DISCUSSION PAPER SERIES

DP13087

SOME SIMPLE ECONOMICS OF PATENT PROTECTION FOR COMPLEX TECHNOLOGIES

Vincenzo Denicolò and Piercarlo Zanchettin

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

SOME SIMPLE ECONOMICS OF PATENT PROTECTION FOR COMPLEX TECHNOLOGIES

Vincenzo Denicolò and Piercarlo Zanchettin

Discussion Paper DP13087
Published 30 July 2018
Submitted 30 July 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

SOME SIMPLE ECONOMICS OF PATENT PROTECTION FOR COMPLEX TECHNOLOGIES

Abstract

We analyze patent protection when innovative technologies are "complex" in that they involve sequential and complementary innovations. We argue that complexity affects the classic Nordhaus trade-off between innovation and static monopoly distortions. We parametrize the degree of sequentiality and that of complementarity and show that the optimal level of patent protection increases with both. We also address the issue of the optimal division of profit among different innovators.

JEL Classification: O30, O40

Keywords: Sequential innovation, Complementarity, Patent design, Elasticity of the supply of inventions, Division of profit

Vincenzo Denicolò - vincenzo.denicolo@unibo.it
University of Bologna and CEPR

Piercarlo Zanchettin - pz11@le.ac.uk
University of Leicester

Some Simple Economics of Patent Protection for Complex Technologies*

Vincenzo Denicolò
University of Bologna and CEPR

Piercarlo Zanchettin
University of Leicester

July 27th, 2018

Abstract

We analyze patent protection when innovative technologies are "complex" in that they involve sequential and complementary innovations. We argue that complexity affects the classic Nordhaus trade-off between innovation and static monopoly distortions. We parametrize the degree of sequentiality and that of complementarity and show that the optimal level of patent protection increases with both. We also address the issue of the optimal division of profit among different innovators.

Key words: sequential innovation, complementarity, patent design, elasticity of the supply of inventions, division of profit

JEL classification numbers: O30, O40

*We thank Hugo Hopenhyan, Matthew Mitchell, Volker Nocke, Konrad Stahl and seminar audiencies at Helsinki, Lisbon, Mannheim, Tilburg and Chicago for useful comments and discussion. The usual disclaimer applies. E-mail addresses: vincenzo.denicolo@unibo.it (Vincenzo Denicolò) and pz11@le.ac.uk (Piercarlo Zanchettin).

1 Introduction

Over the past few years, economists have become increasingly aware that traditional models of discrete, independent innovations may not represent well the reality of many industries. In the modern economy, many technologies are "complex," meaning that a single product or process may involve a number of different patents, each protecting a separate piece of innovative knowledge. In some cases, the distinct innovative components can be invented independently of each other. In others, basic discoveries open the way for subsequent improvements and applications. In short, innovation is often complementary and sequential. When this is so, the aggregate value of the inventions is greater than the sum of their stand-alone values.

Many scholars have argued that the complexity of the technology increases the social costs of patent protection. The main reason for this is that the fragmentation of intellectual property rights may create problems of coordination among the patent holders. Coordination failures may generate transaction costs (Heller and Eisenberg, 1998), pricing externalities (Lerner and Tirole, 2004), and greater scope for opportunistic behaviour (Farrell and Shapiro, 2008). In addition, it has been argued that patents may impede the sharing of intermediate technological knowledge among innovative firms. In some models, this effect may be so strong that patent protection may actually stifle technological progress rather than spurring it (Bessen and Maskin, 2009; Fershtman and Markowich, 2010). In the light of these potential issues, the conventional wisdom is that patent protection should be weakened as the technology becomes more complex.

Without questioning the relevance of these problems, this paper highlights a countervailing effect that seems to have gone unnoticed so far. This effect implies that, anything else equal, sequentiality and complementarity demand stronger, not weaker patent protection.

The new effect we highlight relates to a most basic trade-off in the analysis of optimal patent protection, namely, that between rapid innovation and monopoly deadweight losses. Since Nordhaus (1969), this trade-off has been at the centre of the analysis of optimal patent design for isolated innovations. The economic literature on sequential and complementary innovations, in contrast, has focused on the externalities arising among the different innovators, and on the division of profit among them. However, sequentiality and complementarity affect also the Nordhaus trade-off and hence may have an impact on its resolution. That is precisely the focus of the present paper.

With isolated innovations, the optimal resolution to the Nordhaus trade-off hinges on the elasticity of the supply of inventions. This elasticity is the percentage increase in the number of inventions (or, equivalently, in the probability of success x) associated with a one percent increase in R&D expenditure X. The intuitive reason why this elasticity matters is simple. Patents create incentives to invest in R&D (the uppercase X), but what policymakers really care about is technological progress (the lowercase x). The elasticity of the supply of inventions, $\varepsilon = \frac{dx}{dX}\frac{X}{x}$, measures precisely how effective additional R&D investment is in delivering more innovations. No wonder this must be a key determinant of the optimal level of patent protection. In simple models of isolated innovations, the relationship between the optimal level of protection and the elasticity is one of direct proportionality.

Consider now the case of multiple, related innovations. For example, think of two innovations, 1 and 2, that are strictly complementary, meaning that the stand-alone value of each is nil but the aggregate value of both is positive. In this case, research is effectively successful only if both innovations are achieved, so the relevant probability of success is $x_1 x_2$. As a consequence, the relevant elasticity is now the *sum* of the individual elasticities, $\varepsilon_1 + \varepsilon_2$. This implies that, anything else equal, the optimal level of patent protection is higher than in the discrete innovation case.

That is, in a nutshell, the main message of this paper. The rest of the paper formalizes the above argument and extends it in various ways. The effect we uncover relates to the technology of R&D and so is very robust. We demonstrate that it operates in many different set-ups and under quite diverse institutional conditions.

Although there is by now a vast literature on sequential and complementary innovation, to the best of our knowledge this paper provides the first thorough analysis of the impact of sequentiality and complementarity on the Nordhaus trade-off. The paper that comes closest to ours is the classic article by Green and Scotchmer (1995) on sequential innovation. This article analyzes how patent policy should jointly determine the total profit earned by the successive innovators and the division of the profit between them. The former depends on the protection accorded to the patent holders against imitators (*backward* protection), the latter on the protection accorded to the first inventor against the second (*forward* protection).

Green and Scotchmer however drastically simplify the Nordhaus trade-off, assuming that innovations can be achieved with probability one by sinking a fixed R&D investment. Since patents are distortionary, this implies that the optimal level of backward protection is simply the one that allows the innovator to just cover his R&D cost. With two-stage innovation, this simple principle implies that each innovator must cover his own cost. However, Green and Scotchmer argue that the policymaker cannot control the division of profit finely, implying that the division may not reflect precisely the firms' respective R&D costs. Specifically, if the first inventor covers his cost, in their model the second will inevitably obtain a positive rent. This has two implications. First, with sequential innovations backward protection should be stronger than with isolated innovations (or when a single firm can achieve both innovations), as more total profit is needed to cover both innovators' costs. Second, forward protection should transfer as much profit as possible from the second inventor to the first one.

We differ from Green and Scotchmer in that we posit a smooth relationship between R&D investment and innovation. As Green and Scotchmer recognize in their discussion of endogenous R&D investment (p. 31), this is necessary for a richer analysis of the Nordhaus trade-off. Furthermore, we assume that the division of profit can be fine tuned. This allows us to abstract from the effect highlighted by Green and Scotchmer and to focus on the new mechanism mentioned above.

The subsequent literature on sequential innovation has focused almost exclusively on forward protection. The literature has analyzed circumstances in which it is not necessarily desirable to favour the first inventor, and it has compared different ways in which forward protection can be provided, but it has neglected the impact of sequentiality on the level of backward protection.[1] Likewise, the literature on complementary innovations has focused on the division of profit but has not addressed the issue of the overall level of protection.[2]

The rest of the paper proceeds as follows. In Section 2, we present and analyze our baseline model of sequential innovation. We parametrize the degree of sequentiality and show that as sequentiality increases, both backward and forward protection should be strengthened. In Section 3, we show that our results are robust to several variations of the baseline model. Section 4 presents similar results for the case of complementary (but simultaneous) innovations. Section 5 summarizes the results and provides some final remarks. Proofs are collected in the Appendix.

2 Sequential innovation

In this and the following section, we analyze the case of sequential innovation. We start from the case of two innovations, 1 and 2. These innovations come in sequence in that innovation 2 cannot be searched for unless innovation 1 has already been achieved. In Section 3, we extend the analysis to, among others, the case of an infinite sequence of innovations.

[1] See, for instance, Scotchmer (1996), Matutes, Regibeau and Rockett (1996), O'Donoghue, Scotchmer and Thisse (1997), O'Donoghue (1998), Denicolò (2000), Hunt (2004), Hopenhayn, Llobet and Mitchell (2006), and Parra (2014).

[2] See, for instance, Shapiro (2007), Clark and Konrad (2008), Gilbert and Katz (2011) and Schmidt (2014).

