v and $\bar{L} \leq L(v)$ for other realizations of v , then firm 2 asks for a licensing agreement ex-post, once it has observed v and it knows whether there is infringement. In this way it pays the licensing fee if and only if there is infringement and, provided that \bar{L} is offered, it selects the most favorable contract paying a fee equal to the $\min\{L(v), \bar{L}\}$. This strategy is preferred to choosing $L(v)$ ex-ante, since firm 2 pays $\min\{L(v), \bar{L}\} \leq L(v)$, whenever \bar{L} is offered. Similarly, approaching firm 1

ex-post is preferred to choosing \bar{L} ex-ante since $\min \{L(v), \bar{L}\} \leq \bar{L}$, and the license is paid in case of infringement only. However, if firm 2 approaches firm 1 ex-post, then firm 1 is better-off not offering \bar{L} : firm 2 has already made the investment (there is no investment to be incentivated through a fixed licensing fee), the variable contract $L(v)$ is acceptable by firm 2 and therefore by offering \bar{L} with positive probability firm 1 only lowers its licensing revenues;

iii) when $\bar{L} < L(v)$, then again firm 2 approaches firm 1 ex-post and pays \bar{L} if and only if there is infringement. Also in this case firm 1 is better off not offering \bar{L} .

Similar arguments apply to the other possible licensing fees that might be included in the menu that firm 1 proposes; namely they apply both to the case of *i*) a variable licensing fee $\tilde{L}(v)$ payable both when there is infringement or not, and *ii*) a licensing fee $\hat{L}(v)$ different from $L(v)$ and payable only in case of infringement. ■

Proposition 1 can be intuitively interpreted. Being offered a menu of contracts, firm 2 will certainly approach firm 1 ex-post, in order to exploit its informational advantage; by asking for the licensing agreement once $c(r)$ has been sunk, firm 2 enjoys a double benefit: it pays the fee if and only if there is infringement, and, whenever a menu of contracts is offered, it also selects the most favorable one. This implies that by offering a menu of contracts firm 1 certainly reduces its licensing revenues; therefore, it is better-off by proposing $L(v)$ only.

Note that Proposition 1 does not necessarily imply that all licensing agreements occur ex-post; indeed, there are two equilibria of the bargaining game. In the first equilibrium, the follow-on inventor approaches firm 1 ex-post and it is offered $L(v)$. The second equilibrium is in mixed strategies: *i*) firm 2 randomizes and approaches firm 1 ex-ante and ex-post with positive probability, and *ii*) the probability according to which firm 2 goes ex-ante is small enough and such that firm 1 best response is still offering no other contracts than $L(v)$ payable contingent on the infringement.

2.2 R&D investment and licensing revenues

9

Having defined the equilibrium in the licensing stage, we can solve for the optimal amount of R&D activities performed by firm 2. Before that, it is useful to look at 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 is:

$$r^{eff} = \operatorname{argmax}_r \{rV^G + (1-r)V^B - c(r)\}.$$

The first order condition is simply:

$$V^G - V^B = c'(r^{eff}). \quad (1)$$

Let us now consider the investment level that firm 2 actually chooses. From Proposition 1 we know that whenever the follow-on innovation infringes the patent of the first inventor then firm 2 gets only a share ρ of the commercial value generated by its innovation. Therefore, firm 2 chooses r in order to maximize the following expression:

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

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:

$$(V^G - V^B) [\gamma(b) + \rho(1-\gamma(b))] + (\gamma(b) - \beta(b)) V^B (1-\rho) = c'(r^*). \quad (2)$$

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

⁹In most of the arguments of this section we assume interior solutions for the R&D choice.

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 expression (2). The commercial effect is represented by the first term of the expression which is proportional to $V^G - V^B$. The second term is proportional to the reduction in the probability of not infringing, $\gamma(b) - \beta(b)$, and represents indeed the infringement effect.

From a simple comparison of expressions (1) and (2) it is immediate to determine under which conditions firm 2 under or overinvests .

