



HAL
open science

Choosing Intellectual Protection: Imitation, Patent Strength and Licensing

David Encaoua, Yassine Lefouili

► **To cite this version:**

David Encaoua, Yassine Lefouili. Choosing Intellectual Protection: Imitation, Patent Strength and Licensing. 2006, pp.1-31. halshs-00177619

HAL Id: halshs-00177619

<https://shs.hal.science/halshs-00177619>

Submitted on 10 Oct 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Choosing intellectual protection: imitation, patent strength and licensing

David Encaoua* and Yassine Lefouili**

March 22, 2006 (revised version)

Abstract

Patents are probabilistic rights. We set up a multi-stage model in which choosing between patent and trade secrecy is affected by three parameters : the *patent strength* defined as the probability that the right is upheld by the court, the *cost of imitating* a patented innovation relative to the cost of imitating a secret innovation, and the *innovation size* defined as the extent of the cost reduction. The choice of the protection regime is the result of two effects: the damage effect evaluated under the unjust enrichment doctrine and the effect of market competition that occurs under the shadow of infringement. We find that large innovations are likely to be kept secret whereas small innovations are always patented. Furthermore, medium innovations are patented only when patent strength is sufficiently high. Finally, we investigate a class of patent licensing agreements used to settle patent disputes between patent holders and their competitors.

Keywords : *patent, trade secrecy, imitation, licensing*

JEL Classification : *D45, L10, O32, O34*

⁰*University of Paris I Panthéon Sorbonne (CES); e-mail: encaoua@univ-paris1.fr

**University of Paris I Panthéon Sorbonne (CES); e-mail: yassine.lefouili@m4x.org

We thank two anonymous referees for their constructive suggestions.

1 Introduction

The traditional view that firms always prefer patents to other forms of protection for their innovations has been empirically challenged for a long time. It is now well known that trade secrecy, first mover advantage and exploitation of lead time may be preferred forms of protection, at least in some industries. Even in the same industry, forms of protection may differ according to the nature and importance of the innovation and to the disclosure effect. Early studies by Scherer (1965, 1967, 1983) have shown that the propensity to patent varies significantly across industries and that interindustrial variations in patenting activity are not explained by R&D expenditures. Pakes and Griliches (1980) were among the first to find that the degree of randomness in the patenting activity within industries was not explained by R&D variations. They have shown that the residual patenting behaviour was explained by the potential imitation allowed by the disclosed information and by the innovator's capability to appropriate the rents generated by the innovation. Mansfield (1986) obtained similar results based on a survey where US manufacturing firms were asked what fractions of inventions they would not have developed in the absence of patents between 1980 and 1983. These fractions were very low in many industries (less than 10% in electrical equipment, primary metals, instruments, motor vehicles and others) and relatively high in industries like pharmaceuticals (60%) and chemicals (40%). Two more recent surveys (Yale Survey by Levin *et al.*, 1987 and Carnegie Mellon Survey by Cohen *et al.* 2000) confirm these trends: it is only in industries where knowledge is strongly codified that patents appear to be substantially preferred to other forms of protection.

Despite this accumulated empirical evidence, theoretical explanations of why and when an innovator would prefer to keep an innovation secret rather than to patent it remain rather scarce. Before turning to the related literature, note that even if patenting is not considered as the best form of protection, innovators have a lot of reasons to apply for patents serving purposes different from protection (Hall and Ziedonis, 2001, Encaoua *et al.* 2004). This feature complicates the problem. Indeed, a theoretical explanation of why and when patenting is not the best form of protection must also be compatible with the more cumbersome issue of why, despite the existence of preferred forms of protection, patents remain so widespread (Scotchmer, 2004).

At least three types of theoretical arguments are required to explain the protection choice.

First, patents must be recognized as not being ironclad property rights but rather *probabilistic rights*. If patents offered perfect protection against any imitation, there would be no doubt about the best protection regime. Lemley and Shapiro (2005) qualify the uncertain intellectual property rights in a suggestive way: "*A patent does not confer upon its owner the right to exclude a potential imitator but rather a right to try to exclude by asserting the patent in court. When a patent holder asserts its patent against an alleged infringer, the patent is*

rolling the dice. If the patent is found invalid, the property right will have evaporated". Thus patent strength refers to the probability to recover damages, with the consequence that only strong patents give in principle to the patent holder the right to exclude an infringer or to force him to buy a license. But as we argue below, even holders of weak patents may escape the uncertain litigation process. They succeed through their licensing strategies to capture a significant part of the consumer's surplus. This is why the notion of patent quality enters so forcefully in the agenda of antitrust authorities nowadays, especially through criticism of the examination system of the Patent Office (Merges, 1999, Lemley, 2001).

Second, the traditional view that knowledge is a blueprint has also been challenged. Replicating an existing invention may be costly and time consuming because knowledge is more or less embedded in individuals and firms rather than in physical products or equipment. One has to distinguish innovations according to their *secrecy effectiveness*, which is the main determinant of the imitation cost (Anton *et al.* 2005b). Many innovations involve hidden know-how even if the allowed performance is perfectly observable. Consider for instance a process innovation leading to a cost reduction that is reflected back in the market price but for which the technological knowledge is neither perfectly revealed nor easily reverse-engineered. In this case, imitating the process innovation or building around it may be rather costly. Moreover, the imitation cost may depend on whether the invention has been patented or not. As the patent discloses some enabling technological information, it is clear that imitating a patented innovation should be at most as costly as if it was kept secret. Our paper offers a natural framework to analyze the classical tradeoff between getting a legal protection involving a compulsory disclosure of enabling information and keeping secrecy by giving up legal protection.

Third, even if patents do not always appear as the best form of protection, innovators may nevertheless prefer to patent their innovations because holding a patent offers the possibility to settle a dispute against an alleged infringer through a licensing agreement (Farrell and Shapiro, 2005, Lemley and Shapiro, 2005). Alleged infringers may also prefer to avoid a litigation process not only because litigation is costly but also because winning the lawsuit against the patent holder involves a free-riding aspect, as other competitors benefit from the asserted patent's invalidity. Therefore, even when they are weak, patents generate substantial revenues through licensing royalties that may harm consumers. This is why patent settlements raise serious concerns for competition policy authorities (Shapiro, 2003, Encaoua and Hollander, 2004).

The main objective of this paper is to introduce these arguments in a simple model allowing a discussion of the following issues:

- i/ What are the different forces that interact in the choice of a protection regime (patent or secrecy)?
- ii/ How are these forces affected by the patent strength, imitation cost and innovation

size?

iii/ What sort of licensing agreements are likely to emerge in order to avoid patent litigation?

Two main contributions have explicitly explored the decision whether to patent an innovation or not.¹ Hortsman *et al.* (1985) assume that an innovating firm possesses private information about profits available to competitors and that patent coverage may not exclude profitable imitation. Conceived as an information transfer mechanism, a patent that covers full information is not optimal. The optimal innovator's choice is a mixed strategy between patenting and keeping secrecy while the follower's optimal choice is to stay out of the market when the innovator patents and to imitate when the innovator does not. The peculiarities of this model, in terms of the signaling aspect of the patent and the *a priori* restrictions put on the follower's action, explain why imitation of a patented innovation does not occur in equilibrium. Since our paper is close to Anton and Yao (2004), we describe more thoroughly their framework. Starting from the premise that disclosure provides competitors with usable information and focusing on the innovator's decision about how much of an innovation should be disclosed, their model is particularly relevant for a special type of secrecy effectiveness. They describe a situation where the real innovation performance is not directly observable while the disclosed know-how enables a competitor to costlessly replicate it. Therefore, by choosing the amount to be disclosed, the innovator directly controls the behaviour of the potential imitator. Their model is a signalling game where the innovator has private information on the innovation size and decides to reveal partially or fully this information, letting the potential imitator infer the leader's advance. The follower chooses either to imitate or not under the risk of infringement. A refined perfect bayesian equilibrium of the signalling game involves a separating strategy in which: i/ small innovations are patented and fully disclosed; ii/ medium innovations are patented and partially disclosed; iii/ large innovations are kept secret and partially disclosed through a public announcement. This result is illustrated by their suggestive title: "little patents and big secrets".

In our model, we maintain the general tradeoff with which an innovator is confronted, when choosing the protection regime. However, rather than focusing on the signaling aspect we assume that: i/ the process innovation size, measured by the cost reduction, is directly observable; ii/ a patent reveals technological information that lowers the imitation cost relatively to the situation where the innovation is kept secret.² Choosing to patent may expose

¹Among other papers related to the choice of an intellectual property regime, one can include Crampes (1986), Gallini (1992) and Scotchmer and Greene (1990). Crampes examines the tradeoff between keeping secret an invention during an indefinite time or obtaining a legal protection over a finite duration (the statutory patent life). Gallini introduces the idea that breadth governs the cost of inventing around the patent. However, it is entry cost rather than imitation cost that matters. Scotchmer and Greene (1990) focuses on the impact of patent policy on the incentives to innovate. Their model involves a binary choice as the innovation would not be realized if it were not patented. They also assume full disclosure of technological know-how.

²If the reduction of the imitation cost directly depends on the disclosed level of enabling knowledge, using

the innovator either to an increased imitation level or to a lower one because the imitation level does not only depend on the imitation cost but also on two other crucial parameters: the innovation size and patent strength. It may happen that an innovator benefits from being imitated : this occurs whenever the incurred loss due to imitation is overcompensated by the damages it receives from an imitator if the court upholds the patent validity *and* the patent infringement. If patenting or keeping secret the process innovation were leading to the same imitation level, then patenting would be preferred since damages are expected under the patent regime. This corresponds to the *damage effect*. But as soon as imitation levels differ according to the protection regime, a conflict arises as long as imitation becomes higher under the patent regime. This corresponds to the *competition effect*. Therefore, as soon as the imitation extent is decided by the follower, different interactions may occur between the competition effect and the damage effect. Our paper aims to clarify these interactions. We propose a complete information multistage game in which three common knowledge parameters are important: the innovation size, the patent strength and the relative cost of imitation. We depart from the assumption limiting *a priori* an imitator's behaviour by letting it choose its own cost reduction in response to the process innovation. We also specify that expected damages paid by the potential infringer correspond to the infringer's profit, the risk of infringement being directly linked to the patent strength. Our damages specification only presents a slight difference with the infringer's revenue retained in Anton and Yao (2004).³ Thus our results may be compared to the "little patents and big secrets" results in Anton and Yao.