2.1 Baseline model

To begin with, we follow Green and Scotchmer (1995) in assuming that innovators are specialized. That is, certain firms can invest only in innovation 1, and others only in innovation 2. (Among the extensions, we shall consider the case where a single firm can achieve both innovations.)

2.1.1 R&D technology

The probability that innovation i is achieved, denoted by x_i, is a function of the aggregate R&D expenditure targeted to that innovation, denoted by X_i:[3]

$$x_i = F_i(X_i). \tag{1}$$

The function $F_i(X_i)$ is the "innovation production function." It is assumed to be smooth, increasing and concave, with $F_i(0) = 0$. Concavity captures the assumption of decreasing returns to R&D. Returns may be diminishing because the production of innovative knowledge requires inputs which are in fixed supply, such as for instance talent, or the set of good ideas at any given point in time.[4]

The elasticity of the innovation production function, $\varepsilon_i \equiv \frac{F_i'(X_i)X_i}{x_i}$, will play a crucial role in the following analysis. The elasticity is the percentage increase in the probability of success associated with a one percent increase in R&D expenditure. It is positive, as the function F_i is increasing, and less than one, as F_i is concave with $F_i(0) = 0$. With a large number of potential inventions, the number of innovations achieved is proportional to the probability of success and thus ε_i may be interpreted as the elasticity of the supply of inventions.[5]

[3] Our framework here is static, but it may be viewed as the reduced form of a dynamic setting in which the timing of innovation is uncertain. In that case, the probabilities x_1 and x_2 could be interpreted as the discount factors corresponding to the time lag to innovation, as in Denicolò (2000).

[4] We assume that the first innovation is fully disclosed in any case, so patent protection does not affect the production function of the second innovation. This allows us to abstract from the effects analyzed in Bessen and Maskin (2009), who assume, in contrast, that patents impede technological disclosure.

[5] A large empirical literature has tried to estimate these elasticities. Estimates vary considerably from study to study, though: see Cohen (2010) for an excellent survey.

2.1.2 Product market

To keep the analysis as general as possible, we do not make any specific assumption on the nature of the innovations and the product market. Rather, we use a reduced-form model that is consistent with many different set-ups.

To abstract from the effects of intellectual property fragmentation mentioned in the introduction, we assume that patent holders can perfectly coordinate their behaviour. This rules out issues of business stealing, transaction costs, Cournot complements etc. These factors may be important and should be included in a more complete analysis, but our modeling strategy here is to avoid distracting complications so as to focus on the effect of interest.

In our simplified setting, the social value of the two innovations generally comprises three components. The first one is the total discounted profit accruing to the innovators under full patent protection. We normalize this to 1. The second component, denoted by $U \geq 0$, is the increase in consumer surplus brought by the innovations under full patent protection.[6] This is positive for drastic innovations, but vanishes when innovations are non-drastic.[7] The third component, denoted by $D > 0$, is the monopoly deadweight loss. This component of the value does not materialize under full patent protection, but it does when patent protection expires, or is somewhat limited. Thus, D captures the social costs of patent protection.

Therefore, with full patent protection the two innovations taken together increase social welfare by $1 + U$, of which 1 accrues to the innovators and U to consumers. With no patent protection, in contrast, the social benefit is $1 + U + D$ and is reaped entirely by consumers.

[6] More generally, U might include positive spillovers enjoyed by firms other than the innovators.

[7] Innovations are non-drastic when competition from outsiders prevents innovators from charging the monopoly price. If this is so, then the innovators must engage in limit pricing and thus the quality-adjusted price does not change until the patent expires.

2.1.3 Patent policy

In fact, patent protection can take on values intermediate between full protection and no protection at all. In our analysis, we shall treat it as a continuous policy variable, denoted by $\mu \in [0, 1]$. We assume that with intermediate levels of protection payoffs are a linear combination of those arising under full or no protection, with weights μ and $(1 - \mu)$, respectively. Thus, the joint discounted profit obtained by the two innovators is μ, monopoly deadweight losses are μD, and the surplus left to consumers is $(1 - \mu)(1 + D) + U$.

The variable μ captures various possible reasons why patent protection may be incomplete. Most obviously, the duration of the protection may be finite: for example, patent life is currently 20 years in most countries. However, even before a patent expires the protection accorded to the patent holder may be imperfect; that is, the breadth of protection may also be limited. Our variable μ encompasses both length and breadth.[8]

2.1.4 Sequentiality

It remains to detail the specific contributions to private and social payoffs of the two innovations. For simplicity, we take the parameters D and U, as well as the level of patent protection μ, to be the same for both innovations. We then denote by s the share of the payoffs due to the second innovation, and by $(1 - s)$ that of the first. Thus, s may be taken to be an index of sequentiality. When $s = 0$, all of the value is attached to the first innovation and so we are back to the case of discrete inventions. When instead $s = 1$, the first innovation has no direct utility. For example, it may be thought of as a pure research tool that is valuable only as it enables the search for the second innovation.[9]

Besides determining the total profit obtained by the two innovators, μ, patent

[8] See Gilbert and Shapiro (1990) for a classic analysis of breadth v. length. The assumption that payoffs are linear in μ fits particularly well the case where μ is interpreted as length.

[9] Sometimes, s is also referred to as the "option value" of the first innovation, as achieving the first innovation makes it possible to invest in the second one, which has an extra value of s.

policy determines also the division of the profit between them. Specifically, we assume that innovator 1 gets the entire profit from the first innovation, i.e. $\mu(1-s)$, plus a fraction λ of the profit from the second, μs. Innovator 2 gets the remaining share $(1-\lambda)$ of the profit from the second innovation. The policy variable $\lambda \in [0,1]$ may be interpreted as the level of *forward* patent protection, i.e. the protection accorded to the first inventor against the second.[10] In contrast, μ may be viewed as the level of *backward* protection, i.e. the protection against imitators.

2.1.5 Patent races

At this point, one can consider two variants of the baseline model. In the first one, there is monopoly in the search for each innovation. That is, there are two firms, 1 and 2: firm 1 only can invest in innovation 1, and firm 2 in innovation 2. Firms are risk neutral and maximize their expected profits:

$$\pi_1 = x_1\mu(1-s) + x_1x_2\mu\lambda s - X_1 \tag{2}$$

and

$$\pi_2 = x_1\left[x_2\mu(1-\lambda)s - X_2\right]. \tag{3}$$

In the second variant, for each innovation a number of firms race to be the first inventor. In this variant of the model, we assume free entry in both patent races. We further assume that the returns to R&D are decreasing at the industry level but not at the firm level. This implies that each individual firm's probability of getting the patent, conditional on the innovation being achieved, equals its share in the aggregate R&D investment in that innovation.[11]

[10] In practice, patent law affects the division of profit only indirectly, by determining what each patent holder is entitled to do and hence the disagreement point in the bargaining process. However, taking λ as a continuous variable simplifies the analysis and allows us to abstract from the difficulties of fine tuning the division of profit analyzed by Green and Scotchmer (1995).

[11] Griliches (1990) argues that it may indeed be realistic to assume that the returns to R&D are diminishing at the industry level but not at the firm level. Accordingly to him, at the firm level,

> in the major range of the data [...] there is little evidence for diminishing returns, at least in terms of patents per R&D dollar. That is not surprising, after all. If there were such diminishing returns, firms could split themselves into divisions or separate enterprises and escape them. (p. 1167)

With free entry, a zero-profit condition must then hold for each race. Thus, the R&D investment levels are determined by the condition that $\pi_1 = \pi_2 = 0$. In both cases, investment levels are chosen sequentially: investment in innovation 2 can be made only if innovation 1 has been successfully achieved.

Initially, we focus on the free-entry variant of the model. In the next section, we shall show that similar results can be obtained when there is monopoly in the search for each innovation.

2.2 Equilibrium

The zero-profit conditions are:

$$x_1\mu(1-s) + x_1x_2\mu\lambda s = X_1 \tag{4}$$

and

$$x_2\mu(1-\lambda)s = X_2. \tag{5}$$

These conditions determine the equilibrium R&D investments, denoted by X_i^*, and the corresponding probabilities of success, denoted by x_i^*. Notice that since firms invest in innovation 2 only if innovation 1 has already been achieved, X_2 does not depend on the probability x_1. The equilibrium is then found by proceeding backwards: condition (5) yields X_2^* and hence x_2^*, which may then be plugged into (4) to obtain X_1^*. The existence of a positive solution may be guaranteed by the standard Inada conditions $\lim_{X_i \to 0} F_i'(X_i) = \infty$. Concavity of $F_i(X_i)$ guarantees uniqueness. We assume interior solutions: $x_i^* < 1$.

This simple model delivers comparative statics results that are, for the most part, natural.

Lemma 1 *Both X_1^* and X_2^* increase with the level of backward patent protection μ. An increase in forward protection λ decreases investment in the second innovation*

Returns to R&D may be constant at the firm level and decreasing at the industry level because of the risk that two or more firms may duplicate the same innovation.

and, as long as $\lambda < 1 - \varepsilon_2$, increases investment in the first innovation.