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 not too large with respect to $\beta(b)$ (recall that $\gamma(b) \geq \beta(b)$). 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 $\gamma(b) = \beta(b)$, there is underinvestment since the infringement effect disappears and firm 2 does not obtain the full commercial value of its innovation. Conversely, when V^G and V^B are close in magnitude, the infringement effect tends to dominate and firm 2 overinvest.¹⁰

Note that for each idea $\{F^G(v), F^B(v), \gamma(b), \beta(b)\}$, 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 imply that the inability of firm 1 to observe whether firm 2 has already undertaken its R&D activities or not always prevents that the licensing of the first innovation occurs efficiently. This is in sharp contrast to what shown in Green

¹⁰A practical observation supporting this result comes from the software industry. 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.

and Scotchmer (1995) where, in a symmetric information context, ex-ante contracts always ensure an efficient licensing agreement between subsequent inventors.

It is interesting to note that, under the mild condition that the probabilities of non-infringement decrease smoothly with b , the following corollary applies:

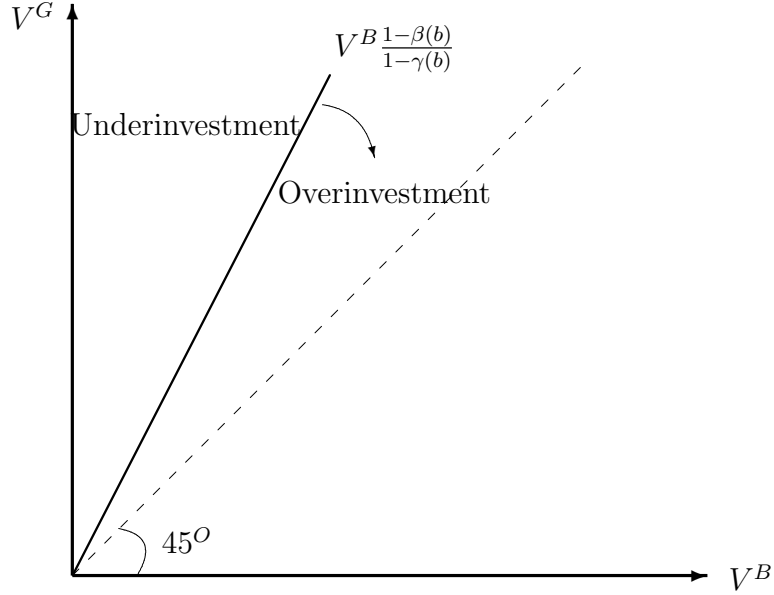


Figure 1: regions of over underinvestment as b grows above \tilde{b}

Corollary 3. *Assuming that $\gamma''(b) > 0$ and $\beta''(b) > 0$, $\forall b$, then there exists a unique level of the breadth \tilde{b} such that:*

- when $b > \tilde{b}$, an increase in the patent breadth makes underinvestment more likely.

Proof. In order to prove this corollary, recall that $\gamma(0) = \beta(0) = 1$, $\lim_{b \rightarrow \infty} \gamma(b) = \lim_{b \rightarrow \infty} \beta(b) = x$, and that $\gamma(b) \geq \beta(b)$, $\forall b$; therefore the assumption $\gamma''(b) > 0$ and $\beta''(b) > 0$, $\forall b$ implies that there exists a unique value of the breadth, \tilde{b} , such that $\forall b < \tilde{b}$, $\beta'(b) < \gamma'(b)$ while $\forall b > \tilde{b}$, $\beta'(b) > \gamma'(b)$.

According to Proposition 2, firm 2 underinvests whenever $\frac{V^G}{V^B} \geq \frac{(1-\beta(b))}{(1-\gamma(b))}$. It can be easily shown that the rhs of the previous expression decreases as b grows larger than \tilde{b} . Therefore,

as the patent breadth gets bigger the set of combinations of V^G and V^B such that there is infringement enlarges. ■

The explanation for this corollary relies on the strength of the infringement effect: when $b > \tilde{b}$, the difference $\gamma(b) - \beta(b)$ reduces with the patent breadth; therefore an increase in b lessens the infringement effect and therefore reduces the investment incentives. Figure 1 provides a graphical representation of Corollary 3. All the combinations of $\{V^G, V^B\}$ that lie above the $V^B \frac{1-\beta(b)}{1-\gamma(b)}$ line represent cases in which there is underinvestment; contrarily, below the line there is overinvestment. As b grows the $V^B \frac{1-\beta(b)}{1-\gamma(b)}$ rotates clockwise and thus the region where underinvestment occurs enlarges.

