

Dynamic Protection of Innovations through Patents and Trade Secrets*

Paul Belleflamme[†]

Université catholique de Louvain

Francis Bloch[‡]

Paris I, Paris School of Economics

September 19, 2014

Abstract

This paper analyzes the optimal protection strategy for an innovator of a complex innovation who faces the risk of imitation by a competitor. We suppose that the innovation can be continuously fragmented into sub-innovations. We characterize the optimal mix of patent and trade secrets when the innovator faces a strict novelty requirement and can only patent a fraction of the innovation once. We also study the optimal dynamic patenting policy in a soft novelty regime, when the innovator can successively patent different fragments of the process. We compare a regime with prior user rights, when the innovator can use the secret part of the process, even when it is patented by an imitator with a regime without prior user rights.

Keywords: Patents, Trade Secrets, Dynamic Protection of Innovation, Intellectual Property Rights

JEL-Classification: O31, O34

*We are grateful to Emeric Henry, Gaston Llanes, Gerard Llobet, Yann Ménière, Mathieu Parenti and Francisco Ruiz-Aliseda for discussions. The research was started when Paul Belleflamme visited the Department of Economics at Ecole Polytechnique.

[†]CORE and Louvain School of Management, Université catholique de Louvain, 34 Voie du Roman Pays, B-1348 Louvain la Neuve, Belgium, Paul.Belleflamme@uclouvain.be; also affiliated with CESifo.

[‡]Department of Economics, Université Paris I and Paris School of Economics, 106-112 Boulevard de l'Hopital, 75647 Paris CEDEX 13, France, francis.bloch@univ-paris1.fr; also affiliated with CORE.

1 Introduction

Many product and process innovations can be considered as “complex” insofar as they can be fragmented into a set of sub-innovations. Examples of complex innovations are multi-stage production processes (where each stage can be patented independently), production processes requiring the use of different complementary components (where each component can be patented individually), or processes combining the use of a specific ingredient with a production method (where both the characteristics of the ingredient and the production method can be patented).

Although the possibility exists to patent each sub-innovation, innovators may decide to keep some or all of them secret. As Anton *et al.* (2006, p. 9) state: “Because innovations are rarely composed of a monolithic piece of knowledge, a combination of patenting and secrecy is common.” Combinations of that sort are documented in a number of industries. Arora (1997) describes how firms in the organic chemical industry resorted to both patenting and secrecy to protect their innovations. Jorda (2007) gives the examples of the artificial manufacture of diamonds for industrial use (GE patented much of the technology for making these diamonds but “also kept distinct inventions and developments secret”), and of Premarin, a hormone-therapy drug (Wyeth owned patents on the manufacturing process but also held a number of related trade secrets). The latter example is no exception: the survey results of Cohen *et al.* (2000, Tables 1 and 2) show that firms in the pharmaceutical industry declare relying heavily on both secrets and patents. Perng Pan and Mion (2010) describe the strategy of Coskata, a producer of biofuel, that “has several pending patent applications on the bioreactor segment of the process”, while “[t]he identity of the micro-organism fed into the bioreactor is protected by trade secret”; it is further explained that “this does not rule out the possibility of a patent on the biological component”. Examples also abound in the food industry where recipes, lists of ingredients or formula are kept secret, while cooking, manufacturing or packaging processes are patented.¹ Another indication that complex innovations are

¹*Kentucky Fried Chicken* holds secret the recipe of 11 herbs and spices that go into its fried chicken, and owns, e.g., a patent on “Device and method for frying and grilling” (EP 1648235 A1). *McDonald* keeps secret the Big Mac special sauce, and has patented “Method and apparatus for making a sandwich” (WO 2006068865 A3). It is well known that *Coca-Cola*’s syrup formula is a trade secret, but the company also owns patents on

quite common is that the average number of US patents per innovation is larger than 5 (Levêque and Ménière, 2007).

Objective. As illustrated by these examples, inventors of complex innovations face a rich set of strategies when it comes to protect their intellectual property. They may indeed choose between patenting and secrecy for each fragment of their innovation, which theoretically opens up a large number of combinations.² However, the patent regime that is in force may restrict these possibilities. For instance, a strict utility requirement may prohibit the patenting of fragments of innovations, or a strict novelty requirement may prevent inventors from patenting long held trade secrets.³

The objective of this paper is to provide a systematic study of the protection of complex innovations. From a positive point of view, we want to analyze the innovator’s choice of patent/secrecy mix under various patent regimes. From a normative point of view, our aim is to conduct a welfare comparison of these various patent regimes. To this end, we build the following model. The starting point is the discovery of a complex innovation by an inventor. As this innovation can be fragmented into sub-innovations, the inventor has to choose which fragments to patent and which fragments to keep secret. For simplicity, we consider that the innovation fragments are interchangeable and symmetric, so that the inventor’s decision amounts

“Coffee cola beverage composition” (EP 1736063 B1). To be fair, recipes are often kept secret because it is hard to patent them (in the U.S and in Europe alike). Yet, as Tarazano (2013) explains, it is not impossible to obtain a patent for a food recipe and there is thus room for a choice.

²In accordance with empirical surveys (see, e.g., Levin *et al.*, 1987, or Cohen *et al.*, 2000), we take patents and secrecy as the two main instruments of, respectively, formal and informal protection of intellectual property (IP). As reported by Hall *et al.* (2014), other forms of formal IP are copyright, trademarks and designs, while alternative informal appropriation mechanisms are lead time and complexity.

³A first requirement for patentability is that the invention be of practical use. A second requirement is that the invention show an element of novelty; that is, it must show some new characteristic that is not known in what is called the “prior art”, i.e., the body of existing knowledge in the technical field of the claimed invention (see, e.g., www.wipo.int/patentscope/en/patents/). According to Quinn (2012), a claimed invention can fail the utility requirement if the applicant fails “to disclose enough information about the invention to make its utility immediately apparent to those familiar with the technological field of the invention.” As far as novelty is concerned, whether and/or when trade secrets are considered as prior art varies across countries and through time (see the discussion in Section 4).

to choose the (continuous) fraction of the innovation that is protected. We assume that the innovator faces a single imitator. That is, there is only firm that has a sufficient absorptive capacity to appropriate the innovation and, thereby, compete with the inventor.⁴ In particular, the imitator is able, with some probability, to discover or circumvent any part of the innovation that has been kept secret; the imitator is also able to exploit any patented part of the innovation as soon as the patent expires.

We consider and compare four different regimes of patent protection. First, the *binary patent regime* corresponds to a strict utility requirement, which leads the patent office to reject patent applications that only concern a fragment of the innovation; as a result, the innovator is left with a binary decision: seek patent protection for the entire innovation or for nothing. The second regime, called the *single patent regime*, relaxes the utility requirement by allowing patents on fragments, but imposes a strict novelty requirement, insofar as the innovator cannot introduce more than a single patent for the innovation, which covers either a part or the full innovation; the patent office would reject any attempt to patent another fragment of the same innovation as the second fragment would not be deemed as novel enough with respect to the first one. Under these two regimes, the innovator faces a static optimization problem as his decision is to choose, once and for all, which parts of the complex innovation to patent and to keep secret.

When the novelty requirement is softer, the innovator's maximization problem becomes dynamic as it is now possible to patent at a later date parts of the innovation that were previously kept secret. As we discuss it in Section 4, the new Patent Reform Act in the U.S. (known as the *Leahy-Smith American Invents Act*) implements such a softening of the novelty requirement. It is then a so-called *sequential patent regime* that prevails, where the innovator chooses both the fraction of the innovation to patent in the first place, and the time at which a second patent is introduced.⁵ Two different sequential patent regimes are considered according to whether the innovator is granted *prior user rights* or not, i.e., whether or not the inno-

⁴According to Cohen and Levinthal (1990), the absorptive capacity is defined as a firm's ability to recognize the value of new information, assimilate it, and apply it to commercial ends.

⁵In our baseline model, we simplify the analysis by assuming that the innovator can only introduce two successive patents and that the second patent must cover the remainder of the innovation. We relax these assumptions in Section 6.

vator is allowed to continue using parts of the innovation that are patented by the imitator.

One feature that proves crucial for the analysis of the four patent regimes is the sensitivity of the innovator's profit to the fraction of the innovation that the imitator has access to. Quite naturally, we assume that the innovator's profit decreases in this fraction (as the more of the innovation the imitator can exploit, the stronger the competition he exerts). What really matters is whether the innovator's profit function is *concave* or *convex* in the fraction that the imitator can exploit. Intuitively, the innovator's profit function is *concave* if the benefit that the competitor obtains from the fraction he has access to is convex, i.e., if the imitator must learn a large fragment of the innovation in order to exploit it; in that case, we will talk of *increasing harm from imitation*. Conversely, the profit function is *convex* if the imitator benefits even from learning small fractions of the innovation, and the marginal benefit of learning a larger part of the innovative process is decreasing (*decreasing harm from imitation*). The curvature of the innovator's profit function naturally depends on the degree of complementarity among the sub-innovations constituting the complex innovation: the more complementary they are, the less useful a single sub-innovation is for the imitator and the more the harm from imitation tends to increase (concavity) rather than to decrease (convexity) as the imitator acquires a larger share of the innovation.

Results. The main results of our analysis are the following. First, in the static optimization problem (corresponding to a strict novelty requirement), we show that the innovator's optimal conduct may involve mixing patenting and secrecy; for this to happen, the innovator's profit function must be concave (i.e., the competitor needs to attain a critical share of the innovation to be able to exploit it); the optimal mix contains then more patents when the innovation is easier to reverse engineer, when the patent length is longer and when the discount rate is higher. Otherwise, when the profit function is convex (meaning that the competitor benefits from small increments in the innovation), the innovator optimally makes a binary choice between patenting the entire innovation (if secrets are relatively easy to leak) or keeping it entirely secret (otherwise). Hence, if the innovator's profit function is convex in the fraction of the innovation that the imitator can exploit, softening

the utility requirement by allowing patents on fragments of innovations has no effect whatsoever.

Second, when the novelty requirement is softer, the innovator dynamically chooses which fraction of the innovation to patent and when to patent the remaining fraction. Under prior user rights, strategies are again quite simple when the innovator's profit is convex in the imitator's fraction: the optimal dynamic patenting strategy is identical to the static strategy, meaning that the innovator never introduces a second patent. However, if the profit function is concave, the innovator may exploit the additional degree of freedom of the soft novelty regime and optimally choose to patent sequentially two fragments of the innovation. This situation happens when the probability of reverse engineering the innovation is intermediate. We note that the region of parameters for which the innovator chooses to fragment the innovation is larger in the dynamic patenting regime than in the static regime; intuitively, when the innovator is allowed to file a second patent on the innovation, he has an incentive to decrease the share of the innovation which is patented the first time. Finally, we show that the main qualitative results of the analysis remain unchanged if we endogenize the imitation effort, or if we allow a second patent that does not necessarily cover the remaining secret part of the innovation.