Intuitively, an increase in μ increases the innovators' profits and hence their R&D efforts. In addition to this direct effect, the impact of μ on x_1^* involves also an indirect effect. That is, an increase in μ increases x_2^* and hence the probability that the first innovator gets the share λ of the second innovation's profits. Both the direct and the indirect effects are positive.

Turning to the forward protection variable λ, it is evident that an increase in λ harms the second innovator. What is perhaps surprising is that stronger forward protection may stifle the first innovation, too. This happens when $\lambda > 1 - \varepsilon_2$, and the intuition is as follows. If the first inventor gets a share of the common profit that is already large, a further increase in that share may reduce the investment in the second innovation by so much that the first inventor is actually harmed. Clearly, this result by itself implies that it will never be optimal to set λ above $1 - \varepsilon_2$, as in this region more forward protection would discourage both innovations.

2.3 Optimal policy

Let us now suppose that the policy variables μ and λ are chosen so as to maximize expected social welfare:

$$W = [1 + U + (1 - \mu)D]\,[(1 - s)x_1^* + s x_1^* x_2^*] - X_1^* - x_1^* X_2^*. \tag{6}$$

Using the zero-profit conditions (4) and (5), social welfare reduces to:

$$W = [(1 - \mu)(1 + D) + U]\,[(1 - s)x_1^* + s x_1^* x_2^*]. \tag{7}$$

That is, social welfare coincides with expected consumer surplus, as the profits from the innovations are entirely "dissipated" in the patent races.

The policymaker then maximizes (7), keeping in mind that x_1^* and x_2^* are given by (4) and (5). Consider the optimal level of backward protection first.

Lemma 2 *The optimal level of backward patent protection is implicitly given by the following condition:*

$$\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U} = \frac{\varepsilon_1}{1-\varepsilon_1} + S(1)\frac{\varepsilon_2}{1-\varepsilon_2} + S(\lambda)\frac{\varepsilon_1}{1-\varepsilon_1}\frac{\varepsilon_2}{1-\varepsilon_2}, \tag{8}$$

where

$$S(\lambda) \equiv \frac{\lambda s x_2^*}{(1-s)+\lambda s x_2^*}. \tag{9}$$

Expression (8) immediately implies that we always have $\mu^* > 0$. This is intuitive, as social welfare vanishes when $\mu = 0$ (which in our simple model implies zero R&D investments). If condition (8) delivers a value of μ greater than 1, the optimal policy is to provide full patent protection ($\mu^* = 1$).

The variable $S(\lambda)$ may be interpreted as an index of "effective" sequentiality. This index vanishes not only when $s = 0$, in which case the second innovation has no value, but also when $x_2^* = 0$. In this latter case, too, we are effectively back to the single innovation framework as the second innovation would be valuable but is never achieved. On the other hand, if all of the social value is associated with the second innovation, i.e. $s = 1$, then the index of effective sequentiality is equal to one. Furthermore, the index of effective sequentiality accounts for the fact that the first innovator obtains only a share λ of the profit from the second innovation.

Two further remarks are in order. First, since the left-hand side of (8) is increasing in μ, the solution μ^* is an increasing function of the right-hand side. Thus, it appears that the optimal level of backward protection is an increasing function of the elasticities of the innovation production functions. It also appears that the optimal level of protection depends on the degree of sequentiality; that dependence will be the focus of the comparative statics analysis in the next subsection. Second, backward and forward protection are related. Specifically, μ^* depends also on the level of forward protection λ, *via* the index of effective sequentiality $S(\lambda)$ that appears in the last term on the right-hand side of (8).

As for forward protection itself, we have:

Lemma 3 *The optimal level of forward patent protection is*

$$\lambda^* = (1 - \varepsilon_2) \frac{\frac{\varepsilon_1}{1-\varepsilon_1} - [1 - S(1)] \frac{\varepsilon_2}{1-\varepsilon_2}}{\frac{\varepsilon_1}{1-\varepsilon_1} + S(1)\varepsilon_2} \tag{10}$$

Expression (10) immediately implies that we always have $\lambda^* < 1-\varepsilon_2$, as we noted above. The level of forward protection is positive when

$$\frac{\varepsilon_1}{1 - \varepsilon_1} > [1 - S(1)] \frac{\varepsilon_2}{1 - \varepsilon_2}. \tag{11}$$

When this inequality is reversed, we have a corner solution $\lambda^* = 0$. That is, the optimal level of forward protection vanishes if the elasticity of the supply of the first innovation is lower than that of the second one and sequentiality is low.

2.4 Comparative statics

We are now ready to address the issue of how changes in the degree of sequentiality affect optimal patent policy. Consider first the optimal level of backward protection, μ^*. When $s = 0$, both $S(\lambda)$ and $S(1)$ vanish, and thus we are back to the case of isolated innovations, where

$$\mu^* = \frac{1 + U + D}{1 + D}\varepsilon_1. \tag{12}$$

That is, the optimal level of protection is proportional to the elasticity of the supply of inventions.[12]

When s is positive, the sequentiality index becomes positive. Thus, two extra positive terms appear on the right-hand side of (8). On the other hand, when s is positive X_1 is lower than when $s = 0$ (this follows from condition (4)). The decrease in X_1 may affect the elasticity ε_1. These observations immediately imply the following result:

Proposition 1 *When innovation is sequential, the optimal level of backward protection is higher than in the single innovation case, provided that the elasticity* ε_1 *is non-increasing in* X_1.

[12] This simple "elasticity rule" was first noted by Denicolò (2007).

To get some intuitive insight, remember that in the case of discrete innovations ($s = 0$) the optimal level of protection is simply proportional to the elasticity of the supply of inventions ε_1. When instead $s > 0$, a fraction of the value from the innovations is obtained with probability $x_1 x_2$. For this fraction, the relevant elasticity is $\varepsilon_1 + \varepsilon_2$. Since the elasticity is greater, the optimal level of protection is higher.

In particular, when the elasticities are identical ($\varepsilon_1 = \varepsilon_2 = \varepsilon$) and $U = 0$, neglecting the last term in (8) the optimality condition reduces to:[13]

$$\mu^* = [1 + S(1)]\varepsilon. \tag{13}$$

Expression $[1 + S(1)]\varepsilon$ may be viewed as the effective elasticity of the supply of inventions in the sequential case. For a fraction of the value, the elasticity is ε but for the other fraction it is 2ε. The term $[1 + S(1)]\varepsilon$ is weighted average of the two elasticities, with a weight that reflects the relative importance of the second innovation.

In the general case, the exact relationship between the elasticities and the optimal level of backward protection is more complicated. However, it remains true that the relevant elasticity is higher than in the case of isolated innovations.

In Proposition 1, the condition that ε_1 is non-increasing in X_1 guarantees that the change in ε_1 due to the decrease in X_1 is non negative. The condition is sufficient but not necessary. If ε_1 increases with X_1, the result is reversed only if this negative effect dominates the additional terms on the right-hand side of (8). While in general we cannot rule out this possibility, it should be noted that the extent to which the elasticity may be increasing is limited by the hypothesis of diminishing returns to R&D.[14]

When the elasticities are constant, as we henceforth assume, the relationship

[13] The last term vanishes when there is no forward protection ($\lambda = 0$) or when the elasticities are small enough that ε^2 is negligible.

[14] Most of the empirical literature specifies the innovation production function as log-linear and hence assumes that the elasticity is constant. The few studies that allow for a variable elasticity, however, suggest that the elasticity is decreasing (see e.g. Guo and Trivedi, 2002).

between μ^* and s is monotone increasing. To show this, we first suppose that the level of forward protection, λ is pre-determined. In this case, we have:

Proposition 2 *For any given level of forward protection λ, the optimal level of backward protection $\mu^*$is an increasing function of the degree of sequentiality s.*

However, the optimal level of backward protection depends on forward protection, as we have already noted. Since changes in the degree of sequentiality may affect λ^*, it may be interesting to analyze the impact of sequentiality on backward protection when the level of forward protection λ is optimally adjusted instead of being fixed exogenously. To this end, we must first characterize the direction of the change in λ^*.

Consider, then, the effect of the degree of sequentiality on the optimal level of forward protection, i.e., on the division of the profit among the first and the second innovator. We have:

Proposition 3 *The optimal degree of forward protection λ^* is a non decreasing function of the degree of sequentiality s.*

The intuition here is as follows. With sequential innovations, the first innovator exerts a positive externality on the second, the magnitude of which increases with s. Granting to the first inventor a share of the profits from the second innovation is a way to correct for the externality. Proposition 3 says that the correction should increase with the size of the externality.

We are now in a position to analyze how the optimal degree of backward protection varies with the degree of sequentiality when the level of forward protection λ is optimally adjusted instead of being fixed exogenously.

Proposition 4 *When the level of forward protection is chosen optimally, the optimal level of backward protection is an increasing function of the degree of sequentiality s.*

Taken together, Propositions 3 and 4 show that μ and λ are complementary policy tools: both should increase as the degree of sequentiality increases.