The main results of our approach are as follows. For a given innovation size, patent strength and relative cost of imitation generally act as strategic substitutes. An increase of one of these parameters must be compensated by a decrease of the other in order to keep the same value of the innovator's profit. Inventors of small process innovations always prefer patent protection to trade secrecy. This reminder of the "little patents" result by Anton and Yao (2004) rests however on a different argument in our model. For large process innovations, our results present some difference with the "big secrets" characterization in Anton and Yao. Our model does not totally discard the possibility of patenting some large process innovations, whenever imitation is too costly. This may happen when information is poorly disclosed in the patent. In this case the innovator is indifferent between secret and patent. For medium process innovations, our results differ more significantly from those of Anton and Yao. It is not optimal for a firm producing such an innovation to file a patent of bad quality, that is a patent

a relative imitation cost parameter is equivalent to using a disclosure level. It appears however that working with the relative imitation cost parameter is more convenient since the extent of imitation remains controlled by the imitator, while the choice of a protection regime is made by the innovator.

³In a more recent paper, Anton and Yao (2005a) introduce the "lost" profits of the patentee, defined as the profits that would have occurred in the absence of infringement. They show that at equilibrium, infringement may take one of two forms: a "*passive*" form in which lost profits of the patentee are zero and an "*agressive*" form where they are positive. One of the main results in Anton and Yao (2005a) is that infringement always occurs when damages equal lost profits. This last result does not hold in our model.

having a low probability to be upheld by the court, unless the disclosed information does not significantly lower the imitation cost. We show that there exists a *safe protection level* that is sufficient to deter any imitation and that this level is lower than a 100% protection. As the innovation size decreases, trade secrecy is less likely. Finally, the "one size fits all" principle in the patent design is not validated by our analysis.

These results raise many practical issues. While the model predicts that it is seldom optimal for a firm to file a patent when the probability that it will be upheld by the court is low, bad quality patents (relatively to novelty and non-obviousness requirements) are widespread in real world. How can this be explained? Moreover, why bad quality patents are not litigated more often than we do observe? These issues are at the heart of what has been called the *patent paradox* (Hall and Ziedonis, 2001, Scotchmer, 2004). We devote a brief analysis that suggests a possible answer to explain this paradox. Whenever a patent is not conceived only as a protection against imitation but also as a tool to reach private settlements through licensing agreements (Shapiro, 2003), licensing agreements may act as an alternative to patent litigation. A royalty rate independent of the patent strength combined to a specific fixed fee may serve this purpose.

The model is presented in section 2. The market competition outcome under the shadow of infringement is described in section 3. The imitator's behaviour is analyzed in section 4. The core of the paper, which corresponds to the protection regime choice is examined in section 5. We devote section 6 to licensing agreements. Our conclusions are presented in section 7.

2 The basic set-up

We examine a process innovation in a framework involving two competing firms. We suppose that firm 1 is an innovating firm and firm 2 is a possible imitator. Each firm is risk-neutral and seeks to maximize its expected profit. Initially, both firms produce at the same marginal cost $c > 0$. Fixed production costs are assumed equal to zero. We assume that firm 1 undertakes an R&D investment which allows to reduce the marginal cost to the level $d_1 < c$. The game we study hereafter starts once the innovation is introduced and involves three stages.

First, in the *protection stage*, the innovator has to choose between two protection regimes. The first regime, which we denote by P, is to patent its innovation, and the second regime, which we denote by S, consists in protecting its innovation by the means of trade secrecy.

Second, in the *imitation stage*, after the observation of the innovator's marginal cost d_1 , firm 2 chooses to imitate (or "build around") the innovator's technology. It imitates the innovation by transforming its old technology onto a *follow-up technology* which allows it to reduce its marginal cost to $d_2 \in [d_1, c]$. Note that we do not allow the imitator to improve the innovator's technology. The difference $c - d_2$ represents the "extent of imitation". When $d_2 = c$, there is no imitation at all and when $d_2 = d_1$, imitation is full. We assume that

imitation at a level $d_2 \in [d_1, c[$ induces a fixed imitation cost $I(d_2)$ which depends on whether the innovation is patented or kept secret. Precisely, we assume that the imitation cost under the patent regime, which we denote by $I^P(d_2)$, and the imitation cost under the secrecy regime, which we denote by $I^S(d_2)$, satisfy the following condition :

$$I^P(d_2) = fI^S(d_2)$$

where the parameter $f \geq 0$ measures the relative costs of imitation under the regimes P and S. We assume that $f \leq 1$: since patenting involves a compulsory disclosure, it is likely that imitating a patented innovation turns out to be less costly than imitating a secret innovation.

Third, in the *competition stage*, market outcomes are determined under the shadow of punishment. We assume that, when the innovation is patented, firm 1 sues firm 2 for infringement if firm 2 chooses a follow-up technology that allows to produce at a marginal cost $d_2 < c$. We also assume that firm 2 systematically contests the validity of the patent covering the innovation. We denote by e the probability that an imitation infringes the innovator's patent. This probability can be interpreted as an indicator of the lagging patent's *breadth* : the broader the patent's breadth, the higher the probability that a follow-up technology that reduces the marginal cost c to $d_2 \in [d_1, c[$ is an infringement of the patent on the process innovation d_1 . We denote by g the probability that the patent survives the imitator's legal contestation of the patent's validity. We interpret this parameter as the *patent's quality* : low quality patents have higher chances to be invalidated by a court than high quality patents. Thus, a higher quality patent (in terms of novelty and inventiveness) is less uncertain in the sense that the probability that a court will uphold its validity is higher. Firm 2 is compelled to pay damages, supposed to be equal to its market profit, if and only if the patent is held valid and the imitation infringes the patent. This occurs with probability $\theta = eg$ (we assume that the issues of validity and infringement are independent). The parameter $\theta \in [0, 1]$ corresponds to what is called the *patent strength*. When the innovation is not patented, no damages are paid.

Following Anton and Yao (2004), we model our duopoly market competition as a traditional Cournot competition (quantity setting) with linear market demand:

$$p(x_1 + x_2) = a - (x_1 + x_2)$$

where x_1 is the output of firm 1, x_2 is the output of firm 2 and $p(x_1 + x_2)$ is the market clearing price.

We assume that $c < a < 2c$. The first inequality is usual and means that the marginal cost before innovation is below the choke price. The second inequality expresses that the market is small which is a likely scenario for innovative markets, as it allows the possibility that the

innovative firm becomes at least twice as efficient as it currently is⁴. In other words, the inequality $a < 2c$ implies that there exist innovations d_1 such that $d_1 < 2c - a$, which can also be written as $a - d_1 > 2(a - c)$.

We choose a convex specification for the imitation technology and, to reach analytical results, we use a quadratic expression :

$$I(d_2) = \begin{cases} f \frac{(c-d_2)^2}{2} & \text{if } Patent \\ \frac{(c-d_2)^2}{2} & \text{if } Secret \end{cases}$$

3 Competition stage

The competition stage occurs under the shadow of litigation only if the innovation is patented. Therefore, the outcome of the competition stage depends on whether the innovation is patented or not.

3.1 Patented innovation

We separately examine the cases $d_2 < c$ (the follower imitates the innovator, at least partially) and $d_2 = c$ (the follower does not imitate the innovator), as the profit functions differ in these two cases.

The follower imitates ($d_2 < c$) : Under regime P , the expected gross profits of firm 1 and firm 2 are given by:

$$\Pi_1^P(x_1, x_2, d_1, d_2, \theta) = (a - (x_1 + x_2) - d_1)x_1 + \theta(a - (x_1 + x_2) - d_2)x_2$$

and

$$\Pi_2^P(x_1, x_2, d_2, \theta) = (1 - \theta)(a - (x_1 + x_2) - d_2)x_2$$

From the expected profits, one derives the Cournot-Nash equilibrium outputs $x_1^P(d_1, d_2, \theta)$ and $x_2^P(d_1, d_2, \theta)$. They correspond either to an interior solution where both firms are active : $x_1^P(d_1, d_2, \theta) \geq x_2^P(d_1, d_2, \theta) > 0$ or to a boundary solution where only firm 1 is active : $x_1^P(d_1, d_2, \theta) > x_2^P(d_1, d_2, \theta) = 0$.

Consider first an interior solution. Routine computations lead to:

$$x_1^P = \frac{a(1 - \theta) + d_2(1 + \theta) - 2d_1}{3 - \theta}$$

$$x_2^P = \frac{a - 2d_2 + d_1}{3 - \theta}$$

⁴We acknowledge one of the referees for this suggestion.

Hence, a necessary and sufficient condition for an interior solution to exist is :

$$d_2 < \frac{a + d_1}{2} \quad (1)$$

Note that this condition is always satisfied when $d_1 > 2c - a$. Note also that the market price $p^P(\theta)$ is given by $p^P(\theta) = \frac{a + d_2(1-\theta) + d_1}{3-\theta}$ which is increasing in $\theta \in [0, 1]$ as long as condition (1) is satisfied.