When the inventor does not hold prior user rights, the characterization of the optimal dynamic patenting strategy becomes more challenging. In that case, the innovator is no longer allowed to use the part of the innovation patented by the imitator; this clearly reduces the innovator's profit (as he is forced by the imitator to downgrade his product or resort to a less efficient production process). In this case, it is no longer possible to ascertain the concavity or convexity of the innovator's present discounted flow of profits in the fraction that he initially chooses to patent. The reason is the following: in the absence of prior user rights, a change in this fraction not only affects the fraction of the innovation that the imitator can exploit, but also the fraction of the innovation that the inventor can still exploit if the imitator patents the secret part. As a result, we cannot fully characterize the dynamic patenting strategy. A specific example of a process innovation suggests, however, that the innovator may choose to patent the entire innovation even when successful reverse engineering is relatively unlikely; this is because the absence of prior user rights raises the cost of losing one's trade

secrets.

Finally, we compare the four regimes of patent protection from the point of view of the three stakeholders: the innovator, the imitator and consumers. The analysis is done in the context of two illustrating examples (the innovation may either reduce production costs or result in a quality upgrade). Not surprisingly, in the two examples, the innovator's profits increase as regimes become more flexible; that is, the dynamic regime with prior user rights dominates the single patent regime, which dominates the binary regime. As for the other stakeholders, their preferences depend on the type of innovation. In the case of a process innovation that reduces the cost of production, the welfare of the imitator and of consumers are aligned and in complete conflict with the profit of the innovator: they all favor regimes with less flexibility, putting the binary regime on top. In contrast, in the case of a product innovation that enhances quality, the imitator and the consumers may also prefer the more flexible regimes. In fact, they both rank the static fragmentation regime above the binary regime, preferring to let the inventor segment the innovation in different pieces; the competitor also always prefers the dynamic regime to the static regime, whereas consumers prefer the dynamic regime except for intermediate values of the ease with which the imitator can discover what is kept secret. The analysis of this second example thus shows that the conflict in welfare between innovator, imitator and consumers does not necessarily arise and that there exist circumstances where all three types of agents prefer more flexible intellectual property rights regimes.

Regarding the effects of granting prior user rights, we observe (in the case of the cost-reducing innovation) that the imitator clearly gains when the innovator cannot hold prior user rights. However, consumers lose as this implies that, after the competitor discovers the trade secret, the innovator is not able to exploit the cost reducing innovation, yielding lower quantities in equilibrium.

Related literature. While there exists a large empirical literature on the choice between patenting and secrecy at the firm level, the theoretical literature remains rather scarce. Only a few studies consider secrecy as a viable option to patenting, and most of them regard these two modes of protection as mutually exclusive: the choice is between patenting or keeping secret

the whole innovation, which corresponds to what we call the “binary patent regime”.⁶ While this problem is relevant for “discrete” product industries where a single patent is enough to protect an invention, it does not suit “complex” product industries as the ones described above.

The only two exceptions that we are aware of are Ottoz and Cugno (2008) and (2011), where the possibility of mixing patents and secrets for complex innovations is explicitly considered.⁷ The starting point of Ottoz and Cugno (2008) is the same as ours: an innovator has discovered all the fragments of a complex innovation and uses them directly; his problem is to choose the optimal mix between patents and secrets. The authors study solely the static optimization problem (what we call the “single patent regime”) but they use a different framework than ours. The main difference comes from assuming that, to enter the market, an imitator must not only get access to the secret but also circumvent the patent. The effect of competition is then binary: the innovator is either kicked out of the market (if the imitator manages to enter) or remains a monopoly (otherwise). In contrast, we allow for a much larger set of competitive configurations as the imitator may enter by exploiting only a part of the innovation, while the innovator stays on the market (but faces a lower profit). Another important difference in Ottoz and Cugno’s model is that by choosing to patent a larger fraction of the innovation (and thus to disclose more information), the innovator decreases the probability that trade secrets leak out (since there is less knowledge that must leak) but increases the probability that an imitator invents around the patented part (since more knowledge has been disclosed). These opposite effects (which are absent in our analysis) drive Ottoz and Cugno’s main result, namely that the innovator’s optimum may be to patent some fraction

⁶See Hall *et al.* (2014) for a recent and comprehensive survey of the empirical and theoretical literature. It is instructive to note that early empirical studies carried out in the 1990s conclude that firms prefer secrecy over patenting to protect their innovations; yet, patents have grown more popular over the last two decades for a number of reasons (see Hussinger, 2006, and the references in Hall *et al.*, 2014).

⁷In the U.S. patent system prior 1999, it was possible to maintain patents unpublished until grant. This allowed innovators to combine secrecy and patenting *even for discrete innovations*. They could do so by using a “continuation application” (i.e., an application that is based on and has the same priority date as an original patent application). Before 1999, this procedure indeed combined secrecy (as the knowledge in the pending patent application was not disclosed) with patent protection (as a priority date was issued). Graham (2004) and Graham and Hedge (2012) study the strategic opportunity offered by this possibility.

of the innovation while keeping the remainder secret. This paper therefore reaches a conclusion similar to ours but through a different route.

Ottoz and Cugno (2011) takes a more normative point of view by trying to determine the socially optimal scope of trade secret protection when it can be combined with patent protection for the same innovation. In particular, they find conditions under which a broad scope of trade secret law may be socially beneficial. A broader protection of trade secrets has the twofold advantage of increasing the incentives to innovate and of decreasing the incentives to duplicate parts of existing innovations that are kept secret; more innovation and lower duplication costs may then, under certain conditions, compensate for higher R&D costs and reduced competition.

The papers that regard patents and secrets as mutually exclusive have mainly focused on the optimal patent design when secret is an option (e.g., Friedman *et al.*, 1991; Gallini, 1992; Takalo, 1998) and on the strategic disclosure of secrets (e.g., Horstman, MacDonald, and Slivinski, 1985; Anton and Yao, 2004). The first line of research is more relevant for our work. In particular, the following two papers present similarities with our framework and bring useful insights. First, Denicolo and Franzoni (2004) study the relative merits of secrecy and patents from an innovator's and from society's point of view; like us, they also compare different regimes regarding prior user rights. In contrast with our setting, their model incorporates an innovation stage (the innovator invests in R&D while the imitator invests in duplication efforts); they also consider patent length as a policy variable that affects the relative attractiveness of patents and secrets for the innovator (whereas in our model, it is the probability of leakage that plays this role as we keep patent length fixed). Their main conclusion is that if the patent life is set optimally and if the initial innovator relied on secrecy, a successful imitator should be allowed to patent and no prior user rights should be granted to the innovator. Second, Kultti, Takalo and Toikka (2007) consider a situation in which several firms innovate independently and can choose between patenting and secrecy. They show that innovators may prefer patenting to secrecy even when patents offer a lower protection than secrets; this is because innovators are concerned by the threat that their secret could be discovered and patented by another firm. The same intuition applies in our dynamic setting, in particular in the absence of prior user rights.

Finally, two other strands of the economic literature on patents bear some connections with our study. First, a number of papers, starting with Scotchmer and Green (1990), consider situations where follow-up innovations are built on several complementary basic innovations. In this context, the choice between patenting and secrecy does not affect the product market (as in our setting) but the research industry. For instance, Schneider (2008) studies a patent race model where firms choose between patenting and secrecy; he shows that an innovator relies on secrecy when he has a higher speed of discovery of a subsequent invention than his competitor. More recently, Kwon (2012a) considers a patent portfolio race where firms compete for complementary patents; he shows that there exists an equilibrium where firms patent some innovations and keep other innovations secret. Kwon (2012b) makes a similar point in the case of a patent race for a single innovation; here, firms may randomize between secrecy and patenting. The second strand of literature is concerned with complementary pieces of knowledge that are owned not by the same innovator (as in this paper) but by different firms. Although secrecy is most often not considered as an alternative form of protection in these papers, they provide useful insights about issues of great relevance that are not addressed here, such as patent thickets, strategic patenting and licensing (e.g., Shapiro, 2001), or patent pooling and standard-setting (e.g., Lerner and Tirole, 2004).

The rest of the paper is organized as follows. In Section 2, we describe our baseline model. In Section 3, we analyze the static optimization problem where the innovator chooses which fraction of the innovation to patent under a strict novelty requirement. In Section 4, we study the innovator's dynamic patenting strategy under the sequential patent regimes with and without prior user rights. In Section 5, we perform welfare comparisons of the four patent regimes by using two specific examples. Finally, we test the robustness of our results by extending the model in two directions in Section 6 and we offer some concluding remarks in Section 7.

2 The model

2.1 Imitation and protection strategies

We consider the protection strategy of an inventor who has discovered a complex innovation. As illustrated in the introduction, the complex innovation

can be fragmented into sub-innovations, and each fragment can be patented in its own rights. We normalize the entire innovation to 1, and denote by $\sigma \in [0, 1]$ the (continuous) fraction of the innovation that is protected.

The innovator faces a single imitator, which is the only firm that has the capacity to compete with the inventor. We let (σ_1, σ_2) denote the fractions of the innovation that the innovator and imitator have access to respectively. Any couple (σ_1, σ_2) results in profits for the innovator and imitator denoted by $P(\sigma_1, \sigma_2)$ and $Q(\sigma_1, \sigma_2)$ respectively. Throughout the paper, we assume that $P(\cdot, \cdot)$ is increasing in σ_1 and decreasing in σ_2 , and $Q(\cdot, \cdot)$ is increasing in σ_2 . It will be of particular interest to check whether P is concave or convex in σ_2 . Intuitively, P is concave in σ_2 if the benefit that the competitor obtains from σ_2 is convex, i.e., the imitator must learn a large fragment of the innovation in order to exploit it. Conversely, P is convex in σ_2 if the imitator benefits even from learning small fractions of the innovation, and the marginal benefit of learning a larger part of the innovative process is decreasing.

If a fraction of the innovation is covered by a patent, the patenting firm has exclusive access to this fragment during the lifetime of the patent, and the fragment becomes accessible to both firms after the expiry of the patent.⁸ Fractions that are not covered by a patent are trade secrets, which may be discovered by the imitator according to a Poisson process with exogenous parameter $\lambda > 0$. The parameter λ measures the ease with which the imitator can discover or circumvent the part of the process covered by trade secret. The length of the patent is denoted T and both firms discount the future at the same rate $r > 0$.

2.2 Examples

The innovation fragments (σ_1, σ_2) can be interpreted in different ways. We give below two examples of interpretation.

⁸For simplicity, we assume that patents are ironclad, meaning that it is impossible to enter the market with an imitating product before the expiration of the patents. Assuming that patents are “probabilistic” (i.e., have a chance of being litigated and declared invalid, as argued by Lemley and Shapiro, 2005) would complicate the analysis without offering additional insights.