When the first innovation is relatively less elastic than the second ($\varepsilon_1 < \varepsilon_2$), we find a particularly interesting pattern. In this case, when s is close to 0 inequality (11) holds and thus the optimal division of profit entails no forward protection (i.e., $\lambda^* = 0$). As a consequence, as the degree of sequentiality s increases, it is initially optimal to increase only the level of backward protection. However, as s further increases, the optimal level of forward protection becomes positive. At the same time, backward protection should continue to increase. In other words, the first policy response to higher sequentiality should be to increase backward protection only. Forward protection should be used as an additional policy tool only for sufficiently high degrees of sequentiality.

3 Robustness

In this section, we demonstrate the robustness of our results by analyzing several extensions of the baseline model. To begin with, we explore different normalizations of the value of innovations. We then consider alternative assumptions on who can do the research. Finally, we discuss models with an infinite sequence of innovations.

3.1 Alternative normalizations

In the baseline model, we have assumed that changes in the degree of sequentiality s do not affect the aggregate *ex post* value of the two innovations, $1 + U + D$. However, for any given level of R&D investments the aggregate *ex ante* value decreases as s increases, as the second innovation is achieved with a lower probability than the first one. One may wonder that the impact of s on optimal patent policy found in Section 2 may depend on the implicit assumption that as s increases the innovations effectively become less valuable.

To show that this is not so, let us assume that the aggregate *ex post* value is

$v\,(1+U+D)$, of which a share proportional to $(v-s)$ is due to the first innovation and a share proportional to s to the second. As s changes, we then adjust v in such a way that the first-best social welfare

$$W^{FB} = \max_{X_1, X_2} \left\{ (1+U+D)\left[(v-s)x_1 + s x_1 x_2\right] - X_1 - x_1 X_2 \right\} \tag{14}$$

stays constant. In other words, we now keep constant the *ex ante* value of the two innovations, net of R&D costs.

By the envelope theorem, in order to keep W^{FB} constant when s increases, the parameter v must change as follows:

$$\frac{dv}{ds} = 1 - x_2. \tag{15}$$

Proceeding as in the proof of Lemma 2, it can be shown that with this new normalization, the optimal level of backward protection is given by an expression identical to (8) except that the sequentiality index $S(\lambda)$ is now replaced by:

$$\hat{S}(\lambda) \equiv \frac{\lambda s x_2^*}{(v-s) + \lambda s x_2^*}. \tag{16}$$

However, a glance at the proof of Propositions 2, 3 and 4 reveals that the only property of the sequentiality index which is used in the proofs is that the index is increasing in s. For the new index (16), we have:[15]

$$\frac{d\hat{S}(\lambda)}{ds} = \frac{\lambda s(v-s)\frac{dx_2^*}{ds} + \lambda x_2^*(v-s) + \lambda s x_2^{*2}}{\left[(v-s) + \lambda s x_2^*\right]^2} > 0, \tag{17}$$

which implies that all our results continue to hold.

3.2 Specialized research monopolists

Consider now the variant of the baseline model where only one firm (firm 1) can invest in innovation 1, and another firm (firm 2) in innovation 2. With monopoly in each innovation, the R&D investments are no longer determined by the zero-profit conditions (4)-(5). Rather, they are determined by the first-order conditions

[15] Note that $\frac{dx_2^*}{ds} = \frac{\varepsilon_2 x_2^*}{(1-\varepsilon_2)s} > 0$ remains the same as in the baseline model.

$$F_1'(X_1)\,[\mu(1-s)+x_2\mu\lambda s]=1 \tag{18}$$

and

$$F_2'(X_2)\mu(1-\lambda)s=1. \tag{19}$$

It is easy to check that the second-order conditions are always met.

With constant elasticities, the comparative statics results are exactly the same as under free entry. Using consumer surplus as a welfare criterion, the optimality conditions are also exactly the same as in the baseline model, i.e. (8) and (10).

Formulas become more complicated, but the qualitative conclusions do not change, when profits are included in the social welfare function. For example, with profits in the social welfare function and symmetric elasticities, social welfare (6) becomes

$$W=[\mu(1-\varepsilon)+U+(1-\mu)(1+D)]\,[(1-s)x_1^*+sx_1^*x_2^*]. \tag{20}$$

The optimal level of backward protection is then implicitly given by

$$\frac{\mu^*(\varepsilon+D)}{(1-\mu^*)\,(1+U+D)}=\frac{\varepsilon}{1-\varepsilon}\left[1+S(1)+S(\lambda)\frac{\varepsilon}{1-\varepsilon}\right]. \tag{21}$$

The formula is very similar to (8), and the solution has, indeed, the same qualitative properties. When the elasticities differ, formulas are more cumbersome but retain the same flavour.[16]

3.3 Complete monopoly in research

With sequential innovation, the first inventor exerts a positive dynamic externality on the second one. This tends to create underinvestment in R&D – the more so, the higher is sequentiality. One may wonder that the reason why sequentiality increases the optimal level of patent protection is the need to correct for this externatility.

[16] With variable elasticities, and going back to the social welfare function (7), the optimality condition obtained is identical to (8), except that the terms $\frac{\varepsilon_i}{1-\varepsilon_i}$ must be relaced by $\frac{\varepsilon_i}{1-\varepsilon_i(1-\eta_i)}$, where η_i is the elasticity of the elasticity ε_i. Once again, the effect uncovered by or analysis is modified but does not disappear.

To show that this is not the only reason, let us assume that there is only one firm that can invest in both innovations. In this case, all externalities among innovators are fully internalized. Of course, in this case the division of profit is irrelevant, so we are left with only one policy tool, namely, the level of backward protection μ.

With monopoly in research, the expected profit of the monopolistic innovator is:

$$\pi = x_1\mu(1-s) - X_1 + x_1\left[x_2\mu s - X_2\right]. \tag{22}$$

The first-order conditions for a maximum are

$$F_1'(X_1)\left[\mu(1-s) + x_2\mu s - X_2\right] = 1 \tag{23}$$

and

$$F_2'(X_2)\mu s = 1. \tag{24}$$

One can easily check that the second-order conditions hold.

It is easy to confirm that the optimal R&D investments are still increasing in the level of backward protection.[17] The optimal level of protection is now implicitly given by

$$\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U} = \frac{\varepsilon_1}{1-\varepsilon_1}\left(1+\tilde{S}\right) + S(1)\frac{\varepsilon_2}{1-\varepsilon_2} \tag{25}$$

where the new sequentiality index is

$$\tilde{S} \equiv \frac{sx_2^*\varepsilon_2}{(1-s)+sx_2^*(1-\varepsilon_2)} \tag{26}$$

As before, all that matters for our results is that $\tilde{S}$ increases with s. Indeed, we have

$$\frac{d\tilde{S}}{ds} = \varepsilon_2\frac{1+s\frac{dx_2^*}{ds}\left[(1-s)+sx_2^*(1-\varepsilon_2)\right]}{\left[(1-s)+sx_2^*(1-\varepsilon_2)\right]^2} > 0. \tag{27}$$

[17] To be precise, the elasticities of the probabilities of success with respect to the level of patent protection are

$$\frac{dx_2^*}{d\mu}\frac{\mu}{x_2^*} = \frac{\varepsilon_2}{1-\varepsilon_2},$$

just as in the baseline model case, and

$$\frac{dx_1^*}{d\mu}\frac{\mu}{x_1^*} = \frac{\varepsilon_1}{1-\varepsilon_1}\left(1+\tilde{S}\right)$$

where $\tilde{S}$ is given by (26).

This implies that even with complete monopoly in research, the optimal level of patent protection μ^* increases with the degree of sequentiality s. This clarifies that the reason why more sequentiality demands stronger patent protection is not to remedy the dynamic externality, but that the effective elasticity of the supply of invention increases. This is a "technological" effect and as such it must be at work under the most diverse institutional conditions.

3.4 Infinite sequence of innovations

The baseline model considers two innovations that come in sequence, as in the pioneering paper of Green and Scotchmer (1995). The main theoretical alternative to two-stage models are models with an infinite sequence of innovations, as in O'Donoghue, Scotchmer and Thisse (1997) and Hunt (2004). In this subsection, we show that our insights extend also to this framework.

For reasons of tractability, we assume stationarity, as nearly all models with an infinite sequence of innovations do. In each period t, innovative firms make an aggregate investment X_t and obtain innovation t with probability $x_t = F(X_t)$. The innovation production function is assumed to be time invariant to guarantee stationarity. For the same reason, we assume that the profits generated by the innovations are stationary. Specifically, we now normalize to 1 the perpetual flow of profit generated by each innovation under full patent protection. With patent protection set at level μ, the flow of profit is then μ. For simplicity, we assume that successive innovators do not compete with each other, so past innovators suffer no profit erosion as new innovations arrive.

However, future values are discounted at rate r, so the present value of the perpetual flow of profits is $\frac{\mu}{r}$. The lower is r, the greater is the value of future innovations relative to the current one. Therefore, in models with an infinite sequence of innovations it seems natural to take the interest rate r as an inverse measure of the degree of sequentiality.