Consider now a boundary solution. Such a solution arises when condition (1) is not satisfied and is characterized by:

$$\begin{aligned} x_1^P &= \frac{a - d_1}{2} \\ x_2^P &= 0 \end{aligned}$$

The follower does not imitate ($d_2 = c$) : The equilibrium outputs in this case can be derived from those of the previous case by taking $\theta = 0$ and $d_2 = c$. Hence:

- If $d_1 > 2c - a$, then the equilibrium outputs are given by:

$$\begin{aligned} x_1^P &= \frac{a + c - 2d_1}{3} \\ x_2^P &= \frac{a + d_1 - 2c}{3} \end{aligned}$$

- If $d_1 \leq 2c - a$, then we have the same boundary solution as in the imitation case:

$$\begin{aligned} x_1^P &= \frac{a - d_1}{2} \\ x_2^P &= 0 \end{aligned}$$

Summing up all these cases, the expected equilibrium gross profits depend on d_1 , d_2 and θ in the following way :

• If $d_1 \leq 2c - a$ then:

$$\Pi_1^P(d_1, d_2, \theta) = \begin{cases} \frac{[a - d_1(2-\theta) + d_2(1-\theta)][a(1-\theta) - 2d_1 + d_2(1+\theta)] + \theta[a - 2d_2 + d_1]^2}{(3-\theta)^2} & \text{if } d_2 < \frac{a+d_1}{2} \\ \frac{(a-d_1)^2}{4} & \text{if } \frac{a+d_1}{2} \leq d_2 \leq c \end{cases} \quad (2)$$

$$\Pi_2^P(d_1, d_2, \theta) = \begin{cases} (1-\theta) \frac{(a-2d_2+d_1)^2}{(3-\theta)^2} & \text{if } d_2 < \frac{a+d_1}{2} \\ 0 & \text{if } d_2 \geq \frac{a+d_1}{2} \end{cases} \quad (3)$$

- If $d_1 > 2c - a$ then:

$$\Pi_1^P(d_1, d_2, \theta) = \begin{cases} \frac{[a-d_1(2-\theta)+d_2(1-\theta)][a(1-\theta)-2d_1+d_2(1+\theta)]+\theta[a-2d_2+d_1]^2}{(3-\theta)^2} & \text{if } d_2 < c \\ \frac{(a+c-2d_1)^2}{9} & \text{if } d_2 = c \end{cases} \quad (4)$$

$$\Pi_2^P(d_1, d_2, \theta) = \begin{cases} (1-\theta)\frac{(a-2d_2+d_1)^2}{(3-\theta)^2} & \text{if } d_2 < c \\ \frac{(a+d_1-2c)^2}{9} & \text{if } d_2 = c \end{cases} \quad (5)$$

Therefore, under Cournot competition, firm 2 is driven out of the market if it keeps its old technology when the innovation is *large enough* ($d_1 < 2c - a$) and remains active on the market (even without imitating firm 1) when the cost reduction innovation is *small enough* ($d_1 > 2c - a$). This result depends on the small market assumption ($a < 2c$). In a large market ($a > 2c$), firm 2 would remain in the market whatever the innovation size. Thus, the small market assumption captures the strategic aspect in a more complete way.

3.2 Unpatented innovation

Equilibrium outcomes under trade secret regime are derived from those under the patent regime by taking $\theta = 0$. This simply means that no damages are paid when imitation occurs under secrecy. One obtains :

$$\Pi_1^S(d_1, d_2) = \begin{cases} \frac{(a-2d_1+d_2)^2}{9} & \text{if } d_2 < \frac{a+d_1}{2} \\ \frac{(a-d_1)^2}{4} & \text{if } d_2 \geq \frac{a+d_1}{2} \end{cases} \quad (6)$$

$$\Pi_2^S(d_1, d_2) = \begin{cases} \frac{(a-2d_2+d_1)^2}{9} & \text{if } d_2 < \frac{a+d_1}{2} \\ 0 & \text{if } d_2 \geq \frac{a+d_1}{2} \end{cases} \quad (7)$$

4 Imitation stage

Firm 2 aims to maximize its net profit when it chooses its imitation level $d_2 \in [d_1, c]$. Since the follower's gross profit and imitation cost depend on whether the innovation is patented or not, we have to distinguish these two regimes.

4.1 Patented innovation

Under this regime, the imitator's net profits when it chooses a follow-up technology allowing to produce at marginal cost $d_2 \in [d_1, c]$ is given by :

$$G_2^P(d_1, d_2, f, \theta) = \Pi_2^P(d_1, d_2, \theta) - \frac{1}{2}f(c - d_2)^2$$

The follower's optimal imitation level when the innovation is patented is determined as:

$$d_2^P(d_1, f, \theta) = \underset{d_2 \in [d_1, c]}{\text{Arg max}} G_2^P(d_1, d_2, f, \theta)$$

Define $A(\theta) = \frac{1-\theta}{(3-\theta)^2}$. It is a decreasing function of $\theta \in [0, 1]$ such that $A(0) = \frac{1}{9}$ and $A(1) = 0$.

The function $d_2 \rightarrow H(d_1, d_2, f, \theta) = A(\theta)(a - 2d_2 + d_1)^2 - \frac{1}{2}f(c - d_2)^2$, which is the expression of $G_2^P(d_1, d_2, f, \theta)$ when $d_2 < \frac{a+d_1}{2}$ and $d_2 < c$, is necessarily either convex or concave over its whole definition domain. The following preliminary results are easy to show:

1- The function $d_2 \rightarrow H(d_1, d_2, f, \theta)$ is strictly convex if $f < 8A(\theta)$ and is strictly concave if $f > 8A(\theta)$.

2- The unconstrained extremum of $d_2 \rightarrow H(d_1, d_2, f, \theta)$ is easily obtained by the the FOC:

$$d_2^{int}(d_1, f, \theta) = c + \frac{4A(\theta)(d_1 - 2c + a)}{8A(\theta) - f} \quad (8)$$

In order to obtain the value of $d_2^P(d_1, f, \theta)$, it is necessary to know whether $H(d_1, d_2, f, \theta)$ is convex or concave and to compare $d_2^{int}(d_1, f, \theta)$ to d_1 and c . For instance, when $d_1 < 2c - a$ and $f > 8A(\theta)$ (H strictly concave), equation (8) leads to $d_2^{int}(d_1, f, \theta) > c$. Moreover, as $G_2^P(d_1, d_2, f, \theta)$ is a discontinuous function of d_2 for $d_2 = c$, it is necessary to compare the value of $G_2^P(d_1, c, f, \theta)$ obtained in the absence of imitation to the value of $\underset{d_2 \in [d_1, c]}{\text{Arg max}} G_2^P(d_1, d_2, f, \theta)$ obtained with imitation.

4.1.1 Large innovations ($d_1 < 2c - a$)

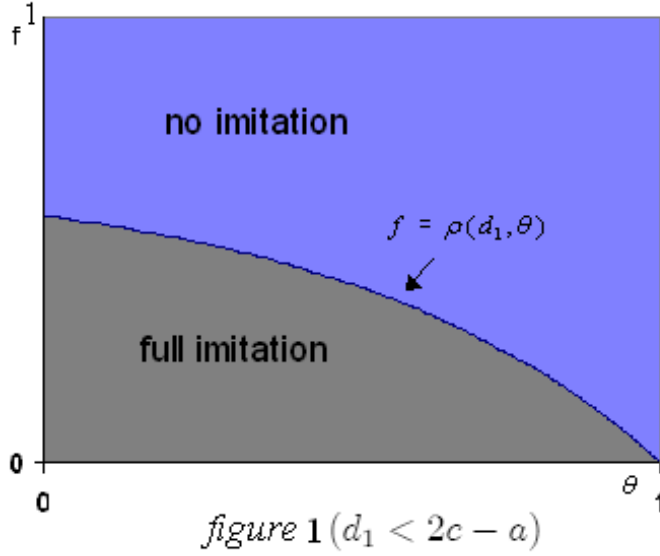
With large innovations, partial imitation never occurs. The following proposition (proof in appendix A1) distinguishes two areas in the (θ, f) space according to whether the optimal imitation level is maximal ($d_2^P(d_1, f, \theta) = d_1$) or minimal ($d_2^P(d_1, f, \theta) = c$). In the (θ, f) space, the extent of these two areas depends on d_1 .

Proposition 1 *For large innovations ($d_1 < 2c - a$), there exists a threshold function $\rho(d_1, \theta) = 2A(\theta) \left(\frac{a-d_1}{c-d_1} \right)^2$ which is decreasing in the patent strength θ and the innovation size $c - d_1$ such that :*

If $f < \rho(d_1, \theta)$, then the follower fully imitates: $d_2^P(d_1, f, \theta) = d_1$.

If $f > \rho(d_1, \theta)$, then the follower does not imitate $d_2^P(d_1, f, \theta) = c$ and is driven out of the market.

The interpretation of this proposition, illustrated in figure 1, is clear. When the process innovation is large enough ($d_1 < 2c - a$), there exists a threshold imitation cost $\rho(d_1, \theta)$ such that if f is below this threshold it is optimal for a follower to fully imitate the patented

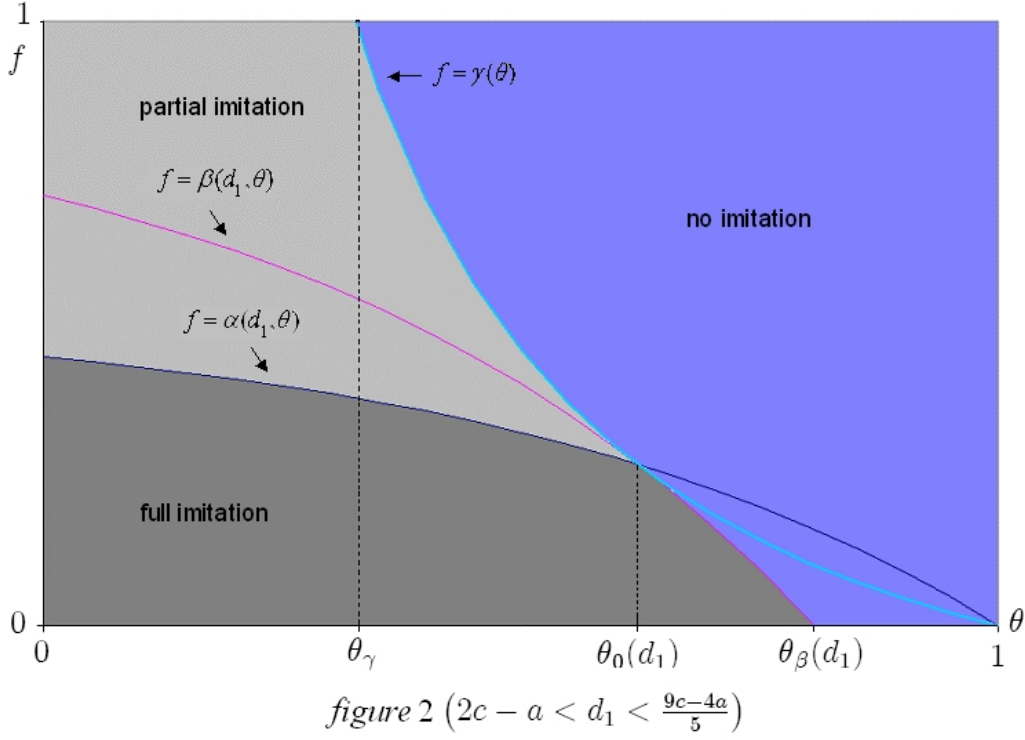


innovation ($d_2^P(d_1, f, \theta) = d_1$), whereas if f is above the threshold, it does not pay to imitate. Note that $\rho(d_1, 0) = \frac{2}{9} \left(\frac{a-d_1}{c-d_1} \right)^2 < \frac{8}{9}$ for any $d_1 < 2c - a$. Therefore, sufficiently large patented innovations, even if they are protected with a weak patent (θ not far from 0), will not be imitated as long as the imitation cost parameter f is sufficiently high ($f > \frac{8}{9}$). This result means that for a sufficiently high lead advance of the innovator and a sufficiently high imitation cost, imitation of the patented innovation never occurs and the technological follower is driven out of the market. This result, which occurs under a low intensity of competition in the product market (Cournot) is also true for a higher intensity of competition.⁵ Another important result is that the threshold imitation cost $\rho(d_1, \theta)$ decreases as the patent is stronger (higher θ) and as the innovation is larger (lower d_1). The patent strength θ and the imitation cost parameter f act as strategic substitutes, because both f and θ include a cost dimension for the imitator, directly via f and indirectly via θ . As θ increases, the expected damages paid by the infringer increase and correspond to a higher cost of infringement. Therefore an increase of one of these cost parameters must be compensated by a decrease of the other in order to keep the same expected profits of the imitator.