2.2.1 Cost reducing innovations

Suppose that the innovation is a process innovation, which results in a reduction in the firm's marginal costs. By exploiting only a fragment σ of the innovation, each firm obtains a reduction in marginal cost that is increasing in σ , so that

$$C'(q) = 1 - f(\sigma),$$

with $f'(\sigma) > 0$. If $f(\cdot)$ is concave, the marginal effect of the innovation on cost reduction is decreasing; if $f(\cdot)$ is convex, the marginal effect is increasing. Suppose in addition, that the innovator and imitator are Cournot competitors on a market with linear demand $p = 2 - \frac{1}{9}(q_1 + q_2)$. The profits of the two firms are easily computed as

$$\begin{aligned} P(\sigma_1, \sigma_2) &= (1 + 2f(\sigma_1) - f(\sigma_2))^2, \\ Q(\sigma_1, \sigma_2) &= (1 - f(\sigma_1) + 2f(\sigma_2))^2. \end{aligned}$$

We also compute

$$\frac{\partial^2 P}{\partial \sigma_2^2} = -2f''(\sigma_2)[1 + 2f(\sigma_1) - f(\sigma_2)] + 2f'(\sigma_2)^2.$$

Hence, whenever $f(\sigma)$ is concave, P is convex in σ_2 . In order for P to be concave in σ_2 , we require the function $f(\sigma)$ to be highly convex – i.e., firms only benefit from the innovation when a large fragment is obtained. In particular, we note that when $f(\sigma) = \sigma^2$ and $\sigma_1 > \sigma_2$,

$$\frac{\partial^2 P}{\partial \sigma_2^2} = -4 - 8\sigma_1^2 + 6\sigma_2^2 < 0,$$

so that the profit of the innovator is concave in σ_2 .

2.2.2 Quality upgrades

We assume in this example that innovation results in quality upgrades, so that σ_1 and σ_2 measure the quality of the products sold. We consider a model of vertical differentiation, where each firm faces two markets: a *captive market* of size κ on which it is a monopolist, and a *competitive market* on which it competes with the other firm. On both markets, consumers have a utility function $U_i = \theta_i s - p$, where s denotes the quality of the product bought at price p , and θ_i is a measure of sensitivity to quality, which is distributed uniformly over $[0, 1]$. With this specification, the market

is not covered and consumers with low sensitivity to quality do not buy the product. Let (σ_1, σ_2) denote the qualities of the two firms with $\sigma_1 \geq \sigma_2$. On the captive market, each firm sets the monopoly price $m_i = \sigma_i/2$ and obtains a profit $\Pi_i = \kappa\sigma_i/4$. On the competitive market, the low quality firm sells to consumers in the interval $[\frac{p_2}{\sigma_2}, \frac{p_1-p_2}{\sigma_1-\sigma_2}]$ whereas the high quality firm sells to consumers in the interval $[\frac{p_1-p_2}{\sigma_1-\sigma_2}, 1]$. At equilibrium, prices are given by

$$p_1^* = \frac{2(\sigma_1 - \sigma_2)\sigma_1}{4\sigma_1 - \sigma_2}, p_2^* = \frac{(\sigma_1 - \sigma_2)\sigma_2}{4\sigma_1 - \sigma_2},$$

and profits are

$$P(\sigma_1, \sigma_2) = \frac{\kappa\sigma_1}{4} + \frac{4\sigma_1(\sigma_1 - \sigma_2)}{(4\sigma_1 - \sigma_2)^2},$$

$$Q(\sigma_1, \sigma_2) = \frac{\kappa\sigma_2}{4} + \frac{\sigma_1\sigma_2(\sigma_1 - \sigma_2)}{(4\sigma_1 - \sigma_2)^2}.$$

We verify that P is increasing in σ_1 and, for κ sufficiently high, Q is increasing in σ_2 . In addition, we observe that P is always decreasing and concave in σ_2 whereas Q is increasing in σ_1 .

2.3 Regimes of innovation protection

We consider and compare four different regimes of patent protection. First, the *binary* regime is the regime that has been considered in the literature heretofore, where the innovator is either granted protection for the entire innovation or for nothing. This regime corresponds to a strict *utility requirement* where the patent office rejects patent applications that only concern a fragment of the innovation (a step in the production process, or the characteristics of one component, which are deemed useless if they are not pooled with other innovations). Second, the *single patent regime* describes a situation where the innovator can only introduce one patent for the process – protecting either the full innovation or a fragment of the innovation. Any attempt to introduce another patent on the same process is rejected by the patent office as it does not satisfy the strict *novelty requirement*. Finally, the *sequential patent regime* corresponds to a setting where the innovator can sequentially patent different fragments of the innovation. This assumes that the novelty requirement is soft, in that the patent office employs a weak prior art test and accepts to consider patents over processes that have already been used before. For tractability reasons, we assume for now that

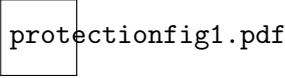


Figure 1: THE FOUR REGIMES OF PATENT PROTECTION

the innovator can only introduce two successive patents and that the second patent must cover the remainder of the innovation. In the sequential patent regime, we distinguish between a situation *with prior user rights*, where the innovator can continue to use parts of the processes that are patented by the imitator, and a situation *without prior user rights*, where the innovator is prevented from using parts of the processes that are later patented by the imitator.

We divide the analysis into two parts. We first focus on a *static optimization problem* where the innovator chooses which parts of the process to patent and to keep secret – this corresponds to the strict novelty requirement and covers both the binary and the single patent regimes. We then analyze the *dynamic optimization process* of the innovator choosing both the fraction of the innovation and the time at which a second patent is introduced – this corresponds to the soft novelty requirement and covers the sequential patent regimes with and without prior user rights. We summarize the four patent regimes in Figure 1.

3 Static optimization: The optimal mix of patent and secret

In this section, it is assumed that the patent office applies a strict novelty requirement, implying that only one patent can be introduced for a particular complex innovation. In this context, we contrast two static optimization problems according to whether the patent can only cover the entire innovation, or whether it can cover any part of it.

3.1 Binary choice

We assume here that the innovator faces a binary choice between patenting the entire innovation or keeping it as a trade secret. By patenting, the innovator obtains the deterministic profit

$$V^P = \frac{1}{r}P(1, 0) - \frac{e^{-rT}}{r}(P(1, 0) - P(1, 1)),$$

whereas by keeping the innovation secret, he obtains an expected profit

$$V^T = \frac{1}{r}P(1, 1) + \frac{1}{r+\lambda}(P(1, 0) - P(1, 1)).$$

The innovator prefers patenting over secret if the probability of discovery exceeds a threshold λ_0 :

$$V^P > V^T \iff \frac{1-e^{-rT}}{r} > \frac{1}{r+\lambda} \iff \lambda > \frac{re^{-rT}}{1-e^{-rT}} \equiv \lambda_0,$$

where λ_0 is the rate of discovery that equates the discounting-adjusted durations of the patent and of the secret (as defined, e.g., by Denicolo and Franzoni, 2004).

3.2 Single patent fragmentation

We now suppose that the innovator can choose the fragment $\sigma \in [0, 1]$ of the innovation which is patented. Because the innovation can only be patented once, even if the imitator discovers the secret she will not be authorized to patent it. Hence, the inventor will always have access to the full innovation, $\sigma_1 = 1$. On the other hand, the imitator has access to $\sigma_2 = 0$ during the patent life if she does not discover the secret, to $\sigma_2 = 1 - \sigma$ during the patent life after the secret is discovered, to $\sigma_2 = \sigma$ after the expiry of the patent if the trade secret is not discovered and $\sigma_2 = 1$ after the expiry of the patent and when the trade secret is discovered. Summarizing, the expected profit of the innovator is given by:

$$\begin{aligned} V(\sigma) &= \int_0^T \left(P(1, 0)e^{-\lambda t} + P(1, 1 - \sigma)(1 - e^{-\lambda t}) \right) e^{-rt} dt \\ &\quad + \int_T^\infty \left(P(1, \sigma)e^{-\lambda t} + P(1, 1)(1 - e^{-\lambda t}) \right) e^{-rt} dt \\ &= \frac{1-e^{-(r+\lambda)T}}{r+\lambda} P(1, 0) + \left(\frac{1-e^{-rT}}{r} - \frac{1-e^{-(r+\lambda)T}}{r+\lambda} \right) P(1, 1 - \sigma) \\ &\quad + \frac{e^{-(r+\lambda)T}}{r+\lambda} P(1, \sigma) + \left(\frac{e^{-rT}}{r} - \frac{e^{-(r+\lambda)T}}{r+\lambda} \right) P(1, 1). \end{aligned}$$

Increasing the share of the innovation that is patented (σ) has two opposite effects. On the one hand, it increases the per-period value for the innovator when the patent is in force and the secret is discovered; on the other hand, it decreases the value when the patent has expired and the secret is not discovered:

$$\frac{dV}{d\sigma} = \underbrace{\left(\frac{1-e^{-rT}}{r} - \frac{1-e^{-(r+\lambda)T}}{r+\lambda} \right)}_{+} \underbrace{(-P_2(1, 1 - \sigma))}_{+} + \underbrace{\frac{e^{-(r+\lambda)T}}{r+\lambda}}_{+} \underbrace{P_2(1, \sigma)}_{-}. \quad (1)$$

If the function P is convex in σ_2 , the expected profit V is convex in σ and the optimal solution is either to patent the full innovation (if $\lambda > \lambda_0$) or to keep the entire innovation as a trade secret (if $\lambda < \lambda_0$). If the function P is concave in σ_2 , the expected profit V is concave in σ and, for intermediate values of λ , the optimal mix of patent and trade secret results in an interior value $\sigma^* \in (0, 1)$. Define

$$K \equiv \frac{\frac{1-e^{-rT}}{r} - \frac{1-e^{-(r+\lambda)T}}{r+\lambda}}{\frac{e^{-(r+\lambda)T}}{r+\lambda}} = 1 + \frac{(\lambda - \lambda_0)}{(r + \lambda_0) e^{-(r+\lambda)T}},$$

Notice that $K = 1$ for $\lambda = \lambda_0$, $K < 1$ if $\lambda < \lambda_0$ and $K > 1$ if $\lambda > \lambda_0$. As P is concave in σ , $\frac{P_2(1,0)}{P_2(1,1)} < 1$. Define implicitly the two values $\underline{\lambda} < \lambda_0 < \bar{\lambda}$ which are respectively solutions to:

$$\frac{P_2(1,0)}{P_2(1,1)} = K \text{ and } \frac{P_2(1,0)}{P_2(1,1)} = \frac{1}{K}.$$

We are now ready to characterize the optimal fragmentation of the innovation (see the proof in Appendix 8.1).