In this framework, forward protection can be captured by assuming that innovator t is entitled to a share λ of the profits from innovation $t+1$. By stationarity, however, innovator t must in turn leave to innovator $t-1$ a share λ of the profits from innovation t. Because of a front-loading of profits effect (see Segal and Whinston, 2007), it is easy to see that the optimal level of forward protection is $\lambda^* = 0$.

With $\lambda = 0$, the expected profit of inventor t is:

$$\pi_t = x_t \frac{\mu}{r} - X_t. \tag{28}$$

Using either the zero-profit condition $\pi_t = 0$ (with free entry), or the first-order condition for a maximum

$$F'(X_t)\frac{\mu}{r} = 1 \tag{29}$$

(with monopoly in research), one can immediately derive the following quite intuitive comparative statics results

$$\frac{dx^*}{d\mu}\frac{\mu}{x^*} = \frac{\varepsilon}{1-\varepsilon} > 0, \tag{30}$$

and

$$\frac{dx^*}{dr}\frac{r}{x^*} = \frac{\varepsilon}{\varepsilon - 1} < 0, \tag{31}$$

where we have dropped the time index as there is no risk of confusion.

3.4.1 Pseudo sequentiality

Some models with an infinite sequence of innovations assume that failure in period t does not preclude or impede subsequent innovations. If this is so, however, then innovation is not really sequential, as pointed out by Bessen and Maskin (2009). According to Bessen and Maskin, sequentiality does not mean simply that innovation $t+1$ comes after innovation t, but also that innovation $t+1$ is enabled, or at least facilitated, by innovation t. They convincingly argue that a proper model of sequential innovation must therefore posit that if innovation fails in one period, then with some positive probability the entire innovation process stops forever.

We shall consider a truly sequential model presently, but for now let us focus on what may be called the pseudo-sequential case, where the innovative process never stops in any case. In this case, expected consumer surplus may be written as follows:

$$W = x\frac{(1-\mu)(1+D)+U}{r} + \frac{W}{(1+r)} \tag{32}$$

Notice that the second term on the right-hand side (i.e., the continuation value) is not multiplied by the probability of success in the current period, reflecting the assumption that future innovation is not precluded by current failure. From (32), social welfare may be rewritten as

$$W = \frac{1+r}{r^2}x\left[(1-\mu)(1+D)+U\right]. \tag{33}$$

The optimal level of patent protection is then given by

$$\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U} = \frac{\varepsilon}{1-\varepsilon}, \tag{34}$$

exactly as in the case of discrete innovations.

In a pseudo-sequential model, then, the optimal level of protection does not depend directly on the degree of sequentiality r.[18] This is not surprising, as in such a model innovations are connected to each other only temporally but not functionally.

3.4.2 True sequentiality

Following Bessen and Maskin (2009), consider now the truly sequential case. To save on notation, we assume that the industry can continue to innovate in period $t+1$ only if innovation t has been achieved. (The analysis is similar if one assumes that failure in period t does not stop the innovative process altogether, but only with some positive probability).

In this case, social welfare is

$$W = x\frac{(1-\mu)(1+D)+U}{r} + x\frac{W}{(1+r)}, \tag{35}$$

[18] If the elasticity of the supply of inventions is not constant, however, the optimal level of protection can depend on r indirectly. In particular, since a decrease in r increases X, if the elasticity ε decreases (increases) with X then an increase in the degree of sequentiality implies a lower (higher) level of protection.

which rearranging terms may be rewritten as

$$W = x\frac{(1-\mu)(1+D)+U}{r(1+r-x)}. \tag{36}$$

The optimal level of backward protection is now implicitly given by the condition

$$\frac{\mu^*(1+D)}{(1-\mu^*)\,(1+D)+U} = S(r)\frac{\varepsilon}{1-\varepsilon} \tag{37}$$

where

$$S(r) \equiv \frac{1+r}{1+r-x^*} \tag{38}$$

is the index of effective sequentiality for this model. The index is indeed a decreasing function of r for any given x. However, an increase in r has a negative impact on x^*. Even accounting for this indirect effect, though, we have

$$\frac{dS}{dr} = x\frac{(1+r)\frac{\varepsilon}{\varepsilon-1}-r}{(1+r-x)^2} < 0. \tag{39}$$

meaning that S is a decreasing function of r (and hence an increasing function of the degree of sequentiality). This immediately implies that the optimal level of patent protection is an increasing function of the degree of sequentiality also in models with an infinite sequence of innovations.

4 Complementary innovations

In this section, we consider the case in which innovations are complementary (i.e., the aggregate value of the two innovations is greater than the sum of their stand-alone values) but can be achieved independently of each other. This implies that R&D investments may be simultaneous rather then sequential.[19]

[19] Even if firms can invest simultaneously, it may be efficient to target one innovation after the other in a pre-specified order. The reason for this is that when complementarity is particularly strong, this strategy reduces the risk of investing in R&D projects that, even if successful, turn out to be of little value because the complementary innovation is not achieved. This however creates a war of attrition, as investing earlier is riskier than investing later. The equilibrium may therefore be one where firms invest sequentially even if they need not to, as in Biagi and Denicolò (2014). Simultaneous investments may nevertheless arise when complementarity is less pronounced, or when research takes time and so firms must invest before knowing whether the complementary R&D project has succeeded or not.

4.1 Model assumptions

Like in the sequential case, we normalize to 1 the aggregate profit for the two innovators under full patent protection. Now we assume that a share c of this profit is obtained only if both innovations are achieved. Of the remaining share $1-c$, a fraction β_1 is the stand-alone value of innovation 1 and the complementary fraction $\beta_2 = 1-\beta_1$ that innovation 2. Thus, the parameter c is our measure of the degree of complementarity. When $c = 1$, innovations are strictly complementary; when instead $c = 0$, innovations are independent.

Like in the case of sequential innovation, patent policy may affect the division of the profit between the innovators. We assume that each innovator gets the full stand-alone profit (i.e., $\beta_1(1-c)\mu$ for innovator 1 and $\beta_2(1-c)\mu$ for innovator 2). Furthermore, innovator 1 gets a share λ_1 of the "common" profit $c\mu$, and innovator 2 the remaining share $\lambda_2 = 1-\lambda_1$. The innovators' expected profits therefore are

$$\pi_i = \mu\beta_i(1-c)x_i + \lambda_i \mu c x_i x_j - X_i \tag{40}$$

4.2 Equilibrium

Like in the case of sequential innovation, we may assume either free entry or monopoly in each innovation.[20] With free entry, the R&D investment levels are determined by the zero-profit conditions $\pi_1 = \pi_2 = 0$. Under monopoly, in contrast, they are determined by the first-order conditions

$$F'(X_i)\left[\mu\beta_i(1-c) + \lambda_i \mu c x_j\right] = 1. \tag{41}$$

For any given patent policy (μ, λ_i), these equilibrium conditions determine the level of R&D investments, X_i^*, and hence the equilibrium probabilities of success, x_i^*. Both variants of the model deliver the same comparative statics as well as the same optimal policy (provided that profits are not included in the social welfare criterion).

[20] One can also consider the case of complete monopoly in research, as in Section 3.3 above.

4.3 Comparative statics

We now characterize the impact of patent policy on the equilibrium R&D investments. Consider the level of backward protection first. By implicit differentiation of the system of the equilibrium conditions we have:[21]

$$\frac{dx_i^*}{d\mu}\frac{\mu}{x_i^*} = \varepsilon_i \frac{1-\varepsilon_j\left[1-C_j(\lambda_i)\right]}{(1-\varepsilon_i)(1-\varepsilon_j)-\varepsilon_i\varepsilon_j C_i(\lambda_j)C_j(\lambda_i)}, \tag{42}$$

where

$$C_i(\lambda_j) = \frac{c\lambda_j x_i^*}{\beta_j(1-c)+c\lambda_j x_i^*} \tag{43}$$

may be regarded as indexes of effective complementarity, analogously to the indexes of effective sequentiality encountered in sections 2 and 3. Like c, $C_i(\lambda_j)$ ranges from 0 (when $c=0$) to 1 (when $c=1$). However, differently from c (which is a purely exogenous measure of complementarity), the variables $C_i(\lambda_j)$ reflect also the equilibrium level of R&D investments.

Assuming that

$$(1-\varepsilon_i)(1-\varepsilon_j)-\varepsilon_i\varepsilon_j C_i(\lambda_j)C_j(\lambda_i) > 0, \tag{44}$$

it appears that stronger backward protection increases R&D investment. Condition (44) may be viewed as a "stability" condition that ensures that firms get not trapped in a zero-investment equilibrium in which investment in innovation i vanishes for fear that innovation j may not be achieved, and vice versa. The condition is always satisfied if the degree of complementarity is not too high.[22]

As for the division of profit, we have

$$\frac{dx_i^*}{d\lambda_i}\frac{\lambda_i}{x_i^*} = \frac{(1-\varepsilon_j)\,C_j(\lambda_i)-\frac{\lambda_i}{\lambda_j}C_i(\lambda_j)C_j(\lambda_i)\varepsilon_j}{(1-\varepsilon_i)\,(1-\varepsilon_j)-C_i(\lambda_j)C_j(\lambda_i)\varepsilon_i\varepsilon_j}\varepsilon_i. \tag{45}$$

As long as λ_i is not too large, shifting profit from innovator j to innovator i increases x_i^* and decreases x_j^*. When λ_i is already very large (and thus λ_j very low), however, a

[21] Under free entry, equation (42) holds even if the elasticities are not constant. With monopoly in research, in contrast, when the elasticities are variable the formula should be modified as shown in footnote 16 above.