2.2.1 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 (2) 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:

$$\frac{\partial r^*}{\partial b} = \frac{(1 - \rho) (\gamma' (b) (V^G - V^B) + (\gamma' (b) - \beta' (b)) V^B)}{c'' (r)}. \quad (3)$$

The sign of this expression coincides with that of the numerator; more specifically, the numerator simply represents the marginal variation of the commercial and the infringement effect (the first and the second term, respectively). An increase in b always 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 always decreases. The impact of a larger b on the infringement effect is more complicated, and 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 invests more. When $\gamma' (b) - \beta' (b) < 0$ the opposite occurs.

Due to the indeterminacy on how a larger breadth impacts on the infringement effect, in general we are not able to give a clear sign to expression (3). Nonetheless, under the mild

condition that the probabilities of non-infringement, $\gamma(b)$ and $\beta(b)$, decrease smoothly with b , then the following result holds:

Proposition 4. *Assuming that $\gamma''(b) > 0$ and $\beta''(b) > 0$, $\forall b$, then $\frac{\partial r^*}{\partial b} < 0$ for any $b \geq \tilde{b}$, where \tilde{b} is defined in Corollary 3.*

Proof. From the proof of Corollary 3 we know that $\gamma'(b) - \beta'(b) < 0$, $\forall b > \tilde{b}$. This is enough to prove the proposition. ■

As already explained above the assumption of convexity of the probabilities $\gamma(b)$ and $\beta(b)$ implies that for $b \geq \tilde{b}$ the infringement effect shrinks as the patent breadth gets larger. Therefore, in this case, a larger b reduces both the commercial as well as the infringement effect, and this explains Proposition 4.

2.2.2 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.

Firm 1 obtains a share of the commercial value of the second innovation whenever there is infringement; formally, firm 1's licensing revenues are given by:

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

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

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

$$\begin{aligned} \frac{\partial \Pi_1(b)}{\partial b} = & (1 - \rho) \left(- (r^*(b) \gamma'(b) V^G + (1 - r^*(b)) \beta'(b) V^B) + \right. \\ & \left. + \frac{\partial r^*}{\partial b} ((1 - \gamma(b)) V^G - (1 - \beta(b)) V^B) \right) \end{aligned}$$

The sign of this derivative is the combination of two effects, a direct and an indirect one. More specifically, as b gets larger:

1. given the investment level chosen by firm 2, the revenues of firm 1 get larger due to the increased probability of infringement. This is the direct effect which, formally, is given by $- (r^*(b) \gamma'(b) V^G + (1 - r^*(b)) \beta'(b) V^B)$ and it is always non-negative;
2. the indirect effect is the effect mediated by the change in $r^*(b)$; it can be decomposed into two further effects since the change in the R&D activities affects both the expected commercial value of the second innovation and the probability of infringement. Formally, the indirect effect is given by $\frac{\partial r^*}{\partial b} ((1 - \gamma(b)) V^G - (1 - \beta(b)) V^B)$, and it has an indeterminate sign.¹¹

Clearly, the sign of $\partial \Pi_1(b) / \partial b$ depends on the sign and the magnitude of the indirect effect; although we cannot fully determine how firm 1's licensing revenues vary with b , we can still characterise them in some specific but interesting cases.

Proposition 5. *Assuming that $\gamma''(b) > 0$ and $\beta''(b) > 0$, $\forall b$, then $\forall b \geq \tilde{b}$:*

- *when firm 2 overinvests, then the indirect effect has a positive sign and therefore also $\partial \Pi_1(b) / \partial b$ is positive;*
- *when firm 2 underinvests, then the indirect effect is negative and therefore $\partial \Pi_1(b) / \partial b$ has an indeterminate sign,*

¹¹Note that to disentangle the double effect of the change in $\partial r^*(b)$ on the expected commercial value and on the probability of infringement, then the indirect effect should be rewritten as $\frac{\partial r^*}{\partial b} ((1 - \gamma(b)) (V^G - V^B) + (\gamma(b) - \beta(b)) V^B)$.

where \tilde{b} is defined in Corollary 3.