Proposition 1 *If the profit of the innovator is convex in σ_2 , the innovator either patents the entire innovation (if $\lambda > \lambda_0$) or keeps the entire innovation as a trade secret (if $\lambda < \lambda_0$). If the profit of the innovator is concave in σ_2 , the innovator patents the entire innovation if $\lambda > \bar{\lambda}$ and keeps the entire innovation secret if $\lambda < \underline{\lambda}$. If $\underline{\lambda} < \lambda < \bar{\lambda}$, the innovator optimally patents a fragment $\sigma^* \in [0, 1]$ of the innovation where σ^* solves:*

$$\frac{P_2(1, \sigma^*)}{P_2(1, 1 - \sigma^*)} = K. \quad (2)$$

The optimal patented share σ^ is increasing in λ , T and r .*

Proposition 1 clearly distinguishes between situations where the function P is convex (the competitor benefits from small increments in the innovation) and situations where the function P is concave (the competitor needs to attain a critical share of the innovation to be able to exploit it). In the first case, the innovator optimally makes a binary choice between patenting the entire innovation or keeping it entirely secret. In the second case, the innovator chooses to protect the innovation by a mix of patent and secret, where the mix contains more patents when the innovation is easier to reverse engineer, when the patent length is longer and when the discount rate is higher.

We illustrate Proposition 1 by computing the optimal fragmentation of the innovation in the two examples discussed in the previous section, for a patent life $T = 20$ and a discount rate $r = 0.05$, which result in a value $\lambda_0 \sim 0.0291$.

Example 1 (continued) In the model of cost reducing innovation, when the cost function is a quadratic function of σ , $c(\sigma) = 1 - \sigma^2$, the optimal fragmentation σ^* is given below for different values of λ .

λ	0.005	0.01	0.02	0.03	0.04	0.05	0.06	0.08
σ^*	0.056	0.130	0.318	0.517	0.676	0.785	0.856	0.931

Example 2 (continued) In the model of innovation resulting in quality increments, the optimal fragmentation σ^* is given below for different values of λ :

λ	0.005	0.01	0.02	0.03	0.04	0.05	0.06	0.08
σ^*	0	0	0.251	0.502	0.741	0.928	1	1

4 Dynamic optimization: optimal timing of patents

We turn now to a world where the patent office applies a soft novelty requirement. In the U.S., the American Invents Act (AIA, passed in 2011) has implemented such a change. Besides the well-known conversion of the U.S. patent system from a “first to invent” system to a “first inventor to file” system, the AIA also eliminates several types of secret prior art.⁹ In practice, this change (which became effective in March 2013) implies that long held trade secrets are now patentable. In our setting, this means that if the innovator leaves a fraction of the innovation under secret, the secret fragment can be patented later on, either by the innovator or by the imitator.

In this world, the innovator faces a dynamic optimization problem as he has to determine not only which fraction of the innovation to patent in

⁹As Maier (2011) explains: “the forfeiture that previously penalized inventors for maintaining their inventions as trade secrets for some period of time longer than a year is no longer applicable, and inventors are left with the option to practice their invention as trade secrets for now and still patent those same inventions later (assuming, of course, that no other inventor files a patent application claiming the same subject matter first).”

the first place, but also the time at which to introduce a second patent on the remaining fraction of the innovation.¹⁰ As this regime also allows the imitator to patent whichever part of the innovation that was kept secret, the innovator’s decisions heavily depends on whether he is allowed or not to continue exploiting any part of the innovation that the imitator has succeeded in patenting, i.e., whether he can invoke prior user rights or not. We analyze the two sequential patent regimes in turn.¹¹

4.1 Dynamic optimization with prior user rights

We suppose here that the initial innovator is granted prior user rights, so that he may exploit the entire innovation irrespective of the fact that the competitor discovers his secret. The innovator now selects two control variables: the fraction of the innovation which is covered by the first patent, σ , and the time η at which the secret part of the innovation is patented if the competitor has not discovered it.

If the imitator discovers the secret before η , he gains access to the secret part of the innovation during the time of the first patent and to the entire innovation after the expiry of the first patent. The expected payoff to the inventor is thus $P(1, 0)$ until the discovery of the secret, $P(1, 1-\sigma)$ during the life of the patent after the secret has been discovered, and $P(1, 1)$ after the expiry of the patent. If on the other hand, the imitator has not discovered the secret before η , she will stop making efforts to invent around the patent after η , and the inventor’s profit is given by $P(1, 0)$ during the first patent, $P(1, \sigma)$ after the expiry of the first patent and before the expiry of the

¹⁰In order to simplify the analysis, we suppose that the innovation can only be divided into two parts, and that the innovator must patent the remainder of the innovation in the second stage. This corresponds to many realistic situations where firms divide their production processes into two parts, patenting the process itself but keeping the components secret.

¹¹The AIA has also modified provisions regarding prior user rights. As explained by Kappos and Stanek Rea (2012, p. 1): “The AIA also expands the “prior user rights” defense to infringement and broadens the classes of patents that are eligible for the new limited prior user rights defense. (...) U.S. law already provided a prior user rights defense that was limited to patents directed to methods of conducting business. The AIA, by contrast, extends the prior user rights defense to patents covering all technologies, not just business methods.” According to the Tegernsee Experts Group (2012, p. 2), “[p]rior user rights are provided for by the different national legislations (...) [which] have common ground, but also have differences in the conditions under which they may be acquired.”

second, and $P(1, 1)$ after the expiry of the second patent.

We distinguish between the case where the innovator patents the secret before the expiry of the first patent ($\eta < T$) and after the expiry of the first patent ($\eta > T$). In the first case, the expected value of the innovator as a function of σ and η is given by

$$\begin{aligned} V(\sigma, \eta) = & \int_0^\eta \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-r\tau}) P(1, 0) + (e^{-r\tau} - e^{-rT}) P(1, 1 - \sigma) \right. \\ & \left. + e^{-rT} P(1, 1) \right\} d\tau \\ & + \frac{1}{r} e^{-\lambda\eta} \left\{ (1 - e^{-rT}) P(1, 0) + (e^{-rT} - e^{-r(T+\eta)}) P(1, \sigma) \right. \\ & \left. + e^{-r(T+\eta)} P(1, 1) \right\}. \end{aligned}$$

In the second case, the expected value is given by

$$\begin{aligned} V(\sigma, \eta) = & \int_0^T \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-r\tau}) P(1, 0) + (e^{-r\tau} - e^{-rT}) P(1, 1 - \sigma) \right. \\ & \left. + e^{-rT} P(1, 1) \right\} d\tau \\ & + \int_T^\eta \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-rT}) P(1, 0) + (e^{-rT} - e^{-r\tau}) P(1, \sigma) \right. \\ & \left. + e^{-r\tau} P(1, 1) \right\} d\tau \\ & + \frac{1}{r} e^{-\lambda\eta} \left\{ (1 - e^{-rT}) P(1, 0) + (e^{-rT} - e^{-r(T+\eta)}) P(1, \sigma) \right. \\ & \left. + e^{-r(T+\eta)} P(1, 1) \right\}. \end{aligned}$$

4.1.1 Optimal timing of patents

We first consider, for a fixed σ , the optimal choice of the time of the second patent, η . By increasing η , the innovator increases the length of the patent protection of the innovation, but also increases the time during which the trade secret is not protected, making it more likely that the imitator discovers the unpatented part of the innovation. As we will see, this trade-off may result in the innovator choosing a *finite time* η^* *at which the secret part of the innovation is patented*.

The trade-off crucially depends on two measures: the loss in utility due to the fact that the second patent expires,

$$\Delta P^1 = P(1, \sigma) - P(1, 1),$$

and the loss in utility due to the fact that the imitator discovers the secret,

$$\Delta P^2 = P(1, 0) - P(1, 1 - \sigma).$$

Notice that, if P is concave in σ_2 , $\Delta P^1 > \Delta P^2$ whereas if P is convex in σ_2 , $\Delta P^1 < \Delta P^2$. In order to derive the optimal timing of the second patent η^* , we compute the derivative of the expected value $V(\sigma, \eta)$ with respect to η , both for $\eta < T$ and for $\eta > T$. For $\eta < T$,

$$\begin{aligned} \frac{\partial V}{\partial \eta} &= e^{-\lambda \eta} \left[e^{-r(T+\eta)} \Delta P^1 \right. \\ &\quad \left. - \frac{\lambda}{r} \left((e^{-r\eta} - e^{-rT}) \Delta P^2 + (e^{-rT} - e^{-r(T+\eta)}) \Delta P^1 \right) \right]. \end{aligned}$$

For $\eta > T$,

$$\frac{\partial V}{\partial \eta} = e^{-(\lambda+r)\eta} \lambda_0 + r(\lambda_0 - \lambda) \Delta P^1.$$

Building on the computation of these derivatives, the following proposition (whose proof can be found in Appendix 8.2) summarizes the optimal choice of η as a function of the parameters.

Proposition 2 *Given σ , the optimal timing of the second patent is given as follows. (1) If $\Delta P^2 > \Delta P^1$,*

- $\eta^* = 0$ for $\lambda_0 \Delta P^1 \leq \lambda \Delta P^2 - \lambda_0 (1 - e^{-\lambda T}) (\Delta P^2 - \Delta P^1)$,
- $\eta^* = \infty$ otherwise.

(2) *If $\Delta P^1 > \Delta P^2$,*

- $\eta^* = 0$ for $\lambda_0 \frac{\Delta P^1}{\Delta P^2} \leq \lambda$,
- $0 < \eta^* < T$ for $\lambda_0 \leq \lambda \leq \lambda_0 \frac{\Delta P^1}{\Delta P^2}$,
- $\eta^* = \infty$ for $\lambda \leq \lambda_0$.

Proposition 2 characterizes the optimal delay between the two patents for a fixed value of σ . As in the static case, the concavity of the profit of the innovator in σ_2 plays a crucial role in determining the optimal delay. If P is convex, the innovator either immediately patents the entire innovation, or chooses never to file a second patent. If P is concave, for intermediate values of the probability of discovery, the innovator chooses to file a second

patent before the expiry of the first patent. We note that one situation never arises: it is never optimal for the inventor to file a second patent in finite time but after the expiry of the first patent.