[22] When condition (44) fails, positive R&D investments will be nonetheless obtained in the equilibrium with sequential R&D investments discussed in footnote 19 above.

further increase in λ_i may stifle both innovations. This effect is similar to the effect of increasing forward protection with sequential innovation, and it has the same intuitive explanation.

4.4 Optimal policy

Let us not turn to the optimal policy. We take social welfare to be the expected consumer surplus from the innovations:[23]

$$W = [(1-\mu)(1+D)+U]\,[\beta_1(1-c)x_1^* + \beta_2(1-c)x_2^* + cx_1^*x_2^*]. \tag{46}$$

4.4.1 Backward protection

For backward protection, we have (denoting $C_i \equiv C_i(1)$ to simplify notation):

Lemma 4 *With either free-entry or monopoly in research, the optimal level of backward protection is implicitly given by the following condition*

$$\frac{\mu^*(1+D)}{(1-\mu^*)\,(1+D)+U} = \sum_{i=1}^{2} \frac{C_i}{C_i + C_j - C_iC_j}\frac{dx_i^*}{d\mu}\frac{\mu}{x_i^*}. \tag{47}$$

Comparing the case where $c > 0$ to the case $c = 0$ we get

Proposition 5 *When innovation is strictly complementary ($c = 1$), the optimal level of backward protection is higher than in the case of independent innovations ($c = 0$).*

The intuition is similar to the intuition for Proposition 1. With independent innovations, the level of protection is determined by a weighted average of the elasticities ε_1 and ε_2, with weights that reflect the relative importance of the two innovations. With strictly complementary innovations, in contrast, the level of protection is determined by the sum of the elasticities.

Proposition 5 contrasts the extreme cases $c = 1$ and $c = 0$. The analysis of intermediate cases is more complicated as there are two possible channels through which

[23] With monopoly in research, expression (46) holds as long as one does not include profits in the social welfare calculation. Inclusion of profits complicates the formulas but does not change the main results, as in the case of sequential innovations.

c may affect the optimal level of patent protection. Firstly, a change in c affects the indexes of complementarity C_i directly. For any given level of R&D investment, the effect is positive, meaning that more complementarity implies that stronger patent protection is desirable. However, a change in c may also affect the equilibrium levels of R&D. To be precise, the equilibrium R&D investments decrease as the degree of complementarity increases. The reason for this is that when innovations are complementary, firms investing in innovation 1 exert a positive externality on innovator 2, and vice versa. This positive externality implies that the market equilibrium must be characterized by underinvestment. The stronger is the complementarity among innovations, the stronger will be the externality and hence the more serious the underinvestment problem. Thus, an increase in the degree of complementarity reduces x_i^*, and this in turn affects the indexes of effective complementarity, C_i. Thus, we have two, possibly opposing effects at work here.

To make some progress, we focus on the case in which innovations are symmetric. Therefore, we assume that $\beta_i = \frac{1}{2}$ and $F_1(.) = F_2(.)$, which implies $\varepsilon_i = \varepsilon$. This symmetry obviously implies that it is optimal to split the common profit μc evenly, so we can directly set $\lambda_i = \frac{1}{2}$. This in turn implies that R&D investments will be symmetric: $x_i^* = x^*$.

In this case, condition (47) simplifies to

$$\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U} = \frac{\varepsilon(1+C)}{1-\varepsilon(1+C)}, \tag{48}$$

where

$$C \equiv \frac{cx^*}{(1-c)+cx^*}. \tag{49}$$

For this case, we have

Proposition 6 *For symmetric complementary innovations, the optimal level of patent protection is an increasing function of the degree of complementarity, c.*

4.4.2 Division of profit

Finally, we briefly turn to the issue of the division of profit. We have:

Proposition 7 *With either free-entry or monopoly in research, the optimal division of profit is implicitly given by*

$$\lambda_i = \frac{\varepsilon_i - \varepsilon_j + \varepsilon_j C_j}{\varepsilon_i C_i + \varepsilon_j C_j}. \tag{50}$$

It should be noted that equation (50) determines the division of profit only implicitly, as the indexes of complementarity depend on the R&D investments which in turn depend on λ_i. At any rate, for given R&D investments condition (50) implies that patent policy should favour the innovation with the higher elasticity of supply.[24]

However, the R&D investments are in fact endogenous. Even taking the elasticities as constant, the R&D investments depend on the βs and also on multiplicative shifters in the innovation production functions that do not affect the elasticities. As for these factors, the general rule seems to be that the division of profit should be biased in favour of the "weaker" innovation, i.e. the innovation which is less valuable, or more difficult to achieve.

For example, suppose that $\varepsilon_i = \frac{1}{2}$, a case for which closed form solutions are available. If both innovations have the same stand-alone value, i.e. $\beta_i = \frac{1}{2}$, but innovation 2 is twice more costly to achieve than innovation 1 (i.e., $x_i = \phi_i \sqrt{X_i}$ with $\phi_2 = 2\phi_1$), then direct calculation shows that

$$\lambda_1^* = \frac{2(2+c)^2}{24 + 16c + 3c^2}.$$

If instead $\phi_2 = \phi_1$but $\beta_1 = 2\beta_2$, we have

$$\lambda_1^* = \frac{8(1+c)^2}{72 + 32c + 9c^2}.$$

[24] It may be interesting to contrast (50) with the rule derived by Shapiro (2007), which is

$$\left(\frac{\lambda_i}{\lambda_j}\right)^2 = \frac{\frac{dx_i^*}{d\lambda_i}\frac{\lambda_i}{x_i^*}}{\frac{dx_j^*}{d\lambda_j}\frac{\lambda_j}{x_j^*}}.$$

Shapiro's rule is obtained by maximizing the aggregate expected profit rather than social welfare and is more opaque than (50) but has the same general flavour.

In both cases, $\lambda_1^* < \frac{1}{2}$, confirming that the weaker innovator should obtain a share of the common profit greater than the stronger one.

5 Conclusion

In this paper, we have highlighted an important effect that arises when innovation is sequential or complementary. In these cases, the effective elasticity of the supply of inventions is higher than in the case of isolated innovations. Since the elasticity is a key determinant of the optimal level of protection, this implies that, anything else equal, sequentiality and complementarity demand stronger patent protection. The effect we have uncovered is robust as it refers to the technology of R&D. As such, it does not depend on the fine details of the model and operates in many different circumstances.

The previous literature has pointed to various reasons why sequentiality and complementarity may increase the social costs of patent protection. This has created a conventional wisdom that patents should be weakened as the technology becomes more complex. Our analysis does not call into question that literature but suggests more caution in drawing general conclusions about the impact of technological complexity on the optimal level of patent protection.

References

[1] Bessen, J. and E. Maskin (2009). Sequential innovations, patents, and imitation, *RAND Journal of Economics*, 40, 611-635.

[2] Biagi, A., and V. Denicolò (2014). Timing of discovery and the division of profit with complementary innovations, *Journal of Economics & Management Strategy*, 23, 89-102.

[3] Clark, D. and K. Konrad (2008). Fragmented property rights and incentives for R&D, *Management Science*, 54, 969-981.

[4] Cohen, W. M. (2010). Fifty years of empirical studies of innovative activity and performance, in *Handbook of the Economics of Innovation*, Vol. 1, Amsterdam, North-Holland, 129-213.

[5] Denicolò, V. (2000). Two-stage patent races and patent policy, *RAND Journal of Economics*, 31, 488-501.

[6] Denicolò, V. (2007). Do patents overcompensate innovators?, *Economic Policy*, 22, 680-729.

[7] Green, J. and S. Scotchmer (1995). On the division of profit with sequential innovation, *RAND Journal of Economics*, 26, 20-33.

[8] Farrell, J. and C. Shapiro (2008). How strong are weak patents?, *American Economic Review*, 98, 1347-1369.

[9] Fershtman, C. and S. Markovich (2010). Patents, imitation and licensing in an asymmetric dynamic R&D race, *International Journal of Industrial Organization*, 28, 113-126.

[10] Gilbert, R. and M. Katz (2011). Efficient division of profits from complementary innovations, *International Journal of Industrial Organization*, 29, 443-454.

[11] Gilbert, R. and C. Shapiro (1990). Optimal patent breadth and length, *RAND Journal of Economics*, 21, 106-112.

[12] Griliches, Z. (1990). Patent statistics as economic indicators: a survey, *Journal of Economic Literature*, 28, 1661–707.