The first innovator tends to benefit from a larger patent breadth (the direct effect is always positive). However, as shown in the above proposition a larger b may have a negative indirect effect that moves in the opposite direction. In particular, this happens when firm 2 underinvests and the patent breadth is increased beyond \tilde{b} ; in this case, the reduction in firm 2's R&D activities hurts firm 1 since the decrease in the commercial value that is generated is not compensated by the increase in the probability of infringement.

It is worth noting that the presence of the indirect effect is the main consequence of the inefficiency in firm 2's investment decision and this is due to the fact that licensing contracts are not lump-sum. The inability of the first inventor to observe when the follow-on innovator carries out its R&D activities is what drives these effects.

2.2.3 Examples

In this section we provide two extremely simplified examples whereby we show that the first innovator can indeed be harmed when the breadth of the patent protecting its invention is too broad. The key drivers of these results are the commercial effect (first example) and the infringement effect (second example) of the R&D activity of the second innovator. In both examples, we assume that $c(r) = \frac{r^2}{2} + r$, $\beta(b) = \max\{0, 1 - b\}$.

Example 1

Consider the case where $V^B = 0$ and $\gamma(b) = \beta(b)$, that is the probability of non-infringement is the same in the two states of the world. In this simplified setting it is possible to show the following result:

Remark 6. *When $V^G \in \left(2, \frac{1}{\rho}\right)$, firm 1 obtains larger profits when the patent has a limited breadth, namely when b is smaller than 1.*

Proof. Firm 2 chooses $r \in [0, 1]$ in order to maximize $rV^G (\beta(b) + (1 - \beta(b))\rho) - c(r)$. It can be easily shown that when $V^G \in \left(2, \frac{1}{\rho}\right)$ firm 2 selects $r^* = 0$ when $b = 1$ and $r^* = 1$ for

any $b \in [0, \bar{b}]$, where $\bar{b} = \frac{V^G - 2}{V^G(1-\rho)}$, that is such that $V^G (\beta(\bar{b} + (1 - \beta(\bar{b})))) = c'(1)$. Firm 1's profits are $V^G(1 - \rho)(1 - \beta(\bar{b}))$, when $b = \bar{b}$ and 0 when $b = 1$. Note that the set $V^G \in \left(2, \frac{1}{\rho}\right)$ is non-empty provided that $\rho < \frac{1}{2}$. ■

When $\gamma(b) = \beta(b)$ the infringement effect is absent and $r^*(b)$ decreases with b . In particular, when the patent breadth is set at its maximum level, the second innovator does not invest in R&D, and this drives firm 1's licensing revenues to zero.

Example 2

Consider the case where $\gamma(b) = 1$, and $V^B = V^G = V$, that is the R&D carried out by the second innovator has no commercial effect but reduces the probability of infringement. In this setting the following result holds true:

Remark 7. *When $V \geq \frac{2}{1-\rho}$, firm 1 obtains larger profits when the patent has a limited breadth, namely when b is smaller than 1.*

Proof. Firm 2 chooses $r \in [0, 1]$ in order to maximize $rV + (1-r)V(\beta(b) + (1 - \beta(b))\rho) - c(r)$. It can be easily shown that when $V \geq \frac{2}{1-\rho}$ firm 2 selects $r^* = 1$ when $b = 1$ and $r^* = 0$ for any $b \in [0, \hat{b}]$, where $\hat{b} = \frac{1}{V(1-\rho)}$, that is such that $V(1 - \beta(\hat{b}))(1 - \rho) = c'(0)$. Firm 1's profits are $\hat{b}V(1 - \rho)$ when $b = \hat{b}$ and 0 when $b = 1$. ■

In this example firm 2's R&D does not affect the expected commercial value of the innovation and it is driven by the infringement effect only. In this case $r^*(b)$ increases with the patent breadth and firm 1 is harmed by a too large protection given that the probability of infringement is reduced without any increase in the commercial value.

2.3 Policy implications

Having examined the impact of patent breadth on the follow-on innovator's investment decision, and on the first inventor licensing revenues, we can now discuss the effects of patent breadth on the Social Welfare, here represented by the sum of the profits of the two firms.