The first order condition for optimal timing of the second patent, when $\Delta P^1 > \Delta P^2$ and $\lambda_0 \leq \lambda \leq \lambda_0 \frac{\Delta P^1}{\Delta P^2}$ is given by:

$$e^{-r\eta} = \frac{\Delta P^1 - \Delta P^2}{\frac{r}{\lambda} \Delta P^1 - \frac{r}{\lambda_0} \Delta P^2 + \Delta P^1 - \Delta P^2}. \quad (3)$$

4.1.2 Optimal innovation fragmentation

We now consider the optimal choice of the share of the innovation covered by the first patent, σ . If $\eta = 0$, the innovator immediately patents the full innovation. If $\eta = \infty$, his optimal choice is similar to that of Section 3.2. As shown by Proposition 2, the case $\eta > T$ is irrelevant. When $0 < \eta \leq T$, the derivative of V with respect to σ can be computed as

$$\begin{aligned} \frac{\partial V}{\partial \sigma} &= - \left[\frac{1 - e^{-rT}}{r} - \frac{1 - e^{-(\lambda+r)\eta}}{\lambda + r} - \frac{e^{-\lambda\eta}}{r} (e^{-r\eta} - e^{-rT}) \right] P_2(1, 1 - \sigma) \\ &+ \frac{e^{-(\lambda\eta+rT)}}{r} (1 - e^{-r\eta}) P_2(1, \sigma). \end{aligned}$$

Hence, when P is convex in σ_2 , the optimal choice is either $\sigma^* = 0$ or $\sigma^* = 1$, and when P is concave, the solution is either $\sigma^* = 0$ or $\sigma^* = 1$ or the interior solution satisfying

$$\frac{P_2(1, \sigma^*)}{P_2(1 - \sigma^*)} = \frac{\frac{1 - e^{-rT}}{r} - \frac{1 - e^{-(\lambda+r)\eta}}{\lambda + r} - \frac{e^{-\lambda\eta}}{r} (e^{-r\eta} - e^{-rT})}{\frac{e^{-(\lambda\eta+rT)}}{r} (1 - e^{-r\eta})}. \quad (4)$$

4.1.3 Optimal dynamic patenting strategy

We now piece together the analysis of the optimal timing and the optimal fragmentation to characterize the optimal dynamic patenting strategy of the innovator. Suppose first that P is convex in σ . Then $\Delta P^2 > \Delta P^1$ and either $\eta^* = 0$ (and the innovator immediately patents the full innovation) or $\eta^* = \infty$ (and the innovator optimally chooses, as in Section 3.2, either to patent the full innovation if $\lambda \geq \lambda_0$ or to keep the entire innovation secret if $\lambda \leq \lambda_0$).

Consider next the more interesting case where P is concave in σ_2 . If $\lambda \leq \lambda_0$, $\eta^* = \infty$ for all values of σ , and the optimal choice is the same σ^*

as in Section 3.2, with $\sigma^* = 0$ for $\lambda \leq \underline{\lambda}$ and $\sigma^* \in (0, \frac{1}{2})$ for $\lambda_0 > \lambda > \underline{\lambda}$. If $\lambda \geq \lambda_0$, according to Proposition 2, the regime of repatenting depends on the value of σ . This requires an understanding of the effect of changes in σ on the values of ΔP^2 and ΔP^1 . Let

$$H(\sigma) \equiv \frac{P(1, \sigma) - P(1, 1)}{P(1, 0) - P(1, 1 - \sigma)} = \frac{\Delta P^1}{\Delta P^2}.$$

We show in Appendix 8.3 that $H(\sigma)$ is increasing from $H(0) = 1$ to $H(1) = \frac{P_2(1,1)}{P_2(1,0)}$ and that $\lambda_0 H(1) > \bar{\lambda}$. For any $\lambda_0 < \lambda < \lambda_0 H(1)$, we can thus compute $\hat{\sigma}$ such that $\frac{\Delta P^1}{\Delta P^2} \leq \frac{\lambda}{\lambda_0}$ if and only if $\sigma \leq \hat{\sigma}$. We claim that at the optimum, the innovator must choose $\sigma \in (\hat{\sigma}, 1]$. The argument is based on revealed preference. If the innovator chooses $\sigma \in (\hat{\sigma}, 1]$, by Proposition 2, the optimal timing choice is $\eta^* > 0$, so that the utility of the innovator must be greater than the utility obtained when $\eta^* = 0$, and the second patent is filed immediately. If the innovator chooses $\sigma \in [0, \hat{\sigma}]$, by Proposition 2, the second patent is immediately filed, resulting in a profit which, as we argued, must be strictly lower than the profit obtained by choosing $\sigma > \hat{\sigma}$ followed by $\eta^* > 0$. We summarize this discussion in the following proposition.

Proposition 3 *If the function P is convex in σ_2 , the optimal dynamic patenting strategy is identical to the static strategy, and the innovator chooses $\sigma^* = 0$, $\eta^* = \infty$ if $\lambda < \lambda_0$ and $\sigma^* = 1$ if $\lambda > \lambda_0$. If the function P is concave in σ_2 ,*

- *If $\lambda < \lambda_0$, the optimal dynamic patenting strategy is identical to the static strategy and the innovator chooses $\sigma^* \in (0, 1)$, $\eta^* = \infty$, where σ^* solves Equation (2).*
- *If $\lambda_0 < \lambda < \lambda_0 H(1)$, the optimal dynamic patenting strategy is to select $\sigma^* \in (0, 1)$, $\eta^* \in (0, T)$ which are solutions to Equations (3) and (4).*
- *If $\lambda > \lambda_0 H(1)$, the optimal dynamic patenting strategy is identical to the static strategy and the innovator chooses $\sigma^* = 1$.*

Proposition 3 identifies situations where the innovator optimally chooses to patent sequentially two fragments of the innovation, and the dynamic patenting strategy differs from the static choice of the optimal mix of patent and secrets. These situations arise when the function P is concave in σ_2 and the probability of reverse engineering the innovation is intermediate. We

also note that, because $\lambda_0 H(1) > \bar{\lambda}$, the region of parameters for which the innovator chooses to fragment the innovation is larger in the dynamic patenting regime than in the static regime. The following computations characterize the optimal interior values of η and σ in the two examples discussed in Section 2.2 for values of λ greater than λ_0 .

Example 1 (continued) In the model of cost reducing innovation, when the cost function is a quadratic function of σ , $c(\sigma) = 1 - \sigma^2$, the optimal fragmentation σ^* and timing η^* are given below for different values of λ .

λ	0.03	0.04	0.05	0.06	0.08
σ^*	0.513	0.626	0.700	0.752	0.816
η^*	19.40	14.78	12.10	10.33	8.04

Example 2 (continued) In the model of innovation resulting in quality increments, the optimal fragmentation σ^* is given below for different values of λ :

λ	0.03	0.04	0.05	0.06	0.08
σ^*	0.517	0.683	0.823	0.958	1
η^*	19.77	18.37	18.33	19.41	-

The numerical computations first show, quite naturally, that the size of the first fragment is lower in the dynamic setting than in the static setting (see the two tables in Section 3.2). As the innovator will file a second patent on the innovation, he has an incentive to decrease the share of the innovation that is patented the first time. Computations in Example 1 suggest that the share of the innovation and the delay between the second patent play complementary roles in the strategy of the innovator. As the threat of imitation increases, the innovator simultaneously increases the share of the innovation covered by the first patent and reduces the delay before filing the second patent. However, Example 2 shows that this phenomenon is not general. In the example of quality upgrades, as the threat of imitation increases, the share of the innovation covered by the patent increases but the delay before the second patent (which remains very large and close to the patent length $T = 20$) is not always decreasing.

In general, the share of the innovation patented σ^* and the delay before re-patenting η^* are neither substitutes nor complements. An inspection of the cross derivative $\frac{\partial^2 V}{\partial \eta \partial \sigma}$ indicates that the two instruments are complements when $P_2(1, \sigma)$ is high and $P_2(1, 1 - \sigma)$ low and substitutes when $P_2(1, 1 - \sigma)$ is high and $P_2(1, \sigma)$ low. Given that profit is concave in σ_2 , this implies that there exists a threshold value of the patent share $\underline{\sigma}$ such that patent share and time before re-patenting are complements when the patent share is lower than $\underline{\sigma}$ and substitutes when the patent share is higher than $\underline{\sigma}$. This explains why, in Example 2, for low values of λ , when the optimal patent share is low, patent share and time before re-patenting increase co-monotonically. But when the patent share becomes high, the two instruments become substitute, and an increase in λ results in opposite effects on η and σ .

4.2 Dynamic optimization without prior user rights

If the innovator does not possess prior user rights, he may become unable to exploit the part of the innovation patented by the imitator. The competitor can force the initial inventor to downgrade his product or resort to a less efficient production process, resulting in a payoff of $P(\sigma, 1 - \sigma)$ during the lifetime of the first patent and $P(\sigma, 1)$ between the expiry of the first patent held by the innovator and the expiry of the second patent held by the imitator. The expected profit of the inventor when $\eta < T$ is now given by

$$\begin{aligned}
V(\eta, \sigma) = & \int_0^\eta \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-r\tau}) P(1, 0) + (e^{-r\tau} - e^{-rT}) P(\sigma, 1 - \sigma) \right. \\
& \left. + (e^{-rT} - e^{-r(T+\tau)}) P(\sigma, 1) + e^{-r(T+\tau)} P(1, 1) \right\} d\tau \\
& + \frac{1}{r} e^{-\lambda\eta} \left\{ (1 - e^{-rT}) P(1, 0) + (e^{-rT} - e^{-r(T+\eta)}) P(1, \sigma) \right. \\
& \left. + e^{-r(T+\eta)} P(1, 1) \right\}.
\end{aligned}$$

and when $\eta > T$ by

$$\begin{aligned}
V(\eta, \sigma) = & \int_0^T \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{r\tau})^P(1, 0) + (e^{-r\tau} - e^{-rT}) P(\sigma, 1 - \sigma) \right. \\
& \left. + (e^{-rT} - e^{-r(T+\tau)}) P(\sigma, 1) + e^{-r(T+\tau)} P(1, 1) \right\} d\tau \\
& + \int_T^\eta \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-rT}) P(1, 0) + (e^{-rT} - e^{-r\tau}) P(1, \sigma) \right. \\
& \left. + (e^{-r\tau} - e^{-r(T+\tau)}) P(\sigma, 1) + e^{-r(T+\tau)} P(1, 1) \right\} d\tau \\
& + \frac{1}{r} e^{-\lambda\eta} \left\{ (1 - e^{-rT}) P(1, 0) + (e^{-rT} - e^{-r(T+\eta)}) P(1, \sigma) \right. \\
& \left. + e^{-r(T+\eta)} P(1, 1) \right\}.
\end{aligned}$$

4.2.1 Optimal timing of patents

In order to define the optimal timing of patents, we define a new term,

$$\Delta P^3 = P(1, \sigma) - P(\sigma, 1),$$

which measures the loss experienced by the innovator when the imitator discovers and patents the trade secret after the expiry of the first patent. Notice that $\Delta P^3 \geq \Delta P^1$. We also compute the loss to the innovator when the imitator discovers the secret as

$$\Delta P^4 = P(1, 0) - P(\sigma, 1 - \sigma).$$

Equipped with this notation, we compute the derivative of the firm's profit with respect to η . For $\eta < T$,

$$\begin{aligned}
\frac{\partial V}{\partial \eta} = & e^{-\lambda\eta} \left[e^{-r(T+\eta)} \Delta P^1 \right. \\
& \left. - \frac{\lambda}{r} \left((e^{-r\eta} - e^{-rT}) \Delta P^4 + (e^{-rT} - e^{-r(T+\eta)}) \Delta P^3 \right) \right].
\end{aligned}$$

For $\eta > T$,

$$\frac{\partial V}{\partial \eta} = \frac{e^{-(\lambda+r)\eta}}{\lambda_0 + r} (\lambda_0 \Delta P^1 - \lambda \Delta P^3).$$

Following the same steps as in Subsection 4.1.1, we characterize the optimal timing of patents for a fixed σ in the next proposition (whose proof can be found in Appendix 8.4).