[13] Guo, J.Q. and P.K. Trivedi (2002). Flexible parametric models for long-tailed patent count distributions, *Oxford Bulletin of Economics and Statistics*, 64, 63–82.

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

[15] Hopenhayn, H., Llobet, G. and M. Mitchell (2006). Rewarding sequential innovators: Prizes, patents, and buyouts, *Journal of Political Economy*, 114, 1041-1068.

[16] Hunt, R. (2004). Patentability, industry structure, and innovation, *Journal of Industrial Economics*, 52, 401-425.

[17] Lerner, J. and J. Tirole (2004). Efficient patent pools, *American Economic Review*, 94, 691-711.

[18] Matutes, C., Regibeau, P., and K. Rockett (1996). Optimal patent design and the diffusion of innovations, *RAND Journal of Economics*, 27, 60-83.

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

[20] O'Donoghue, T. (1998). A patentability requirement for sequential innovation, *RAND Journal of Economics*, 29, 654-679.

[21] O'Donoghue, T., Scotchmer, S., and J. F. Thisse (1998). Patent breadth, patent life, and the pace of technological progress. *Journal of Economics & Management Strategy*, 7, 1-32.

[22] Parra, A. (2014). Sequential innovation and patent policy, mimeo, Northwestern University.

[23] Scotchmer, S. (1996). Protecting early innovators: should second-generation products be patentable? *Rand Journal of Economics*, 27, 322-331.

[24] Schmidt, K. (2014), Complementary patents and market structure, *Journal of Economics & Management Strategy*, 23, 68-88.

[25] Segal, I., and M. Whinston,(2007). Antitrust in innovative industries, *American Economic Review*, 97, 1703-1730.

[26] Shapiro, C. (2007), Patent reform: Aligning reward and contribution, *Innovation Policy and the Economy*, 8, 111-156.

Appendix

Proofs omitted in the text follow.

Proof of Lemma 1. To demonstrate the Lemma, we first implicitly differentiate (5) w.r.t μ to obtain:

$$\frac{dx_2^*}{d\mu}\frac{\mu}{x_2^*} = \frac{\varepsilon_2}{1-\varepsilon_2} > 0. \tag{A1}$$

As for the impact of μ on x_1^*, differentiating (4) and (5) we have

$$\frac{dx_1^*}{d\mu}\frac{\mu}{x_1^*} = \frac{\varepsilon_1}{1-\varepsilon_1} + S(\lambda)\frac{\varepsilon_1}{1-\varepsilon_1}\frac{\varepsilon_2}{1-\varepsilon_2} > 0 \tag{A2}$$

The two terms on the right-had side correspond to the direct and indirect effect described in the main text, respectively.

Turning to λ, for the second innovation we have:

$$\frac{dx_2^*}{d\lambda}\frac{(1-\lambda)}{x_2^*} = -\frac{\varepsilon_2}{1-\varepsilon_2} < 0. \tag{A3}$$

Clearly, stronger forward protection discourages the second innovation. The elasticity is the same as that with respect to the level of backward protection.

As for the first innovation, we have

$$\frac{dx_1^*}{d\lambda}\frac{\lambda}{x_1^*} = S(\lambda)\frac{\varepsilon_1}{1-\varepsilon_1}\frac{1-\varepsilon_2-\lambda}{(1-\lambda)(1-\varepsilon_2)} \tag{A4}$$

This is positive if and only if $\lambda < 1-\varepsilon_2$. ■

Proof of Lemma 2. To derive condition (8), it is convenient to rewrite (7) as

$$W = V(\mu) \times \Xi(\mu,\lambda), \tag{A5}$$

where

$$V(\mu) \equiv [1 + U + (1-\mu)D]$$

depends directly on μ, and

$$\Xi(\mu,\lambda) \equiv [(1-s)x_1^* + s x_1^* x_2^*]$$

depends on μ and λ only indirectly, through x_1^* and x_2^*. Assuming an interior solution, the first order condition can be written as:

$$-\frac{dV}{d\mu}\frac{\mu}{V} = \frac{d\Xi}{d\mu}\frac{\mu}{\Xi}.$$

We then calculate:

$$\frac{dV}{d\mu}\frac{\mu}{V} = -\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U}$$

and

$$\frac{d\Xi}{d\mu}\frac{\mu}{\Xi} = \frac{dx_1^*}{d\mu}\frac{\mu}{x_1^*} + S(1)\frac{dx_2^*}{d\mu}\frac{\mu}{x_2^*}.$$

Using (A1) and (A2), condition (8) follows immediately. ■

Proof of Lemma 3. Proceeding as in the proof of Lemma 2, notice that λ enters (A5) only through the second term, $x(\mu, \lambda)$. Assuming an interior solution, the first order condition can be written as:

$$\frac{dx}{d\lambda}\frac{\lambda}{x} = 0.$$

We then calculate:

$$\frac{dx}{d\lambda}\frac{\lambda}{x} = \frac{dx_1^*}{d\lambda}\frac{\lambda}{x_1^*} + S(1)\frac{dx_2^*}{d\lambda}\frac{1-\lambda}{x_2^*}\frac{\lambda}{1-\lambda}.$$

Using (A3) and (A4), condition (10) of the Lemma follows easily. ■

Proof of Proposition 2. The left hand side of (8) is increasing in μ^*. Therefore, to prove that μ^* increases with s we just need to prove that the right-hand side is increasing in s. To this end, it suffices to prove that $S(\lambda)$ is increasing in s for any $\lambda \in [0, 1]$. We calculate:

$$\frac{dS(\lambda)}{ds} = \frac{\lambda s(1-s)\frac{dx_2^*}{ds} + \lambda x_2^*}{[(1-s)+\lambda s x_2^*]^2}.$$

By implicit differentiation of (5) we get

$$\frac{dx_2^*}{ds}\frac{s}{x_2^*} = \frac{\varepsilon_2}{1-\varepsilon_2},$$

and plugging this into the preceding expression we finally have:

$$\frac{dS(\lambda)}{ds} = \lambda \frac{\frac{\varepsilon_2}{1-\varepsilon_2}(1-s)+1}{[(1-s)+\lambda s x_2^*]^2} x_2^* \geq 0$$

where the inequality is strict for $\lambda > 0$. ■

Proof of Proposition 3. When λ^* is given by an interior solution, we have

$$\frac{d\lambda^*}{dS(1)} = \frac{\varepsilon_2^2}{(1-\varepsilon_1)\left[\frac{\varepsilon_1}{1-\varepsilon_1} + S(1)\varepsilon_2\right]^2} > 0,$$

which means that λ^* is an increasing function of the sequentiality index $S(1)$. But we already know from the proof of Proposition 1 that $S(1)$ increases with the degree of sequentiality. This implies that when the solution is interior, the optimal degree of forward protection strictly increases with the degree of sequentiality. In turn, this implies that the solution λ^* is, overall, weakly increasing in s. ■

Proof of Proposition 4. When inequality (11) is reversed, the optimal level of forward protection λ^* is nil and hence does not depend on s. The result then follows directly from Proposition 1. Consider then the case in which inequality (11) holds and so we have an interior solution for λ^*. We know from Proposition 3 that in this case λ^* increases with s. To proceed, then, we must establish how changes in λ affect μ^*. We have:

Lemma 5 *If*

$$\frac{\varepsilon_1}{1-\varepsilon_1} < [1-S(1)]^2 \frac{\varepsilon_2}{1-\varepsilon_2} \tag{A6}$$

then the optimal level of backward protection μ^* decreases with the level of forward protection λ. If instead inequality (A6) is reversed, then for any $0 < s < 1$ there exists a critical value of λ , $\widehat{\lambda} \in (0,1)$, which is implicitly given by

$$\widehat{\lambda} = 1 - \varepsilon_2 - \left[\frac{(1-s)+\lambda s x_2^*}{(1-s)+s x_2^*}\right]^2 \frac{\varepsilon_2}{\frac{\varepsilon_1}{1-\varepsilon_1}} \tag{A7}$$

such that the optimal level of backward protection μ^* increases with the level of forward protection λ for $\lambda < \widehat{\lambda}$, while it decreases for $\lambda > \widehat{\lambda}$.