The impact of an increase in b on firm 2 is obvious: a stronger patent protection undoubtedly reduces its profits. On the contrary, the above discussion shows that the relation between patent breadth and firm 1's licensing revenues is, in general, indeterminate. Nonetheless, a very broad patent protection can go to the detriment of the first innovator too; this may occur when the infringement effect of the R&D investment of the second innovator is very strong (example 2) or when the commercial effect is strong (example 1). In these circumstances, a large patent breadth is Pareto-dominated in that both firms would be better-off with a more limited patent protection.

In the same spirit as example 1, the results for the case of smooth probabilities of non-infringement, indicate that a large patent breadth might be socially not desirable. When b grows further above the threshold level \tilde{b} :

- a) according to Corollary 3, the probability that firm 2 underinvests gets large as b increases;
- b) Proposition 4 shows that a larger breadth induces a lower R&D activity by the follow-on inventor; this effect combined with a) implies that the underinvestment problem becomes more severe;
- c) finally, the combination of the two effects above makes more likely to have a negative indirect effect on firm 1's licensing revenues. Therefore, the indirect effect may partially or entirely compensate the direct one smoothing away the effect of a larger patent breadth on firm 1's profits.

3 Robustness

3.1 Competition between innovators

So far the analysis has been conducted by assuming that the first innovation is a research tool. In this section we show that, qualitatively, our main results hold also when the first innovation has a commercial value and the two innovations/firms compete in the same market.

A simple way to introducing competition in our model is the following. Let us assume that:

- when the bad state of the world occurs, firm 2 fails to obtain any innovation. In this case firm 2 gets a revenue equal to zero while firm 1 maintains its monopoly position, enjoying profits $\pi_1^m(q_1)$, where q_1 denotes the quality/characteristics of the first innovation;
- in the good state of the world, firm 2 obtains the innovation. We assume that in this case the two firms enjoy duopolistic profits, $\pi_1^d(q_1, q_2)$, and $\pi_2^d(q_1, q_2)$, where q_2 represents the quality/characteristics of the second innovation.

Consider first, the choice of R&D made by the follow-on innovator. Firm 2 chooses r to maximize:

$$\max_r r (\gamma(b)\pi_2^d(q_1, q_2) + (1 - \gamma(b))\rho\pi_2^d(q_1, q_2)) - c(r).$$

From the first order condition it follows that the optimal R&D activity, $r^*(b)$, satisfies:

$$\frac{\partial r^*}{\partial b} = \frac{(1 - \rho) \gamma'(b) \pi_2^d(q_1, q_2)}{c''(r)} < 0.$$

The above expression is the equivalent to (3) where $V^G = \pi_2^d(q_1, q_2)$, and $V^B = 0$. Note that due the assumption that in the bad state of the world the second innovator makes no profits, the infringement effect vanishes; this implies that $r^*(b)$ always decreases with the patent breadth. Let us now focus on firm 1; its profits are given by:

$$\Pi_1(b) = r^*(b) [\gamma(b)\pi_1^d(q_1, q_2) + (1 - \gamma(b)) (\pi_1^d(q_1, q_2) + (1 - \rho)\pi_2^d(q_1, q_2))] + (1 - r^*(b))\pi_1^m(q_1).$$

Differentiating this expression with respect to b it is possible to investigate the impact of a change in the patent breadth on the first innovator's profits; formally:

$$\frac{\partial \Pi_1(b)}{\partial b} = -r^*(b)\gamma'(b)(1-\rho)\pi_2^d(q_1, q_2) + \frac{\partial r^*}{\partial b} [\pi_1^d(q_1, q_2) + (1 - \gamma(b))(1 - \rho)\pi_2^d(q_1, q_2) - \pi_1^m(q_1)].$$

This expression can still be interpreted in terms of direct and indirect effects. The first term represents the direct effect and, as usual, it is positive. Having introduced competition

between the two innovators slightly changes the workings of the indirect effect. The change in the R&D activity of the follow-on innovator alters the probability that firm 1 maintains its monopoly position; this additional effect is represented by the term $-\frac{\partial r^*}{\partial b}\pi_1^m(q_1)$, which is clearly positive given that $\frac{\partial r^*}{\partial b} < 0$.

The following example shows that also in case of competition firm 1 can be harmed by a too strong patent protection.