Proposition 4 *Given σ , the optimal timing of the second patent is given as follows. (1) If $\Delta P^4 > \Delta P^3$,*

- $\eta^* = 0$ for $\lambda_0 \Delta P^1 \leq \lambda \Delta P^4 - \lambda_0(1 - e^{-\lambda T})(\Delta P^4 - \Delta P^3)$,
- $\eta^* = \infty$ otherwise

(2) If $\Delta P^3 > \Delta P^4$,

- $\eta^* = 0$ for $\lambda_0 \Delta P^1 \leq \lambda \Delta P^4$,
- $0 < \eta^* < T$ for $\lambda \Delta P^4 \leq \lambda_0 \Delta P^1 \leq \lambda \Delta P^3$,
- $\eta^* = \infty$ for $\lambda_0 \Delta P^1 \leq \lambda \Delta P^3$.

4.2.2 Optimal dynamic patenting strategy

The characterization of the optimal dynamic patenting strategy when the inventor does not hold prior user rights differs from the characterization of the optimal dynamic patenting strategy with prior user rights in two important respects. First, the comparison between ΔP^3 and ΔP^4 does not result from the concavity or convexity of the profit function P . If P is convex, and in addition, the marginal effect of an increase in σ_1 is higher for lower values of σ_2 ($P_{12} < 0$), we obtain

$$\begin{aligned}
\Delta P^4 - \Delta P^3 &= P(1, 0) - P(\sigma, 1 - \sigma) - P(1, \sigma) + P(\sigma, 1), \\
&= [P(1, 0) - P(1, \sigma)] - [P(\sigma, 1 - \sigma) - P(\sigma, 1)] \\
&\geq [P(\sigma, 0) - P(\sigma, \sigma)] - [P(\sigma, 1 - \sigma) - P(\sigma, 1)] \\
&\geq 0.
\end{aligned}$$

It is harder to find conditions under which $\Delta P^4 < \Delta P^3$, as this would require both that P is concave and that $P_{12} > 0$, and the latter condition is unlikely to be satisfied in regular applications. The second difficulty stems from the lack of concavity/convexity of the function V with respect to the fragmentation parameter σ . When the innovation is only patented once ($\eta = \infty$), a simple computation shows that¹²

$$\begin{aligned}
\frac{\partial V}{\partial \sigma} &= \left(\frac{1 - e^{-rT}}{r} - \frac{1 - e^{-(\lambda+r)T}}{\lambda+r} \right) (P_1(\sigma, 1 - \sigma) - P_2(\sigma, 1 - \sigma)) \\
&\quad + \frac{e^{-(\lambda+r)T}}{\lambda+r} P_2(1, \sigma) + \frac{\lambda}{r(\lambda+r)} (1 - e^{-rT}) P_1(\sigma, 1).
\end{aligned}$$

A change in σ affects both the fraction of the innovation that the imitator can exploit (through the derivative P_2) and the fraction of the innovation

¹²A similar result holds for interior values of η .

that the inventor can still exploit if the imitator patents the secret part (through the derivative P_1). Simple intuition suggests that if P is concave in σ_2 , it should be convex in σ_1 and vice versa. But this implies that the function V is unlikely to be concave or convex in σ , so that we cannot characterize the optimal fragmentation share σ^* through a simple first order condition.

We now consider the two examples of Section 2.2. In Example 2, the sign of $\Delta P^4 - \Delta P^3$ is not constant over σ , and we omit the characterization of the optimal patenting policy. In Example 1, we observe that the sign of $\Delta P^4 - \Delta P^3$ is constant, allowing us to compute the optimal policy.

Example 1 (continued) In the model of cost reducing innovation, when the cost function is a quadratic function of σ , $c(\sigma) = 1 - \sigma^2$, we have $\Delta P^2 = 9 - \sigma^2(\sigma + 2)^2 > 9 - 3\sigma^2(\sigma^2 + 2) = \Delta P^3$ for all σ . Hence, the innovator only patents once and the fragment patented is given below for different values of λ .

λ	0.005	0.01	0.02	0.03	0.04	0.05	0.06	0.08
σ^*	0	0	1	1	1	1	1	1

It is interesting to observe that, given that $P_{12} < 0$, even when P is concave in σ_2 , the function V turns out to be convex in σ , so that the innovator will never choose to fragment the innovation. The absence of prior user rights, by making it more costly for the innovator to lose his trade secret, increases the incentive to patent to the point where the innovator chooses to patent the entire process, even when $\lambda < \lambda_0$.

5 Welfare Comparisons

In this section, we compare the four regimes of patent protection from the point of view of the three stakeholders: the innovator, the imitator and consumers. **As we suppose that the innovation has already been made, we do not explicitly analyze the effects of the four regimes on the incentives to innovate. However, incentives to innovate are clearly congruent with the profit of the innovator, so that a comparison of the profit of the innovator under the four regimes gives an indication on the effect of the regimes of patent protection on the speed of innovation.** We conduct the welfare comparisons in the context of the two illustrating examples of the paper.

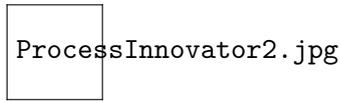


Figure 2: PROCESS INNOVATION: INNOVATOR'S PROFIT

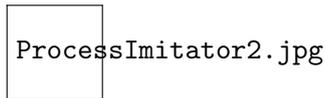


Figure 3: PROCESS INNOVATION: IMITATOR'S PROFIT

5.1 Cournot model with process innovation

In addition to the profits of the innovator and imitator, P and Q , we compute consumer surplus as

$$S(\sigma_1, \sigma_2) = \frac{1}{2}(1 + f(\sigma_1) + f(\sigma_2))^2.$$

Figures 2, 3 and 4 compare, for different values of λ , the welfare of the innovator, imitator and consumers under the four regimes for $f(\sigma) = \sigma^2$.

For the innovator, not surprisingly, as regimes become more flexible, profits increase. Hence, the dynamic regime with prior user rights dominates the static regime, which dominates the binary regime. We also observe that the dynamic regime without prior user rights results in a profit which is higher than the binary regime but lower than any other patent protection regime. Interestingly, in the Cournot model, the welfare of the imitator and of consumers are aligned and in complete conflict with the profit of the innovator. The competitor and consumers favor regimes with less flexibility, preferring the binary regime to the static regime, and to the dynamic regime with prior user rights. Comparing the dynamic regimes with and without prior user rights, we observe that, clearly, the imitator gains when the innovator cannot hold prior user rights. However, consumers lose as this implies that, after the competitor discovers the trade secret for low values of λ , the innovator is not able to exploit the cost reducing innovation, yielding lower

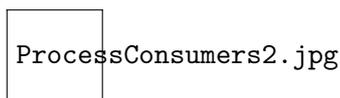


Figure 4: PROCESS INNOVATION: CONSUMER SURPLUS

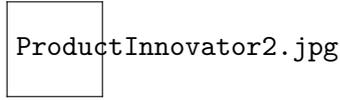


Figure 5: QUALITY INCREMENTS: INNOVATOR'S PROFIT

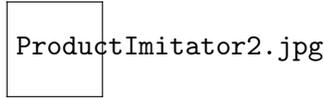


Figure 6: QUALITY INCREMENTS: IMITATOR'S PROFIT

quantities in equilibrium.¹³

On balance, the comparison between the four regimes indicates a conflict in the welfare of the innovator, imitator and consumers. If however, the intellectual property regime puts more weight on the inventor in order to give him incentives to innovate, our analysis suggests that higher flexibility should always be favored.

5.2 Vertical differentiation model with quality increments

In the vertical differentiation model, we compute the surplus of consumers when $\sigma_1 > \sigma_2$ as

$$S(\sigma_1, \sigma_2) = \frac{\kappa}{8}(\sigma_1 + \sigma_2) + \frac{\sigma_1^2(4\sigma_1 + 5\sigma_2)}{2(4\sigma_1 - \sigma_2)^2}.$$

Figures 5, 6 and 7 illustrate the welfare of the innovator, imitator and consumers for different values of λ when $\kappa = 3$ for three regimes of patent protection.¹⁴

¹³Our analysis suggests thus an additional advantage of prior user rights, which was not stressed in the literature considering patents and secrets as mutually exclusive protection mechanisms (as in Denicolo and Franzoni, 2004, or Shapiro, 2006).

¹⁴Recall that the optimal behavior of the innovator is not well defined in the dynamic patent protection regimes without prior user rights.

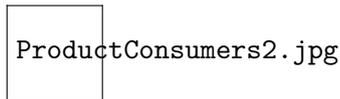


Figure 7: QUALITY INCREMENTS: CONSUMER SURPLUS

As expected, the innovator always prefers regimes with more flexibility. But in the vertical differentiation model, contrary to the Cournot model, the imitator and the consumers also may prefer the more flexible regimes. In fact, they both rank the static fragmentation regime above the binary regime, preferring to let the inventor segment the innovation in different pieces. The competitor also prefers the dynamic regime to the static regime for all values of λ whereas consumers prefer the dynamic regime except for λ close to λ_0 . The analysis of the model of vertical differentiation with quality increments thus shows that the conflict in welfare between innovator, imitator and consumers does not necessarily arise and that *there exist circumstances where all three types of agents prefer more flexible intellectual property rights regimes*.

6 Robustness and extensions

In this section, we test the robustness of our results by extending our framework in two directions: first, we endogenize the imitation effort and second, we modify the sequential patent regime by allowing a second patent that does not necessarily cover the remaining secret part of the innovation.