4.1.2 Small and medium innovations ($d_1 > 2c - a$)

This case is more complicated to analyze. Partial imitation is no more discarded. Indeed, three situations, namely full imitation, partial imitation and no imitation, may occur according to the values of the parameters (d_1, f, θ) . The following proposition (proof in appendix A2)

⁵It is different from the result obtained in Anton and Yao (2005) according to which imitation always occurs under the lost profit damages



summarizes the follower's optimal strategy according to the values of the parameters f , θ and d_1 when $d_1 > 2c - a$.

Proposition 2 Consider small and medium innovations ($d_1 > 2c - a$). For each value of d_1 , there exist three separating functions in the (θ, f) space defined by $\alpha(d_1, \theta) = 4A(\theta)\frac{a-d_1}{c-d_1}$, $\beta(d_1, \theta) = 2A(\theta)\left(\frac{a-d_1}{c-d_1}\right)^2 - \frac{2}{9}\left(\frac{d_1-2c+a}{c-d_1}\right)^2$ and $\gamma(\theta) = \frac{8A(\theta)}{1-9A(\theta)}$ that delineate three regions:

- If (f, θ, d_1) satisfy $f < \text{Min}(\alpha(d_1, \theta), \beta(d_1, \theta))$, then the follower fully imitates: $d_2^P(f, \theta, d_1) = d_1$

- If (f, θ, d_1) satisfy $\alpha(d_1, \theta) < f < \gamma(\theta)$, then the follower partially imitates: $d_2^P(d_1, f, \theta) = d_2^{\text{int}}(d_1, f, \theta)$

- If (f, θ, d_1) satisfy $\beta(d_1, \theta) < f < \alpha(d_1, \theta)$ or $f > \text{Max}(\alpha(d_1, \theta), \gamma(\theta))$, then the follower does not imitate: $d_2^P(f, \theta, d_1) = c$.

Functions $\alpha(d_1, \theta)$ and $\beta(d_1, \theta)$ are decreasing in the patent strength θ and in the innovation size $c - d_1$ and $\gamma(\theta)$ is decreasing in the patent strength θ . Moreover the equations in θ given by $\alpha(d_1, \theta) = \beta(d_1, \theta)$ and $\alpha(d_1, \theta) = \gamma(\theta)$ have the same solution $\theta_0(d_1) \in [0, 1[$ which means that curves $f = \alpha(d_1, \theta)$, $f = \beta(d_1, \theta)$ and $f = \gamma(\theta)$ meet at a same point $\theta_0(d_1)$ in the (θ, f) space for a given $d_1 > 2c - a$.

This proposition is illustrated in figure 2 in which we assume $2c - a < d_1 < \frac{9c-4a}{5}$. It follows from $\alpha(d_1, 0) = \frac{4}{9} \frac{a-d_1}{c-d_1}$ that $\alpha(d_1, 0) < 1 \Leftrightarrow d_1 < \frac{9c-4a}{5}$. Therefore it is worth distinguishing medium innovations ($2c - a < d_1 < \frac{9c-4a}{5}$) from small innovations ($d_1 > 2c - a$).

The imitator's choice of d_2 is affected by three variables: the cost parameter f , the patent strength θ and the innovation size $c - d_1$. Define θ_γ and $\theta_\beta(d_1)$ as the solutions of the respective equations: $\gamma(\theta) = 1$ and $\beta(d_1, \theta) = 0$. For a given innovation d_1 such that $d_1 > 2c - a$, the effect of the cost imitation parameter f on the imitation level d_2 depends on the value of the patent strength θ in $[0, 1]$ in a specific way that we now describe.

When $\theta < \theta_\gamma$, the patent is very weak and imitation occurs whatever the cost imitation parameter f for two reasons. First, the risk of infringing a very weak patent is not sufficiently dissuasive: even if an infringement lawsuit occurs, damages will be paid with a very low probability θ . Second, imitation is not expensive enough to deter imitation of a small or medium innovation ($d_1 > 2c - a$). Therefore, imitation is either partial or full according to the imitation cost f . It is only partial if f is above the threshold $\alpha(d_1, \theta)$ and it is full if f is below this threshold.

When $\theta_\gamma \leq \theta \leq \theta_0(d_1)$, the patent is stronger and imitation becomes more expensive since the payment of damages occurs with a higher probability θ . Therefore, imitation may be either absent, partial or full, according to the imitation cost parameter value f . There is no imitation at all when f is higher than $\gamma(\theta)$. Imitation is only partial when f is below $\gamma(\theta)$ and above the previous threshold $\alpha(d_1, \theta)$ and is full when f is lower than $\alpha(d_1, \theta)$.

A third situation occurs when $\theta_0(d_1) < \theta < \theta_\beta(d_1)$. In this case, infringing is much more expensive because the patent will be upheld by the court with a higher probability θ . However, keeping the old technology $d_2 = c$ is also very detrimental for the follower. Therefore, imitation is either full or absent according to whether f is below or above the lower threshold $\beta(d_1, \theta)$.

Finally, a fourth situation occurs for the highest values of θ ($\theta > \theta_\beta(d_1)$). In this case, it is no more profitable to imitate even when imitation is costless, because the patent protection is very strong. The imitation cost does not matter anymore and the presumptions that the patent will be upheld by the court and that imitation will be judged as being an infringement are so high that the patent protection entirely plays its role against imitation. Since $\theta_\beta(d_1) < 1$, it is interesting to note that less than perfect protection is sufficient to deter imitation. Therefore, it is justified to refer to the value $\theta_\beta(d_1)$ as the *safe protection level*. A patent that protects against imitation does not need to be 100% perfect and the safe protection level depends on the importance of the innovation itself. As the innovation is less important (d_1 increases), the safe protection level $\theta_\beta(d_1)$ increases. This important result suggests that smaller innovations require stronger protection, since they are likely to be imitated. This is a serious argument against the "one size fits all" protection principle.

The effect of d_1 over $d_2^P(f, \theta, d_1)$ for a given (θ, d_1) is interesting: as d_1 decreases, leading to an innovation involving a higher cost reduction, the partial imitation area increases be-

cause $\alpha(d_1, \theta)$ decreases, the full imitation area decreases because both $\alpha(d_1, \theta)$ and $\beta(d_1, \theta)$ decrease and the no-imitation area increases because $\beta(d_1, \theta)$ decreases.

4.1.3 Non patented innovation

The follower's optimal imitation strategy under regime S can be simply derived from its optimal imitation strategy under regime P by taking $f = 1$ and $\theta = 0$. The next proposition summarizes our findings when the innovator chooses to use secrecy to protect its innovation.

Proposition 3 *Under the secrecy regime, the follower's optimal imitation strategy $d_1 \rightarrow d_2^S(d_1)$ is given by:*

- If $d_1 \leq 2c - a$ then the follower does not imitate ($d_2^S(d_1) = c$) and is driven out of the market.

- If $2c - a < d_1 < \frac{9c-4a}{5}$ then the follower partially imitates ($d_2^S(d_1) = 9c - 4(a + d_1) < d_1$).

- If $\frac{9c-4a}{5} \leq d_1 \leq c$ then the follower fully imitates ($d_2^S(d_1) = d_1$).

Note that large innovations ($d_1 \leq 2c - a$) are never imitated under regime S while they are fully imitated under regime P when $f < \rho(d_1, \theta)$. The explanation of this rather unintuitive result simply derives from the previous remark that $\rho(d_1, 0) < 1$ for any $d_1 \leq 2c - a$. Then under the patent regime where some enabling knowledge is disclosed, the cost imitation parameter f may be so low (more precisely $f < \rho(d_1, \theta)$ for a patent strength θ) that it may be profitable to incur the low imitation cost $fI(d_1)$ even when damages are paid with a high probability θ .

5 Protection stage

Which protection regime will the innovator choose, once its process innovation is achieved? To answer this question, we have to compare its expected profit under the patent regime $\Pi_1^P(d_1, d_2^P(d_1, f, \theta), \theta)$ to its expected profits under the secrecy regime $\Pi_1^S(d_1, d_2^S(d_1))$, given its anticipation of the follower's imitation level under each one of the two regimes.

Let us now determine the forces that drive the innovator's protection regime choice. Consider two imitation levels $d_2, d_2' \in [d_1, c]$. The difference $\Pi_1^P(d_1, d_2, \theta) - \Pi_1^S(d_1, d_2')$ can be decomposed in the following way:

$$\Pi_1^P(d_1, d_2, \theta) - \Pi_1^S(d_1, d_2') = (\Pi_1^P(d_1, d_2, \theta) - \Pi_1^S(d_1, d_2)) + (\Pi_1^S(d_1, d_2) - \Pi_1^S(d_1, d_2'))$$

The first term of this decomposition, namely the difference $\Pi_1^P(d_1, d_2, \theta) - \Pi_1^S(d_1, d_2)$, corresponds to what we call the *damage effect*. Given an imitation level d_2 , the innovator can expect some damages if it patents its innovation, which is not the case if it chooses to keep it secret. Let us show that the damage effect is always nonnegative and nondecreasing in θ . This

is equivalent to show that function $\theta \rightarrow \Pi_1^P(d_1, d_2, \theta)$ is nondecreasing for any $d_2 \in [d_1, c]$. For any $d_2 < \min\left(c, \frac{a+d_1}{2}\right)$, one obtains:

$$\frac{\partial \Pi_1^P(d_1, d_2, \theta)}{\partial \theta} = \frac{6\theta(a-d_2)(d_2-d_1) + 6(a-2d_2+d_1) + 2(a-2d_2+d_1)^2}{3-\theta}$$

When $d_2 < \min\left(c, \frac{a+d_1}{2}\right)$, one can check that this derivative is strictly positive. In particular, this leads to :

$$\Pi_1^P(d_1, d_2, \theta) > \Pi_1^P(d_1, d_2, 0) = \Pi_1^S(d_1, d_2)$$

When $d_2 \geq \frac{a+d_1}{2}$, we have shown that $\Pi_1^P(d_1, d_2, \theta)$ does not depend on the parameter θ . In particular, $\Pi_1^P(d_1, d_2, \theta) = \Pi_1^P(d_1, d_2, 0) = \Pi_1^S(d_1, d_2)$. It follows that $\Pi_1^P(d_1, d_2, \theta) - \Pi_1^S(d_1, d_2) \geq 0$ for any $\theta \in [0, 1]$ and $d_2 \in [d_1, c]$.