Example 3

Consider the following extension of example 1 to the case of competing firms. As above we assume that $c(r) = \frac{r^2}{2} + r$, and $\gamma(b) = \beta(b) = \max\{0, 1 - b\}$. Moreover, we assume that there is a mass 1 of consumers with utility function $U(q, p) = q - p$, where q is the quality of the product and p the price. The quality of the good produced by firm 1 is q_1 and that of firm 2 is $q_2 = V^G + q_1$ in the good state of the world and 0 in the bad state. Finally, we assume that production costs are zero and that firms compete in prices; therefore, in the bad state firm 1's product is sold in the market at price q_1 , while in the good state firm 2's product is sold at price V^G .

Remark 8. *When $V^G \in \left(2 + q_1, \frac{1}{\rho}\right)$, firm 1 obtains larger profits when the patent has a limited breadth, namely when b is smaller than 1.*

Proof. From the proof of Remark 6 we have that firm 2 chooses $r^* = 0$ when $b = 1$ and $r^* = 1$ for any $b \in [0, \bar{b}]$, where $\bar{b} = \frac{V^G - 2}{V^G(1 - \rho)}$. Firm 1's profits are $V^G(1 - \rho)(1 - \beta(\bar{b}))$, when $b = \bar{b}$ and q_1 when $b = 1$. Therefore, firm 1 is better-off when $b = \bar{b}$ provided that $V^G > 2 + q_1$. Finally, note that the set $V^G \in \left(2 + q_1, \frac{1}{\rho}\right)$ is non-empty provided that $\rho < \frac{1}{2 + q_1}$. ■

3.2 Different determination of the licensing fee by the court

In Section 2 we have assumed that, in case of infringement, the court determines the licensing fee in proportion to the commercial value of the innovation. One may wonder, whether our results still apply when the court follows a different rule. First of all note that, by definition,

the court cannot implement a lump-sum licensing schedule, since it may impose a payment on the follow-on innovator only in case of infringement. Moreover, it is possible to show that whatever the rule used by the court, the result Proposition 1 generalises as follows: all licensing occurs at the fee implemented by the court. These facts imply that whatever the rule used by the court the investment of firm 2 is inefficient. Obviously, the kind and severity of the inefficiency depends on the specific rule actually endorsed by the court.

4 Conclusions

The definition of the optimal patent policy for industries where innovation is cumulative requires considering issues that go beyond the classic trade-off between incentives to R&D and dead-weight loss. When several inventors sequentially contribute to innovation, patents should also guarantee an appropriate division of profits among them. Early innovators should be rewarded for laying down the foundations of the industry but this should not go to the detriment of follow-on inventors who provide improvements or applications of the existing technologies to other fields. In this sense, there is the danger that a too strong patent protection, while encouraging early inventors, will hold future innovators up.

The existing theoretical contributions on the economics of cumulative innovation have stressed the merits of ex-ante arrangements; the risk that future innovations are held-up can be substantially mitigated in case the follow-on innovator negotiates the licensing agreement with the patent-holder before incurring the R&D costs. Notably, Green and Scotchmer (1995) show that, in a context of symmetric information, ex-ante licensing eliminates the risk of hold-up.

In this paper we present a model based on Green and Scotchmer (1995) and we show that when the first innovator cannot observe whether the follow-on inventor has already undertaken its R&D activities, the possibility of ex-ante licensing does not ensure efficiency. To the contrary, the investment of the follow-on innovator is always inefficient, no matter what the patent breadth is; at the equilibrium, both underinvestment and overinvestment

may occur depending on whether the commercial effect of the R&D activity prevails on the infringement effect or not.

Importantly, the inefficiency in the R&D investment of the follow-on inventor implies that a large patent breadth may be harmful for the first innovator too. In the paper we provide three very stylized examples whereby this occurs.

Contrary to the existing literature, we show that the optimal patent policy does not necessarily have to balance the opposing interests of early and follow-on inventors; a very large patent breadth, may harm both generations of innovators.

Our model rests on the assumption that both the amount of the R&D investment of the second generation inventor, and when such investment is undertaken are not observable by third parties. 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 difficult measurement.

References

- Bessen, J. (2004). Holdup and licensing of cumulative innovations with private information. *Economic Letters*, 82:321–326.
- Chang, H. (1995). Patent scope, antitrust policy, and cumulative innovation. *RAND Journal of Economics*, 26:34–57.
- 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.