6.1 Endogenous imitation effort

We analyze briefly the endogenous choice of the imitator, who selects the imitation rate λ at a cost $C(\lambda)$. In the dynamic model with prior user rights and $\eta < T$, the derivative of the present discounted value of the imitator with respect to λ is given by:

$$\begin{aligned} \frac{\partial W(\lambda, \sigma, \eta)}{\partial \lambda} &= (Q(1, 1 - \sigma) - Q(1, 0)) \left\{ \frac{\eta \lambda e^{-\lambda \eta}}{\lambda + r} (e^{-r\eta} - e^{-rT}) \right. \\ &+ \frac{r}{(\lambda + r)^2} (1 - e^{-\eta(\lambda+r)} - \eta(\lambda + r)e^{-\lambda\eta - rT}) \left. \right\} \\ &+ (Q(1, 1) - Q(1, \sigma)) \eta e^{-(\lambda+r)\eta} (1 - e^{-rT}) \\ &> 0. \end{aligned}$$

Notice that, as the fraction of the innovation that is patented, σ , increases, $Q(1, 1 - \sigma) - Q(1, 0)$ and $Q(1, 1) - Q(1, \sigma)$ decrease, so that the optimal value of λ goes down. Hence, when the competitor chooses her imitation efforts endogenously, higher values of σ result in lower values of λ . By comparison with the baseline model with exogenous imitation efforts,

we remark that *the inventor has an incentive to increase the fragment of the innovation that is patented in order to reduce the imitation efforts of his competitor*. However, the thrust of the analysis, including the existence of parameter regions for which the inventor chooses interior values of σ and η , remains unchanged.

6.2 Continuous fragmentation of the innovation

We examine now a variant of the model where the inventor may choose, in the second patent, only to patent a fraction $\rho < 1 - \sigma$ of the secret.¹⁵ When $\eta < T$, the present discounted value of the inventor is given by:

$$\begin{aligned}
V(\sigma, \eta) = & \int_0^\eta \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-r\tau})P(1, 0) + (e^{-r\tau} - e^{-rT})P(1, 1 - \sigma) \right. \\
& \left. + e^{-rT}P(1, 1) \right\} d\tau \\
& + \int_\eta^T \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-r\tau})P(1, 0) + (e^{-r\tau} - e^{-rT})P(1, 1 - \sigma - \rho) \right. \\
& \left. + (e^{-rT} - e^{-r(T+\eta)})P(1, 1 - \rho) + e^{-r(T+\eta)}P(1, 1) \right\} d\tau \\
& + \int_T^{T+\eta} \frac{\lambda}{r} e^{-\lambda\tau} \left\{ (1 - e^{-rT})P(1, 0) + (e^{-rT} - e^{-r\tau})P(1, \sigma) \right. \\
& \left. + (e^{-r\tau} - e^{-r(T+\eta)})P(1, 1 - \rho) + e^{-r(T+\eta)}P(1, 1) \right\} d\tau \\
& + \int_\eta^\infty \frac{\lambda}{r} e^{-\lambda\tau} \left\{ -\lambda\tau(1 - e^{-rT})P(1, 0) + (e^{-rT} - e^{-r(T+\eta)})P(1, \sigma) \right. \\
& \left. + (e^{-r(T+\eta)} - e^{-r\tau})P(1, \sigma + \rho) + e^{-r\tau}P(1, 1) \right\} d\tau.
\end{aligned}$$

While the computations of the optimal values of η , σ and ρ are clearly more complicated than in the baseline model, we can use the same steps as in Section 4.1.3 to show that, when P is convex in σ_2 , the optimal strategy is a binary strategy whereas when P is concave in σ_2 , the optimal strategy may involve an interior choice of η , ρ and σ for some parameter values. The main qualitative results of the analysis remain unchanged.

7 Conclusion

We have provided a unified model to study the protection of complex innovations, i.e., innovations that can be fragmented into several sub-innovations,

¹⁵Notice that the imitator never has an incentive to patent less than $1 - \sigma$ since the secret is already known and exploited by the initial inventor.

each of them being patentable separately. The model allows us to analyze the innovator's choice of patent/secret mix under various patent regimes, which differ according to the strength of the utility and the novelty requirements. We are also able to perform some welfare comparisons of these patent regimes. Our main result is to find conditions under which the innovator optimally chooses to mix patents and secrets (in a static framework corresponding to a strict novelty requirement), or to patent sequentially two fragments of the innovation (in a dynamic framework corresponding to a softer novelty requirement). We also find examples where the other stakeholders (namely a potential imitator and the consumers) may agree with the innovator's conduct, suggesting that more flexible patent regimes could be welfare-enhancing. It is important to stress that a pre-condition for these results to apply is that the innovator's profit function be concave in the fraction of the innovation that the imitator can exploit. This occurs when the imitator must learn a large fragment of the innovation in order to be able to exploit it usefully. In contrast, if convexity prevails, the innovator will optimally choose an all-or-nothing strategy that consists in patenting the whole innovation or in keeping it altogether secret.

Our framework could be extended in a number of directions that we leave for further research. First, we have assumed in our model that there is no information leakage in the patenting process. In actual situations, patents may convey information about the innovation, so that it becomes easier for a competitor to circumvent the innovation. This effect would diminish the benefit of a patent, and reduce the part of a complex innovation which is protected by a patent. Alternatively, more complex models of information leakage and reverse engineering could be employed to analyze more precisely the optimal fragmentation of an innovation. A second direction in which the analysis could be extended is to analyze the incentive to engage in R & D before the innovation is discovered. In a symmetric model among two competing firms, we could study how different regimes of patent protection affect the incentives to invest in R & D and how they shape the dynamics of the patent race. Finally, we have assumed that the initial competitor does not license his innovation to his competitor. The introduction of licensing agreements would enrich the comparison of different regimes of patent protection in an interesting way.

8 Appendix

8.1 Proof of Proposition 1

If P is concave, and $\underline{\lambda} < \lambda_0 < \bar{\lambda}$, the optimal fraction σ^* is interior and given by the solution to the first order condition:

$$G(\sigma) \equiv \frac{P_2(1, \sigma)}{P_2(1, 1 - \sigma)} = K.$$

Because P is concave in σ_2 , G is increasing in σ . To show that σ^* is increasing in λ , T and r , it suffices to show that K is increasing in λ , T and r . First,

$$\frac{dK}{d\lambda} = \frac{1}{r + \lambda_0} e^{(r+\lambda)T} (1 - \lambda_0 + \lambda) > 0.$$

where the last inequality is due to the fact that $\lambda_0 \leq 1$ as the function $f(r) = \frac{re^{-rT}}{1-e^{-rT}}$ is smaller than 1 for any $r \in \mathfrak{R}_+$. Note also that

$$\frac{dK}{dT} = \lambda e^{T(r+\lambda)} (r + \lambda) \frac{1 - e^{-Tr}}{r} > 0,$$

and

$$\frac{dK}{dr} = \lambda e^{(r+\lambda)T} \frac{e^{-rT} + rT - 1}{r^2} > 0,$$

where the last inequality is due to the fact that the function $g(r) = \frac{1-e^{-rT}}{rT}$ is smaller than 1 for any $r \in \mathfrak{R}_+$.

8.2 Proof of Proposition 2

In order to compute the sign of $\frac{\partial V}{\partial \eta}$ for $\eta < T$, we consider the term

$$A(\eta) = e^{-r(T+\eta)} \Delta P^1 - \frac{\lambda}{r} \left((e^{-r\eta} - e^{-rT}) \Delta P^2 + (e^{-rT} - e^{-r(T+\eta)}) \Delta P^1 \right).$$

Deriving A with respect to η and using the definition of λ_0 ,

$$A'(\eta) = \frac{re^{-r\eta}}{\lambda_0 + r} \left[\lambda \Delta P^2 - \lambda_0 \Delta P^1 + \frac{\lambda \lambda_0}{r} (\Delta P^2 - \Delta P^1) \right].$$

Notice that the sign of $A'(\eta)$ is independent of η . Hence, either $A(\theta)$ is increasing over $[0, T]$ or it is decreasing over $[0, T]$. We also compute

$$\begin{aligned} A(0) &= \frac{1 - e^{-rT}}{r} (\lambda_0 \Delta P^1 - \lambda \Delta P^2), \\ A(T) &= e^{-rT} \frac{1 - e^{-rT}}{r} (\lambda_0 - \lambda) \Delta P^1. \end{aligned}$$

Now suppose that $\Delta P^2 > \Delta P^1$. If $A(0) \geq 0$, then $A(T) \geq 0$ and V is increasing over the entire interval $[0, T]$. In addition, as $\lambda_0 > \lambda$, V is increasing over $[T, \infty)$, so the optimal solution is $\eta^* = \infty$. If $A(0) \leq 0$, then $A'(\eta) > 0$. If $\lambda_0 < \lambda$, $A(T) < 0$, so V is decreasing over $[0, T]$. In addition, V is decreasing over $[T, \infty)$ so the optimal solution is $\eta^* = 0$. Finally, if $\lambda_0 \geq \lambda$ but $\lambda_0 \Delta P^1 \leq \lambda \Delta P^2$, as $A'(\eta) > 0$, the function V is convex in $[0, T]$ and attains its maximum at the boundaries, either at 0 or at T . In addition, the function V is increasing over $[T, \infty)$, so the optimal solution is either $\eta^* = 0$ or $\eta^* = \infty$. Computing the value of V at $\eta = 0$ and $\eta = \infty$, we obtain the condition in the proposition.

Next suppose that $\Delta P^2 < \Delta P^1$. Then $A(0) > A(T)$. If $\lambda_0 > \lambda$, $A(T) > 0$, so the function V is increasing over $[0, T]$ and over $[0, \infty)$ and the optimal solution is $\eta^* = \infty$. If $\lambda_0 \Delta P^1 < \lambda \Delta P^2$, $A'(\eta) < 0$ and $A(0) < 0$, so the function V is decreasing over $[0, T]$ and as $\lambda > \lambda_0$, it is also decreasing over $[T, \infty)$. Hence the optimal solution is $\eta^* = 0$. Finally, if $\lambda > \lambda_0$ but $\lambda_0 \Delta P^1 > \lambda \Delta P^2$, the function V is increasing at 0, and decreasing over $[T, \infty)$. This implies that there exists an interior maximum $\eta^* \in (0, T)$.

8.3 Properties of the function $H(\sigma)$

We first show that H is increasing. We compute

$$H'(\sigma) = \frac{P_2(1, \sigma) \Delta P^2 - P_2(1, 1 - \sigma) \Delta P^1}{(P(1, 0) - P(1, 1 - \sigma))^2}.$$

Hence, $H'(\sigma) > 0$ if and only if

$$\frac{P(1, \sigma) - P(1, 1)}{P(1, 0) - P(1, 1 - \sigma)} > \frac{P_2(1, \sigma)}{P_2(1, 1 - \sigma)}.$$

By concavity of P in σ_2 ,

$$\frac{P(1, \sigma) - P(1, 1)}{P(1, 0) - P(1, 1 - \sigma)} > 1.$$

If $\sigma \geq \frac{1}{2}$, by concavity of P in σ_2 ,

$$1 > \frac{P_2(1, \sigma)}{P_2(1, 1 - \sigma)},$$

and the inequality is true. Suppose now that $\sigma \leq \frac{1}{2}$. Then,

$$\begin{aligned}
& (V_2(1, \sigma) - V_2(1, 1 - \sigma))(V(1, 0) - V(1, 1 - \sigma)) \\
& + V_2(1, 1 - \sigma)(V(1, 0) - V(1, 1 - \sigma)) \\
& - V_2(1, 1 - \sigma)(V(1, \sigma) - V(1, 1)) \\
= & \underbrace{(V_2(1, \sigma) - V_2(1, 1 - \sigma))}_{+} \underbrace{(V(1, 0) - V(1, 1 - \sigma))}_{+} \\
& + \underbrace{V_2(1, 1 - \sigma)}_{-} \underbrace{[(V(1, 0) - V(1, 1 - \sigma)) - (V(1, \sigma) - V(1, 1))]}_{-}
\end{aligned}$$

which is positive, implying again that $H'(\sigma) > 0$.