Turn now to the second term. The difference $\Pi_1^S(d_1, d_2) - \Pi_1^S(d_1, d'_2)$ corresponds to what we call the *competition effect*. As the innovator and imitator products are substitute, the innovator's profits decline as it is more imitated : $\Pi_1^S(d_1, d_2)$ is an increasing function of d_2 , which implies that the sign of $\Pi_1^S(d_1, d_2) - \Pi_1^S(d_1, d'_2)$ is the same as the sign of $d_2 - d'_2$.

Hence, if the innovator anticipates that it will be less (or equally) imitated under the patent regime than under the secrecy regime then both the damage effect and the competition effect dictate the same protection regime, namely the patent regime. But if the innovator anticipates that it will be more imitated under the patent regime than under the secrecy regime, then the damage effect and the competition effect are antagonist. The first pushes the innovator to choose the patent regime while the second suggests to choose the secrecy regime. The following lemma, that summarizes and completes what precedes, is useful for the subsequent analysis:

Lemma 4 *If the innovator is less (or equally) imitated under regime P than under regime S then its optimal protection regime is the patent regime P. In particular, when the innovator anticipates that it will be fully imitated under the secrecy regime or that it will not be imitated at all under the patent regime, it always chooses to patent its innovation.*

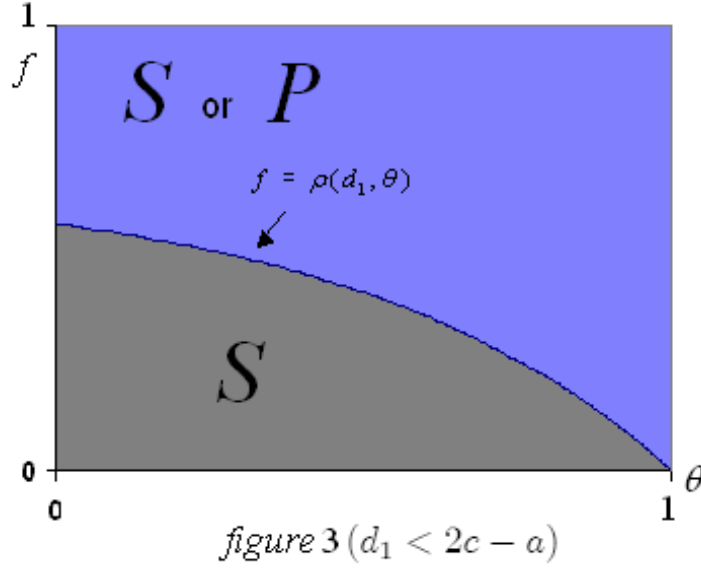
In order to determine the innovator's optimal protection regime, we distinguish the three cases that appeared in the imitation stage discussion.

Case 1: $d_1 < 2c - a$ (large innovations)

In this case, we know that the innovator is not imitated at all when it chooses to keep secrecy ($d_2^S(d_1) = c$). Its expected profit under regime S is then given by:

$$\Pi_1^S(d_1, d_2^S(d_1)) = \Pi_1^S(d_1, d_1) = \frac{(a-d_1)^2}{4}$$

We have also shown that under the patent regime, the innovator is fully imitated or not



imitated at all, according to whether $f < \rho(d_1, \theta)$ or $f > \rho(d_1, \theta)$. Hence, its expected profit under regime P is given by:

$$\Pi_1^P(d_1, d_2^P(d_1, f, \theta)) = \begin{cases} \frac{(a-d_1)^2}{(3-\theta)^2} & \text{if } f < \rho(d_1, \theta) \\ \frac{(a-d_1)^2}{4} & \text{if } f > \rho(d_1, \theta) \end{cases}$$

implying that : $\Pi_1^P(d_1, d_2^P(d_1, f, \theta)) < \Pi_1^S(d_1, d_1)$ if $f < \rho(d_1, \theta)$ and $\Pi_1^P(d_1, d_2^P(d_1, f, \theta)) = \Pi_1^S(d_1, d_2^S(d_1))$ if $f > \rho(d_1, \theta)$.

This leads to the following proposition illustrated in figure 3.

Proposition 5 *When the innovation is large enough ($d_1 < 2c - a$) the innovator prefers to keep its innovation secret if $f < \rho(d_1, \theta)$ and is indifferent between patenting and keeping secrecy if $f > \rho(d_1, \theta)$.*

Hence, keeping secrecy is always an optimal strategy of the innovator when innovation is large ($d_1 < 2c - a$). Such a choice may hinder diffusion of large innovations. This may be detrimental to society since large innovations are likely to be those which bring breakthroughs and open big opportunities for technological improvements (cumulative innovation). Proposition 5 suggests one way to make innovators patent their very inventive innovations: this may be induced either by reducing the level of compulsory disclosure which is equivalent, in our model, to increasing the value of the parameter f or by increasing the value of the expected damages.

Case 2: $2c - a < d_1 < \frac{9c-4a}{5}$ (medium innovations)

In this case, the innovator is partially imitated under regime S ($d_2^S(d_1) = 9c - 4a - 4d_1 < d_1$) and its expected profit under this regime is given by:

$$\Pi_1^S(d_1, d_2^S(d_1)) = \Pi_1^S(d_1, 9c - 4a - 4d_1) = (3c - a - 2d_1)^2$$

Three subcases must be distinguished according to the value of $d_2^P(d_1, f, \theta)$ which affects $\Pi_1^P(d_1, d_2^P(d_1, f, \theta))$.

Subcase 2.1: $f < \text{Min}(\alpha(d_1, \theta), \beta(d_1, \theta))$

We know that for such a set of parameters, the innovator is fully imitated under regime P ($d_2^P(d_1) = d_1$). Its expected profits under regime P are :

$$\Pi_1^P(d_1, d_2^P(d_1, f, \theta)) = \Pi_1^P(d_1, d_1) = \frac{(a - d_1)^2}{(3 - \theta)^2}$$

Some straightforward calculations lead to:

$$\Pi_1^P(d_1, d_1) > \Pi_1^S(d_1, 9c - 4a - 4d_1) \iff \theta > \tilde{\theta}(d_1) = \frac{9c - 4a - 5d_1}{3c - a - 2d_1} \quad (9)$$

Hence the innovator chooses to keep its innovation secret if $\theta < \tilde{\theta}(d_1)$ and to patent it if $\theta > \tilde{\theta}(d_1)$. Note that $\tilde{\theta}(d_1)$ is a decreasing function of $d_1 \in [2c - a, \frac{9c-4a}{5}]$ such that $\tilde{\theta}(\frac{9c-4a}{5}) = 0$ and $\tilde{\theta}(2c - a) = 1$.

It is interesting to compare this new threshold $\tilde{\theta}(d_1)$ to the previously defined safe protection level $\theta_\beta(d_1)$. Since $A(\theta)$ is strictly decreasing, the comparison of $\theta_\beta(d_1)$ and $\tilde{\theta}(d_1)$ can be derived from the comparison of $A(\theta_\beta(d_1))$ and $A(\tilde{\theta}(d_1))$. From $\beta(d_1, \theta_\beta(d_1)) = 0$ we derive: $A(\theta_\beta(d_1)) = \frac{(d_1 - 2c + a)^2}{9(a - d_1)^2}$ and, using the above expression of $\tilde{\theta}(d_1)$, one obtains: $A(\tilde{\theta}(d_1)) = \frac{3(d_1 - 2c + a)(3c - a - 2d_1)}{(a - d_1)^2}$. Therefore we just need to compare $\frac{d_1 - 2c + a}{9}$ and $3(3c - a - 2d_1)$. For any $d_1 \in]2c - a, \frac{9c-4a}{5}[$, we have $\frac{d_1 - 2c - a}{9} < \frac{a-c}{45}$ and $3(3c - a - 2d_1) > 3(a - c)$ and so we get $A(\theta_\beta(d_1)) < A(\tilde{\theta}(d_1))$ which is equivalent to $\tilde{\theta}(d_1) < \theta_\beta(d_1)$.

This result shows that even if a medium innovation is expected to be fully imitated under the patent regime, a patent protection is still preferred by the innovator if the patent's holder expects to recover the infringer's profit with a sufficiently high probability. What is important is that this probability $\tilde{\theta}(d_1)$ is lower than the safe protection level $\theta_\beta(d_1)$ previously defined. Therefore, patents will be filed even if their protection level is strictly lower than the safe protection level warranting perfect protection against imitation (see figure 4).

Subcase 2.2 : $\alpha(d_1, \theta) < f < \gamma(\theta)$

In this subcase, the innovator is partially imitated under regime P ($d_2^P(d_1) = d_2^{int}(d_1, f, \theta)$) and under regime S ($d_2^S(d_1) = 9c - 4(a + d_1) < d_1$). The following lemma (proof in appendix A3) compares these imitation levels under regime P and regime S.

Lemma 6 *When the innovator is partially imitated under both protection regimes, two cases arise:*

- If $f < 9A(\theta)$, then the innovator is more imitated under regime P than under regime S
- If $f > 9A(\theta)$, then the innovator is more imitated under regime S than under regime P

Finally, by combining the two previous lemmas, we reach the conclusion that when $f > 9A(\theta)$, the innovator chooses to patent its innovation since the damage effect and the competition effect go in the same direction and dictate the patent regime choice. However, if $f < 9A(\theta)$ the damage effect and the competition effect are opposite. The following lemma is crucial in order to compare $\Pi_1^P(d_1, d_2^P(d_1, f, \theta), \theta)$ to $\Pi_1^S(d_1, d_2^S(d_1))$ in this case.