Proof. Since the left-hand side of (8) increases with μ, $\frac{d\mu^*}{d\lambda}$ has the same sign as the derivative of the right-hand side of (8). That is:

$$\frac{d\mu^*}{d\lambda} \propto \frac{dS(\lambda)}{d\lambda}\frac{\varepsilon_1}{1-\varepsilon_1}\frac{\varepsilon_2}{1-\varepsilon_2} + \frac{dS(1)}{d\lambda}\frac{\varepsilon_2}{1-\varepsilon_2},$$

where the symbol $\propto$ means "has the same sign as." Next, notice that

$$\frac{dS(\lambda)}{d\lambda} = \frac{s(1-s)}{[(1-s)+s\lambda x_2^*]^2}\left(x_2^* + \lambda\frac{dx_2^*}{d\lambda}\right)$$

and

$$\frac{dS(1)}{d\lambda} = \frac{s(1-s)}{[(1-s)+sx_2^*]^2}\frac{dx_2^*}{d\lambda}.$$

Substituting into the preceding expression, we finally get

$$\frac{d\mu^*}{d\lambda} \propto s(1-s)x_2^*\frac{\varepsilon_2}{1-\varepsilon_2}\frac{1}{[(1-s)+\lambda s x_2^*]^2}\times$$

$$\times\left\{\frac{\varepsilon_1}{1-\varepsilon_1}\left(1-\frac{\lambda}{1-\lambda}\frac{\varepsilon_2}{1-\varepsilon_2}\right) - \frac{[(1-s)+\lambda s x_2^*]^2}{[(1-s)+sx_2^*]^2}\frac{1}{(1-\lambda)}\frac{\varepsilon_2}{1-\varepsilon_2}\right\}$$

and therefore $\frac{d\mu^*}{d\lambda}$ has the same sign as the term inside curly brackets. It is immediate to verify that this term is decreasing in λ and vanishes when condition (A7) holds. At $\lambda = 0$, the derivative has the same sign as

$$\frac{\varepsilon_1}{1-\varepsilon_1} - \frac{(1-s)^2}{[(1-s)+sx_2^*]^2}\frac{\varepsilon_2}{1-\varepsilon_2}$$

or, equivalently, as

$$\frac{\varepsilon_1}{1-\varepsilon_1} - [1-S(1)]^2\frac{\varepsilon_2}{1-\varepsilon_2}.$$

Summarizing, if condition (11) is reversed, then the derivative is always negative. If instead inequality (113) holds, then the derivative is positive for $\lambda < \widehat{\lambda}$ and is negative for $\lambda > \widehat{\lambda}$. ■

It follows from Lemma 5 that a sufficient condition for the proposition to be true is that $\lambda^* < \widehat{\lambda}$. To show that the inequality $\lambda^* \le \widehat{\lambda}$ indeed holds true, notice that the first-order condition (10) can equivalently be rewritten as

$$\left[\frac{(1-\lambda^*)}{\varepsilon_2} - 1\right]\frac{\varepsilon_1}{1-\varepsilon_1} = \frac{(1-s)+\lambda s x_2^*}{(1-s)+sx_2^*}$$

whereas condition (A7) may be rewritten as

$$\left[\frac{\left(1-\widehat{\lambda}\right)}{\varepsilon_2}-1\right]\frac{\varepsilon_1}{1-\varepsilon_1}=\left[\frac{(1-s)+\lambda s x_2^*}{(1-s)+s x_2^*}\right]^2.$$

Since

$$\frac{(1-s)+\lambda^* s x_2^*}{(1-s)+s x_2^*}\leq 1,$$

it is clear that $(1-\lambda^*)\geq(1-\widehat{\lambda})$ and hence $\lambda^*\leq\widehat{\lambda}$. This completes the proof of the Proposition. ■

Proof of Lemma 4. We proceed as in the proof of Lemma 2. Condition (A5) holds also in the complementary case, but now

$$\Xi(\mu,\lambda)\equiv\beta_1(1-c)x_1^*+\beta_2(1-c)x_2^*+c x_1^* x_2^*.$$

Assuming an interior solution, the first order condition is still:

$$-\frac{dV}{d\mu}\frac{\mu}{V}=\frac{d\Xi}{d\mu}\frac{\mu}{\Xi},$$

where

$$\frac{dV}{d\mu}\frac{\mu}{V}=-\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U},$$

as before, and now

$$\frac{d\Xi}{d\mu}\frac{\mu}{\Xi}=\sum_{i=1}^{2}\frac{C_i}{C_i+C_j-C_iC_j}\frac{dx_i^*}{d\mu}\frac{\mu}{x_i^*}.$$

Condition (47) then follows immediately.■

Proof of Proposition 5. When $c=0$, condition (47) reduces to

$$\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U}=\sum_{i=1}^{2}\beta_i\frac{\varepsilon_i}{1-\varepsilon_i}. \tag{A8}$$

The right-hand side is a weighted average of the elasticity terms $\frac{\varepsilon_i}{1-\varepsilon_i}$ for the two innovations, with weights corresponding to their respective stand-alone values.

When $c=1$, in contrast, (47) becomes

$$\frac{\mu^*(1+D)}{(1-\mu^*)(1+D)+U}=\sum_{i=1}^{2}\frac{\varepsilon_i}{1-\varepsilon_i-\varepsilon_j}. \tag{A9}$$

It is immediate to verify that the right-hand side of (A9) is greater than that of (A8), whence the proposition follows. ■

Proof of Proposition 6. Equation (49) implies that the sign of $\frac{d\mu^*}{dc}$ is equal to the sign of $\frac{dC}{dc}$. We calculate:

$$\frac{dC}{dc} = \frac{x^* + \frac{dx^*}{dc} c\,(1-c)}{\left[(1-c) + cx^*\right]^2}.$$

Thus, the sign of $\frac{dC}{dc}$ equals the sign of the numerator:

$$N(x^*, c) \equiv x^* + \frac{dx^*}{dc} c\,(1-c)\,.$$

The derivative $\frac{dx^*}{dc}$ can be found by implicitly differentiation of (41), which gives

$$\frac{dx^*}{dc} = -\frac{\varepsilon}{1-\varepsilon(1+C)} \frac{x^*(1-x^*)}{(1-c)+cx^*} < 0.$$

Notice that $\frac{dx^*}{dc} < 0$, which makes perfect sense but complicates a bit the proof that $N(x^*, s)$ is positive.

To show that $N(x^*(c), c)$ is indeed positive, notice that $N(c) \equiv N(x^*(c), c)$ is continuous on its domain $[0, 1]$. Furthermore, we have $N(0) = x^*(0)$ and $N(1) = x^*(1)$. These are strictly positive if the equilibrium R&D investment is positive. Thus, for $N(c)$ to ever be negative there must exist one value of $c \in [0, 1]$ such that $N(c) = 0$ and $\frac{dN(c)}{dc} < 0$.

We prove below that this can never be true. Specifically, we show that at any c such that $N(c) = 0$, it must be $\frac{dN(c)}{dc} \geq 0$ (with strict inequality whenever $x^*(c) > 0$).

To show this, we first differentiate $N(c)$ with respect to c:

$$\frac{dN(c)}{dc} = 2(1-c)\frac{dx^*}{dc} + c(1-c)\frac{d^2x^*}{dc^2}.$$

The second derivative is:

$$\begin{aligned} \frac{d^2x^*}{dc^2} \quad = \quad & \frac{\varepsilon}{\left[1-\varepsilon(1+C)\right]\left[(1-c)+cx^*\right]^2}\left\{\left[(1-c)(1-2x^*) - cx^{*2}\right]\frac{dx^*}{dc} - x^*(1-x^*)^2\right\} + \\ & + \frac{x^*(1-x^*)}{(1-c)+cx^*}\frac{\varepsilon^2}{\left[1-\varepsilon(1+C)\right]^2}\frac{dC}{dc} \end{aligned}$$

Now, at any c such that $N(x^*(c), c) = 0$, we have $\frac{dC}{dc} = 0$ and hence:

$$\frac{dx^*}{dc} = -\frac{x^*}{c(1-c)}.$$

Furthermore, notice that when $N(x^*(c), c) = 0$ it must be:

$$\frac{\varepsilon}{1-\varepsilon(1+C)} = \frac{(1-c)+cx^*}{c(1-c)(1-x^*)}.$$

Therefore, at any c such that $N(x^*(c), c) = 0$, we have:

$$\frac{d^2x^*}{dc^2} = \frac{x^*\left[(1-c)^2(1-2x^*) - (2c-c^2)x^{*2}\right]}{c^2(1-c)^2(1-x^*)\left[(1-c)+cx^*\right]}$$

From (49), it follows that at any c such that $N(x^*(c), c) = 0$ it must be:

$$\frac{dN(c)}{dc} = \frac{x^*\left[(1-c)+cx^*\right]}{c\,(1-c)\,(1-x^*)} \geq 0.$$

This implies that $N(c)$ can never be negative for $c \in [0,1]$, completing the proof of the proposition. ■

Proof of Proposition 7. The first-order condition for a maximum is

$$\frac{d\Xi}{d\lambda_i}\frac{\lambda_i}{\Xi} = 0,$$

which may be rewritten as

$$\sum_{i=1}^{2} \frac{C_i}{C_i + C_j - C_iC_j}\frac{dx_i^*}{d\lambda_i}\frac{\lambda_i}{x_i^*} = 0,$$

or

$$C_1\frac{dx_1^*}{d\lambda_1}\frac{\lambda_1}{x_1^*} = -C_2\frac{dx_2^*}{d\lambda_2}\frac{\lambda_2}{x_2^*}.$$

Using (45), this condition becomes

$$\varepsilon_1(1-\varepsilon_2) - \lambda C_1\varepsilon_1 = \varepsilon_2(1-\varepsilon_1) - (1-\lambda)C_2\varepsilon_2.$$

Expression (50) then follows easily. ■