Clearly, $H(0) = 1$. We compute $H(1)$ using L'Hospital rule:

$$H(1) = \frac{V_2(1, 1)}{V_2(1, 0)}.$$

Recall that $\bar{\lambda}$ is defined implicitly by the equation:

$$\frac{V_2(1, 0)}{V_2(1, 1)} = K^{-1}(\bar{\lambda}) \Leftrightarrow H(1) = K(\bar{\lambda}), \text{ with } K(\lambda) = 1 + \frac{(\lambda - \lambda_0)}{(r + \lambda_0)e^{-(r+\lambda)T}}.$$

We show that $\lambda_0 H(1) > \bar{\lambda}$. As $K(\lambda)$ is an increasing function of λ , this is equivalent to showing that $K(\lambda_0 H(1)) > K(\bar{\lambda})$, or

$$1 + \frac{\lambda_0(H(1) - 1)}{(r + \lambda_0)e^{-(r+\lambda_0 H(1))T}} > H(1) \Leftrightarrow \frac{\lambda_0}{r + \lambda_0} = e^{-rT} > e^{-(r+\lambda_0 H(1))T},$$

which is always satisfied.

8.4 Proof of Proposition 4

We consider the term

$$B(\eta) = e^{-r(T+\eta)}\Delta P^1 - \frac{\lambda}{r} \left((e^{-r\eta} - e^{-rT})\Delta P^4 + (e^{-rT} - e^{-r(T+\eta)})\Delta P^3 \right).$$

Deriving B with respect to η and using the definition of λ_0 ,

$$B'(\eta) = \frac{re^{-r\eta}}{\lambda_0 + r} \left[\lambda\Delta P^4 - \lambda_0\Delta P^1 + \frac{\lambda\lambda_0}{r}(\Delta P^4 - \Delta P^3) \right].$$

Notice that the sign of $B'(\eta)$ is independent of η . Hence, either $B(\theta)$ is increasing over $[0, T]$ or it is decreasing over $[0, T]$. We also compute

$$\begin{aligned}
B(0) &= \frac{1 - e^{-rT}}{r}(\lambda_0\Delta P^1 - \lambda\Delta P^4), \\
B(T) &= e^{-rT} \frac{1 - e^{-rT}}{r}(\lambda_0\Delta P^1 - \lambda\Delta P^3).
\end{aligned}$$

Now suppose that $\Delta P^4 > \Delta P^3$. Then, if $B(0) > 0$, then $B(T) > 0$ and V is increasing over the entire interval $[0, T]$. In addition, as $B(T) > 0$, $\lambda_0 \Delta P^1 - \lambda \Delta P^3 > 0$ and V is increasing over $[T, \infty)$, so the optimal solution is $\eta^* = \infty$. If $B(0) \leq 0$, then $B'(0) > 0$. If $\lambda_0 \Delta P^1 < \lambda \Delta P^3$, $B(T) < 0$, so V is decreasing over $[0, T]$. In addition, V is decreasing over $[T, \infty)$ so the optimal solution is $\eta^* = 0$. Finally, if $\lambda_0 \Delta P^1 \geq \lambda \Delta P^3$ but $\lambda_0 \Delta P^1 \leq \lambda \Delta P^4$, as $B'(0) > 0$, the function V is convex in $[0, T]$ and attains its maximum at the boundaries, either at 0 or at T . In addition, the function V is increasing over $[T, \infty)$, so the optimal solution is either $\eta^* = 0$ or $\eta^* = \infty$. Computing the value of V at $\eta = 0$ and $\eta = \infty$, we obtain the condition in the proposition.

Next suppose that $\Delta P^4 < \Delta P^3$. Then $B(0) > B(T)$. If $\lambda_0 \Delta P^1 > \lambda \Delta P^3$, $B(T) > 0$, so the function V is increasing over $[0, T]$ and over $[0, \infty)$ and the optimal solution is $\eta^* = \infty$. If $\lambda_0 \Delta P^1 < \lambda \Delta P^4$, $B'(0) < 0$ and $B(0) < 0$, so the function V is decreasing over $[0, T]$ and as $B(T) < 0$, $\lambda_0 \Delta P^1 < \lambda \Delta P^3$, and V is also decreasing over $[T, \infty)$. Hence the optimal solution is $\eta^* = 0$. Finally, if $\lambda > \lambda_0$ but $\lambda \Delta P^4 < \lambda_0 \Delta P^1 < \lambda \Delta P^3$, the function V is increasing at 0, and decreasing over $[T, \infty)$. This implies that there exists an interior maximum $\eta^* \in (0, T)$.

References

- [1] Anton, J. J., Greene, H., and Yao, D. A. (2006). Policy Implications of Weak Patent Rights. In A. Jaffe, J. Lerner, and S. Stern (eds). Innovation Policy and the Economy. Cambridge (MA): MIT Press.
- [2] Anton, J. J., and Yao, D. A. (2004). Little Patents and Big Secrets: Managing Intellectual Property. RAND Journal of Economics 35, 1-22.
- [3] Arora, A. (1997). Patents, Licensing, and Market Structure in the Chemical Industry. Research Policy 26, 391-403.
- [4] Cohen, W. M., and Levinthal; D. A. (1989). Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35, 128-152.
- [5] Cohen, W. M., Nelson, R. R., and Walsh, J. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). NBER Working Paper 7552.

- [6] Denicolo, V. and Franzoni, L.A. (2004). Patents, Secrets, and the First-Inventor Defense. *Journal of Economics and Management Strategy* 13, 517-538.
- [7] Friedman, D.D., Landes, W.M. and Posner, R.A. (1991). Some Economics of Trade Secret Law. *Journal of Economic Perspectives* 5, 61-72.
- [8] Gallini, N. (1992). Patent Length and Breadth with Costly Imitation. *RAND Journal of Economics* 23, 52-63.
- [9] Graham, S. (2004). Hiding in the Patents Shadow: Firms Uses of Secrecy to Capture Value from New Discoveries. Mimeo. Available at <https://smartech.gatech.edu/xmlui/handle/1853/10725>.
- [10] Graham, S., and Hegde, D. (2012). Do Inventors Value Secrecy in Patenting? Evidence from the American Inventors Protection Act of 1999. Mimeo. Available at http://economics.stanford.edu/files/Graham12_4.pdf.
- [11] Hall, B. H., Helmers, Ch., Rogers, M., and Sena, V. (2014). The Choice Between Formal and Informal Intellectual Property: A Review. *Journal of Economic Literature* 52, 1-50.
- [12] Horstmann, I., MacDonald, G. M., and Slivinski, A. (1985). Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent. *Journal of Political Economy* 93, 837-58.
- [13] Hussinger, K. (2006). Is Silence Golden? Patents versus Secrecy at the Firm Level. *Economics of Innovation and New Technology* 15, 735-752.
- [14] Jorda, K.F. (2007). Trade Secrets and Trade-Secret Licensing. In Krattiger, A., Mahoney, R.T., Nelsen, L. et al. (Eds). *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices*, Chapter 11.5. Oxford: MIHR & PIPRA.
- [15] Kappos, D.J., and Stanek Rea, T. (2012). Report on the Prior User Rights Defense. The United States Patent and Trademark Office. <http://www.aiarulemaking.com/rulemaking-topics/group-1/documents/PriorUserRightsReporttoCongress.pdf>
- [16] Kultti, K., Takalo, T., and Toikka, J. (2007). Secrecy versus Patenting. *RAND Journal of Economics* 38, 22-42.

- [17] Kwon, I. (2012a). Patent Thicket, Secrecy, and Licensing. *The Korean Economic Review* 28, 27-49.
- [18] Kwon, I. (2012b). Patent Races with Secrecy. *The Journal of Industrial Economics* 60, 499-516.
- [19] Lemley, M. A., and Shapiro, C. (2005). Probabilistic Patents. *Journal of Economic Perspectives* 19, 75-98.
- [20] Lerner, J., and Tirole, J. (2004). Efficient Patent Pools. *American Economic Review* 94, 691–711.
- [21] Lévêque, F, and Ménière, Y. (2007). Patents and Innovation: Friends or Foes. *Berkeley Center for Law and Technology Working Paper* 28.
- [22] Levin, R. C., Klevorick, A. K., Nelson, R. R. and Winter, S. G. (1987). Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity* 3: 783-831.
- [23] Maier, R.L. (2011). The Big Secret of the America Invents Act. *Intellectual Property Today*, December.
- [24] Ottoz, E. and Cugno, F. (2008). Patent-Secret Mix in Complex Product Firms. *American Law and Economics Review* 10, 142-158.
- [25] Ottoz, E. and Cugno, F. (2011). Choosing the Scope of Trade Secret Law when Secrets Complement Patents. *International Review of Law and Economics* 31, 219–227.
- [26] Perng Pan, S. and Mion, S. (2010). Hybrid Use of Trade Secret and Patent Protection in Green Technology. *Bloomberg Law Reports* 3 (4).
- [27] Quinn, G. (2012). Patentability Overview: When can an Invention be Patented? IPWatchdog, www.ipwatchdog.com (last consulted on August 20, 2013).
- [28] Schneider, C. (2008). Fences and Competition in Patent Races. *International Journal of Industrial Organization* 26, 1348–1364.
- [29] Scotchmer, S., and Green, J. (1990). Novelty and Disclosure in Patent Law. *RAND Journal of Economics* 21, 131–146.

- [30] Shapiro, C. (2001). Navigating the Patent Thicket: Cross-Licenses, Patent-Pools, and Standard-Setting. In Jaffe, A., Lerner, J., and Stern, S. (eds.), *Innovation Policy and the Economy*. Vol. 1. Cambridge, MA: MIT Press.
- [31] Shapiro, C. (2006). Prior User Rights. *American Economic Review* 96, 92-96.
- [32] Takalo, T. (1998). Innovation and Imitation under Imperfect Patent Protection. *Journal of Economics* 67, 229–241.
- [33] Tarazano, L. (2013). Can Recipes Be Patented? *Inventor's Eye* (Vol. 4, Issue 3). Available at www.uspto.gov/inventors/independent/eye/201306/ADVICE.jsp.
- [34] Tegernsee Experts Group (2012). Report on Prior User Rights. Munich: European Patent Office. Available at www.epo.org/news-issues/news/2012/20121108a.html.