Lemma 7 *Along any curve $f = KA(\theta)$ in the (θ, f) space, where K is a strictly positive parameter, the innovator's profit $\Pi_1^P(d_1, d_2^P(d_1, f, \theta), \theta)$ increases with patent strength θ as long as partial imitation occurs.*

Proof. We showed in appendix A3 that the follower's level of imitation $d_2^{int}(d_1, f, \theta)$ depends on the parameters θ and f only through $\frac{f}{A(\theta)}$. This implies that $d_2^{int}(d_1, f, \theta)$ remains constant as one moves on curve $f = KA(\theta)$. Then lemma 7 appears as a simple corollary of the result according to which the function $\theta \rightarrow \Pi_1^P(d_1, d_2, \theta)$ is increasing, for given marginal costs d_1 and d_2 . This result appeared when we introduced the damage effect. ■

Using this lemma, we derive the following result (proof in appendix A4).

Lemma 8 *For medium innovations ($2c - a < d_1 < \frac{9c-4a}{5}$), when $\alpha(d_1, \theta) < f < 9A(\theta)$, there exists a threshold $\lambda(d_1, \theta)$ decreasing in the patent strength θ such that the innovator keeps its innovation secret if $f < \lambda(\theta, d_1)$ and $\theta < \tilde{\theta}(d_1)$ and patents it if $f > \lambda(\theta, d_1)$ or $\theta > \tilde{\theta}(d_1)$. The threshold function $\lambda(d_1, \theta)$ satisfies the following two conditions : $\lambda(d_1, 0) = 1$ and $\lambda(d_1, \tilde{\theta}(d_1)) = \alpha(d_1, \tilde{\theta}(d_1))$.*

The first condition states that the innovator is indifferent between patenting and keeping secrecy when $\theta = 0$ and $f = 1$ and the second that it is indifferent between these two regimes when $\theta = \tilde{\theta}(d_1)$ and $f = \alpha(d_1, \tilde{\theta}(d_1))$ which is consistent with our previous findings (see figure 4).

Subcase 2.3: $\beta(d_1, \theta) < f < \alpha(d_1, \theta)$ or $f > \text{Max}(\alpha(d_1, \theta), \gamma(\theta))$

In this subcase, the innovator is not imitated at all under the patent regime ($d_2^P(d_1) = d_1$). We derive from lemma 4 that the innovator's optimal protection regime is regime (P).

Finally, the following proposition summarizes the case $2c - a < d_1 < \frac{9c-4a}{5}$.

Proposition 9 *For medium process innovations ($2c - a < d_1 < \frac{9c-4a}{5}$), there exist a threshold function $\tilde{\theta}(d_1)$ decreasing in the innovation size $c - d_1$ and a threshold function $\lambda(d_1, \theta)$ decreasing in the patent strength θ such that :*

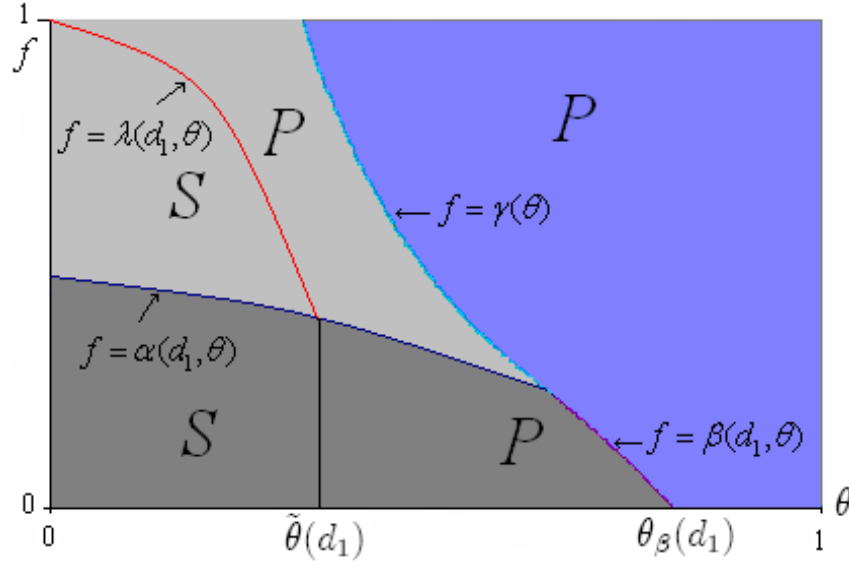


figure4 ($2c - a < d_1 < \frac{9c-4a}{5}$)

- If $\theta < \tilde{\theta}(d_1)$ and $f < \lambda(d_1, \theta)$ then the innovator chooses the secrecy regime
- If $\theta > \tilde{\theta}(d_1)$ or $f > \lambda(d_1, \theta)$ then the innovator chooses the patent regime.

This proposition, illustrated in figure 4, can be interpreted as follows. When the patent is strong enough ($\theta > \tilde{\theta}(d_1)$) the innovator always chooses to patent its innovation whatever the imitation cost parameter f . This means that the imitation cost allowed by disclosure does not matter anymore when the patent strength is above $\tilde{\theta}(d_1)$. In particular, it patents even if disclosure makes innovation costless. This result does not hold for weak patents ($\theta < \tilde{\theta}(d_1)$). In this case, the competition effect and the damage effect may go in opposite directions leading to different protection regimes according to whether the disclosure effect of patenting is high enough ($f < \lambda(d_1, \theta)$) or not. The effect of the innovation size on the secrecy region is clear: since $\tilde{\theta}(d_1)$ decreases as the innovation size decreases (d_1 increases), the corresponding area shrinks as the innovation is smaller.

Case 3: $d_1 > \frac{9c-4a}{5}$ (small innovations)

In this case, the innovator will be fully imitated if it chooses regime S ($d_2^S(d_1) = d_1$). According to lemma 4, the innovator's optimal protection regime is the patent regime (P).

Proposition 10 *Small innovations ($d_1 > \frac{9c-4a}{5}$) are always patented.*

This result can be explained in the following way : since small innovations are fully imitated under secrecy, patenting is preferred for two reasons : first, it may deter imitation

leading to higher market profits; second, even when the patent strength and the disclosure effect are such that imitation cannot be deterred, it allows the innovator to expect some damages compensating its market profits loss due to imitation, which is not the case under secrecy regime.

We develop now an overall discussion of our results by comparing them to those of Anton and Yao (2004) and by embedding them in a broader perspective.

The last proposition states that inventors of small process innovations always prefer protection induced by patent to trade secrecy. This result is a reminder of the "little patents" expression coined by Anton and Yao (2004). However, the argument behind this common result is different in our model. Under the secrecy regime, small process innovations are fully imitated for two reasons that reinforce one another. First, imitating a small innovation is not very costly and second, there is no threat of an infringement lawsuit when the innovation is kept secret. Under patent protection, such a threat exists and it may overrun the benefits that the infringer expects from imitating the leader. Note that in our model, small innovations may be imitated under the patent regime while this does not occur in Anton and Yao.

For large process innovations, our results are similar to the "big secrets" characterization obtained in Anton and Yao. Large process innovations are never imitated when they are kept secret, while the enabling knowledge disclosed by a patent may reduce the imitation cost in a way that renders their imitation attractive. This classical tradeoff in the economics of patents explains why "big secrets" are preferred to "big patents". Note however that our model does not totally discard the possibility of patenting some large process innovations. This may occur when their imitation under the patent regime is too costly. In this case we have shown that the innovator is indifferent between secret and patent, because in both cases, the innovation is not imitated.

Finally, it is for medium process innovations that our results significantly differ from those of Anton and Yao. We have shown that keeping an medium process innovation secret does not avoid imitation. It is precisely for medium process innovations that partial innovation occurs under the secrecy regime. However, under the patent regime, imitation may be either absent, partial and total. We have also shown that the innovator may patent or keep secrecy while in Anton and Yao, medium process innovations are always patented and partially disclosed.

6 Licensing agreement as an alternative to litigation

In this section, we allow licensing agreements between the innovator and the follower. Since our purpose is not to study all the possible agreements that may emerge, we restrict our attention to the simple case of process innovations leading to a small cost reduction, i.e. $d_1 > \frac{9c-4a}{5}$. When introduced in the market, these innovations are fully imitated under regime S which leads the innovator to patent them. We analyze licensing agreements between

the innovator and an imitator that avoid litigation to be completed until the court's decision.

We study two-part tariff licences (r, F) where r is a royalty rate and F a fixed fee. Let us first examine the equilibrium outcomes when the innovator and the follower agree on a licence (r, F) . Gross profits can be written as

$$\Pi_1^L(x_1, x_2, d_1, r, F) = (a - x_1 - x_2 - d_1)x_1 + rx_2 + F$$

$$\Pi_2^L(x_1, x_2, d_1, r, F) = (a - x_1 - x_2 - (d_1 + r))x_2 - F$$

An equilibrium of competition stage under license regime (L) is given by

$$x_1^L = \frac{a - d_1 + r}{3}$$

$$x_2^L = \frac{a - d_1 - 2r}{3}$$

and leads to the following equilibrium price:

$$p^L = \frac{a + 2d_1 + r}{3}$$

We assume hereafter that $0 \leq r \leq \frac{a-d_1}{2}$. This assumption does not entail any loss of generality since any royalty rate such that $r > \frac{a-d_1}{2}$ leads to the same boundary solution as $r = \frac{a-d_1}{2}$:

$$x_1^L = \frac{a - d_1}{2}$$

$$x_2^L = 0$$

Equilibrium gross profits are given by:

$$\Pi_1^L(d_1, r, F) = \left(\frac{a - d_1 + r}{3}\right)^2 + r\frac{a - d_1 - 2r}{3} + F \quad (10)$$

$$\Pi_2^L(d_1, r, F) = \left(\frac{a - d_1 - 2r}{3}\right)^2 - F \quad (11)$$

6.1 Benchmark case

First, we address the following question: does there exist a royalty rate r such that the profits under licence regime replicate the profits under patent regime without licensing:

$$\Pi_1^L(d_1, r, F) = \Pi_1^P(d_1, \theta)$$

and

$$\Pi_2^L(d_1, r, F) = \Pi_2^P(d_1, \theta)$$

Since we do not set any restriction on the fixed fee F (in particular, we allow F to be negative), it is clear that such a question amounts to the existence of a royalty rate r such that the industry profits $\Pi^L(d_1, r) = \Pi_1^L(d_1, r, F) + \Pi_2^L(d_1, r, F)$ under the licence regime replicate the industry profits under the patent regime without licensing, i.e:

$$\Pi^L(d_1, r) = \Pi_1^P(d_1, \theta) + \Pi_2^P(d_1, \theta) \quad (12)$$

This can be rewritten as:

$$\frac{1}{9}[-r^2 + (a - d_1)r + 2(a - d_1)^2] = \frac{2 - \theta}{(3 - \theta)^2}(a - d_1)^2 \quad (13)$$

Solving this equation, we find a unique solution :

$$\tilde{r}(\theta) = \frac{\theta}{3 - \theta}(a - d_1) \quad (14)$$

The equilibrium price is then given by:

$$\tilde{p}^L(\theta) = \frac{a + d_1(2 - \theta)}{3 - \theta}$$

The royalty rate $\tilde{r}(\theta)$ and equilibrium price $\tilde{p}^L(\theta)$ are increasing in θ . We can now derive the fixed fee $\tilde{F}(\theta)$ from equation $\Pi_2^L(d_1, \tilde{r}(\theta), \tilde{F}(\theta)) = \Pi_2^P(d_1, \theta)$:

$$\tilde{F}(\theta) = \frac{\theta(\theta - 1)}{(3 - \theta)^2}(a - d_1)^2$$

Thus, we obtain the following properties of the function \tilde{F} :

- $\tilde{F}(0) = \tilde{F}(1) = 0$.
- $\tilde{F}(\theta) < 0$ for any $\theta \in]0, 1[$.
- $\tilde{F}(\theta)$ is strictly decreasing over $[0, \frac{3}{5}]$ and strictly increasing over $[\frac{3}{5}, 1]$.

Note that licence $(\tilde{r}(\theta), \tilde{F}(\theta))$ not only replicates the expected profits of the innovator and the follower under regime P (without licensing) but also induces the same equilibrium outputs and then the same equilibrium price. It is indeed easy to check that :

$$\tilde{x}_1^L(\theta) = \frac{a - d_1 + \tilde{r}(\theta)}{3} = \frac{(1 - \theta)(a - d_1)}{3 - \theta} = x_1^P(\theta)$$

and

$$\tilde{x}_2^L(\theta) = \frac{a - d_1 - 2\tilde{r}(\theta)}{3} = \frac{a - d_1}{3 - \theta} = x_2^P(\theta)$$

which obviously lead to

$$\tilde{p}^L(\theta) = p^P(\theta)$$

where $p^P(\theta)$ is the equilibrium price under patent regime without licensing.

Thus, the fact that competition stage outcomes are identical under both regimes (patent under the shadow of infringement and license defined by $(\tilde{r}(\theta), \tilde{F}(\theta))$), allows us to consider the licence regime defined by $(\tilde{r}(\theta), \tilde{F}(\theta))$ as a *benchmark* to which we will compare the outcomes of the licensing agreements that are likely to emerge.

6.2 Maximizing the industry profits

One of the agreements which are more likely to emerge is a two-part tariff license such that the generated industry profits is maximum among all the joint profits realized by two-part tariff licenses. The part of the industry profits allocated to each firm is then determined by the fixed fee F .

Consider such a licensing agreement, denoted by $(\hat{r}(\theta), \hat{F}(\theta))$. Since $\hat{r}(\theta)$ is defined by $\hat{r}(\theta) = \underset{0 \leq r \leq \frac{a-d_1}{2}}{\text{Arg max}} \Pi^L(d_1, r)$, the royalty $\hat{r}(\theta)$ does not depend on the patent strength θ . More-

over, it is easy to see that $\Pi^L(d_1, r)$ is an increasing function of r over interval $\left[0, \frac{a-d_1}{2}\right]$. Then,

$$\hat{r}(\theta) = \hat{r} = \frac{a - d_1}{2} = \tilde{r}(1), \quad \forall \theta \in [0, 1]$$

leading to the price:

$$\hat{p}(\theta) = \hat{p} = \frac{a + d_1}{2} = \tilde{p}^L(1) \geq \tilde{p}^L(\theta), \quad \forall \theta \in [0, 1]$$

This result means that the patent strength is no more reflected by the price paid by consumers. In particular, low quality patents which generate lower prices when litigated, generate the same maximal price as would do high quality patents. This is the main concern raised by licensing agreements that harm consumers.

Note that $x_2^L = \frac{a-d_1-2r}{3} = 0$ when $r = \hat{r}$. Hence, when the follower accepts licence $(\hat{r}, \hat{F}(\theta))$ it implicitly accepts to stay out of the market and the industry profits are then captured by the patentee. Nevertheless, the innovator transfers a part of these monopoly profits to the licensee through the negative fixed fee $\hat{F}(\theta)$. In other words, license $(\hat{r}, \hat{F}(\theta))$ is equivalent to an agreement where the innovator pays its competitor to stay out of the market. The profits of the patentee and the licensee when they agree on license $(\hat{r}, \hat{F}(\theta))$ are then

given by:

$$\begin{aligned}\Pi_1^L(d_1, \theta) &= \frac{(a - d_1)^2}{4} + \hat{F}(\theta) \\ \Pi_2^L(d_1, \theta) &= -\hat{F}(\theta)\end{aligned}$$

This type of agreement is accepted by both the innovator and the follower as long as fixed fee $\hat{F}(\theta)$ is in interval $\left[\Pi_1^P(d_1, \theta) - \frac{(a-d_1)^2}{4}, -\Pi_2^P(d_1, \theta)\right]$, that is equivalent to:⁶

$$\hat{F}(\theta) \in \left[-A(\theta)(5 - \theta)(a - d_1)^2, -A(\theta)(a - d_1)^2\right] \quad (15)$$

The possibility that such two-part tariffs involving negative fees emerge in licensing agreements is a big concern for competition authorities. We know that in the pharmaceutical industry, agreements of this kind are allowed under the Hatch-Waxman Act and have been effectively used by some patents holders in their negotiations with generic challengers. They obviously harm consumers and this is why patent settlements, which take the form of licensing agreements, must be under the scrutiny of competition authorities (Shapiro, 2003).

7 Conclusion

Departing from the usual convention that patents are perfect forms of protection opens a lot of research avenues. One of the most important is to know under what conditions a patent is preferred to trade secrecy. Our model provides a complete theoretical answer to this question for a process innovation. For each class of cost reduction (small, medium and large) we have obtained specific results. First, we have determined the imitation level in each regime. Second, in the space of the two key parameters (patent strength and relative imitation cost) we have derived the partition that delineates areas where one protection regime dominates the other. How can one use these results for a policy purpose? This is an interesting and complex question for which we propose preliminary insights. Consider the relative imitation cost. In a world where patent design is independent of the invention, particularly concerning the same compulsory disclosure for all patents, it seems very hard to determine a priori what would be the value of the imitation cost parameter. One can simply reach a rather vague idea of the secrecy effectiveness of the invention, that leads to an idiosyncratic characterization covering a large spectrum of possibilities, running from the "naked idea" case to the "perfectly hidden idea" case. This type of assessment would depend on some priors on whether the the invention could be more or less easily discovered by reverse-engineering. But in a world where a patent

⁶The licence $(\hat{r}, F_0(\theta))$ where $F_0(\theta) = -A(\theta)(5 - \theta)(a - d_1)^2$ is the optimal licence, from the innovator's perspective, among all the licences that maximize industry profits. It is likely to emerge if the innovator has a "take it or leave it" bargaining power. Indeed, it is clear that with such a bargaining power the innovator will pay its competitor the minimum amount that makes it accept to stay out the market.

is not designed around the "one size fits all" principle, some flexibility could be introduced by allowing each innovator to choose a patent inside a menu of characteristics. For instance an innovator may have to choose between a patent with strong property rights and high disclosure requirements and a patent with weak property rights and low disclosure requirements. If an incentive mechanism built around this principle could be achieved, it would be an appropriate answer to the rather disappointing result according to which "little patents and big secrets" are the preferred forms of protection. Small innovations could be easily imitated because their rights are weak. Large innovations could be patented because their rights are strong. In both cases, diffusion of innovation would be enhanced. The construction of such an incentive mechanism is the next step in our agenda.

Our model analyzes also the licensing agreements between a patent holder and a competitor. Such agreements avoid the litigation to go until completion. One of the possible consequences of a patent settlement as an alternative to a trial raises some concerns. The royalty rate paid by the licensee does not depend on the patent strength as a natural benchmark would command. Licensing very bad quality patents may occur with as high royalty rate as if the patent was undoubtful. Moreover, the patentee pays a fixed fee to the licensee to compensate its loss in the market. While the two parties maximize their joint profits, it is clear that such a settlement harms consumers and creates a big concern for society. Shapiro (2003) and Farrell and Shapiro (2005) reach the same conclusion by using quite different models.

Finally, while some economists (Ayres and Klemperer, 1999) find that probabilistic rights open welfare improving opportunities (entry occurs under the shadow of punishment) it is also important to stress some of their negative consequences. Adopting trade secrecy for large inventions may reduce the diffusion of innovation. Moreover, patent settlements of the sort examined in this paper are detrimental to society. This is one reason why patent quality is probably one of the most challenging issues to which we are now confronted.

8 References

Anton, J.J. and D. A. Yao, 2004, "Little patents and big secrets: managing intellectual property", *Rand Journal of Economics* 35, 1-22

Anton, J.J. and D. A. Yao, 2005 a, "Finding "Lost Profits": an equilibrium analysis of patent infringement damages", W.P. Duke University

Anton, J.J., H. Greene and D. A. Yao, 2005 b, "Policy implications of weak patent rights", mimeo

Cohen, W.M, R.R. Nelson and J.P. Walsh, 2000, "Protecting their intellectual assets: appropriability conditions and why US manufacturing firms patent or not", NBER W.P. 7552

- Crampes, C, 1986, "Les inconvénients d'un dépôt de brevet pour une entreprise innovatrice", *L'Actualité Economique, Revue d'Analyse Economique*, 62, 4, 521-534
- Encaoua, D. and A. Hollander, 2004, "Competition policy and innovation", *Oxford Review of Economic Policy* 18, 63-79
- Encaoua, D., D. Guellec and C. Martinez, 2005, "Patent systems for encouraging innovation: Lessons from economic analysis", W.P., University Paris I
- Farrell, J. and C. Shapiro, 2005, "How strong are weak patents?", W. P. University of California at Berkeley
- Gallini, N.T., 1992, "Patent policy and costly innovation", *Rand Journal of Economics* 23, 52-63
- Horstman, I., G.M. MacDonald and A. Slivinski, 1985, "Patents as information transfer mechanisms: To patent or (Maybe) not to patent", *Journal of Political Economy* 95-5, 837-858
- Lemley, M.A. and C. Shapiro, 2005, "Probabilistic patents", *Journal of Economic Perspectives* 19-2, 75-98
- Levin, R.C., A.K. Klevorick, R.R. Nelson and S.G. Winter, 1987, "Appropriating the returns from industrial R&D", *Brooking Papers on Economic Activity*, 783-820
- Mansfield, E., 1986, "Patents and innovation: an empirical survey", *Management Science* 32, 173-181
- Merges, R., 1999, "As many as six impossible patents before breakfast: property rights for business concepts and patent system reform", *Berkeley High Technology Law Journal* 14, 577-615
- Pakes, A. and Z. Griliches, "Patents and R&D at the firm level: a first look", *Economics Letters* 5-4, 377-381
- Scherer, F.M., 1965, "Market structure, opportunity and the output of patented inventions", *American Economic Review* 55, 1097-1125
- Scherer, F.M., 1967, "Market structure and the employment of scientists and engineers", *American Economic Review* 57, 524-531
- Scherer, F.M., 1983, "The propensity to patent", *International Journal of Industrial Organization* 1, 107-128
- Shapiro, C., 2003, "Antitrust analysis of patent settlements between rivals", *Rand Journal of Economics* 34-2, 391-411
- Scotchmer, S., 2004, *Innovation and Incentives*, The MIT Press, Cambridge, Ma.
- Scotchmer, S. and J. Green, 1990, "Novelty and disclosure in patent law", *Rand Journal of Economics* 21, 131-146

9 Appendix

A1. Proof of Proposition 1

Since $\Pi_2^P(d_1, d_2, f, \theta) = 0$ for any $d_2 \in \left[\frac{a+d_1}{2}, c\right]$, it follows from the fact that imitation is costly that the follower's best imitation level over this interval is $d_2 = c$:

$$\underset{d_2 \in \left[\frac{a+d_1}{2}, c\right]}{\text{Arg max}} G_2^P(d_1, d_2, f, \theta) = c \quad (16)$$

which implies that the follower's optimal imitation level is necessarily equal to either c or

$$\underset{d_2 \in \left[d_1, \frac{a+d_1}{2}\right]}{\text{Arg max}} G_2^P(d_1, d_2, f, \theta).$$

In order to determine the maximum value of $G_2^P(d_1, d_2, f, \theta)$ over the interval $\left[d_1, \frac{a+d_1}{2}\right]$, we must distinguish two cases:

Case 1 : $f < 8A(\theta)$

Function $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is convex over $\left[d_1, \frac{a+d_1}{2}\right]$ since $G_2^P(d_1, d_2, f, \theta) = H(d_1, d_2, f, \theta)$ over this interval. Moreover, it is straightforward to show that $d_2^{int}(d_1, f, \theta) < \frac{a+d_1}{2}$. Then, there are two possibilities according to whether $d_2^{int}(d_1, f, \theta) > d_1$ or $d_2^{int}(d_1, f, \theta) \leq d_1$.

- If $d_2^{int}(d_1, f, \theta) > d_1$ then $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is decreasing over the interval $\left[d_1, d_2^{int}(d_1, f, \theta)\right]$ and is increasing over the interval $\left[d_2^{int}(d_1, f, \theta), \frac{a+d_1}{2}\right]$, which entails that $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ reaches its maximum value over the interval $\left[d_1, \frac{a+d_1}{2}\right]$ at $d_2 = d_1$ or $d_2 = \frac{a+d_1}{2}$.

- If $d_2^{int}(d_1, f, \theta) < d_1$ then $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is increasing over the interval $\left[d_1, \frac{a+d_1}{2}\right]$ which implies that it reaches its maximum value at $d_2 = c$.

The crucial point is that in both cases, $d_2^P(d_1, f, \theta) \in \left\{d_1, \frac{a+d_1}{2}, c\right\}$. Since we know that the follower prefers not to imitate rather than imitate at a level $d_2 = \frac{a+d_1}{2}$, it is sufficient to compare $G_2^P(d_1, d_1, f, \theta)$ to $G_2^P(d_1, c, f, \theta) = 0$ in order to determine $d_2^P(d_1, f, \theta)$. Hence, two subcases arise:

- If $f < 2A(\theta) \left(\frac{a-d_1}{c-d_1}\right)^2$ then $G_2^P(d_1, d_1, f, \theta) > 0$ which results in $d_2^P(d_1, f, \theta) = d_1$ (full imitation)

- If $2A(\theta) \left(\frac{a-d_1}{c-d_1}\right)^2 < f < 8A(\theta)$ then $G_2^P(d_1, d_1, f, \theta) < 0$ which results in $d_2^P(d_1, f, \theta) = c$ (no imitation).

Case 2 : $f > 8A(\theta)$

In this case, $d_2^{int}(d_1, f, \theta) > c > \frac{a+d_1}{2}$ and $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is concave over the interval $\left[d_1, \frac{a+d_1}{2}\right]$. Then, the function $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is increasing over the interval $\left[d_1, \frac{a+d_1}{2}\right]$, which results in $\underset{d_2 \in \left[d_1, \frac{a+d_1}{2}\right]}{\text{Arg max}} G_2^P(d_1, d_2, f, \theta) = \frac{a+d_1}{2}$. According to (3), this leads

to $d_2^P(d_1, f, \theta) = c$ (no imitation). QED

A2. Proof of Proposition 2

The imitator must compare the maximal net profit it can get when it imitates, i.e. $\sup_{d_2 \in [d_1, c]} G_2^P(d_1, d_2, f, \theta) = \sup_{d_2 \in [d_1, c]} H(d_1, d_2, f, \theta)$, to the net profit it derives from keeping its old technology, ie. $G_2^P(d_1, c, f, \theta) = \frac{(d_1 - 2c - a)^2}{9}$. Two cases must be distinguished:

Case 1 : $f < 8A(\theta)$

In this case, $d_2^{int}(d_1, f, \theta) > c$ and $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is convex over the interval $[d_1, c]$, which entails that $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is decreasing over the interval $[d_1, c]$, and results in $\sup_{d_2 \in [d_1, c]} G_2^P(d_1, d_2, f, \theta) = G_2^P(d_1, d_1, f, \theta)$ which has to be compared to $G_2^P(d_1, c, f, \theta)$. This leads us to distinguish two subcases.

Set $\beta(d_1, \theta) = 2A(\theta) \left(\frac{a-d_1}{c-d_1} \right)^2 - \frac{2}{9} \left(\frac{d_1-2c+a}{c-d_1} \right)^2$.

- If $f < \beta(d_1, \theta)$ then $G_2^P(d_1, d_1, f, \theta) > G_2^P(d_1, c, f, \theta)$ which results in $d_2^P(d_1, f, \theta) = d_1$ (full imitation).

- If $f > \beta(d_1, \theta)$ then $G_2^P(d_1, d_1, f, \theta) < G_2^P(d_1, c, f, \theta)$ which results in $d_2^P(d_1, f, \theta) = c$ (no imitation).

Case 2 : $f > 8A(\theta)$

In this case, $d_2^{int}(d_1, f, \theta) < c$ and $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is concave over the interval $[d_1, c]$. Two subcases must be distinguished :

- If $d_2^{int}(d_1, f, \theta) < d_1$ then $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ is decreasing over the interval $[d_1, c]$, which implies that $\sup_{d_2 \in [d_1, c]} G_2^P(d_1, d_2, f, \theta) = G_2^P(d_1, d_1, f, \theta)$.

- If $d_2^{int}(d_1, f, \theta) > d_1$ then $d_2 \rightarrow G_2^P(d_1, d_2, f, \theta)$ reaches its maximum over $[d_1, c]$ at $d_2 = d_2^{int}(d_1, f, \theta)$: $\sup_{d_2 \in [d_1, c]} G_2^P(d_1, d_2, f, \theta) = G_2^P(d_1, d_2^{int}(d_1, f, \theta), f, \theta)$.

Consider the condition $d_2^{int}(d_1, f, \theta) < d_1$. It is straightforward to show that this inequality can be rewritten as :

$$f < \alpha(d_1, \theta) = \frac{4(a-d_1)}{c-d_1} A(\theta) \quad (17)$$

Then, the two previous subcases can be written as:

- If $8A(\theta) < f < \frac{4(a-d_1)}{c-d_1} A(\theta)$ then $\sup_{d_2 \in [d_1, c]} G_2^P(d_1, d_2, f, \theta) = G_2^P(d_1, d_1, f, \theta)$ which has to be compared to $G_2^P(d_1, c, f, \theta)$ (this has been previously done).

- If $f > \frac{4(a-d_1)}{c-d_1} A(\theta)$ then $\sup_{d_2 \in [d_1, c]} G_2^P(d_1, d_2, f, \theta) = G_2^P(d_1, d_2^{int}(d_1, f, \theta), f, \theta)$ which has to be compared to $G_2^P(d_1, c, f, \theta)$. Comparing these two terms is equivalent to compare f to the threshold $\gamma(\theta) = \frac{8}{(3-\theta)^2 - 9}$. More precisely :

- If $\frac{4(a-d_1)}{c-d_1} A(\theta) < f < \gamma(\theta)$ then $G_2^P(d_1, d_2^{int}, f, \theta) > G_2^P(d_1, c, f, \theta)$ which leads to $d_2^P(d_1, f, \theta) = d_2^{int}(d_1, f, \theta)$ (partial imitation).

- If $f > \gamma(\theta)$ then $G_2^P(d_1, d_2^{int}, f, \theta) < G_2^P(d_1, c, f, \theta)$ which leads to $d_2^P(d_1, f, \theta) = c$ (no imitation).

Let us now show that the equations $\alpha(d_1, \theta) = \beta(d_1, \theta)$ and $\alpha(d_1, \theta) = \gamma(\theta)$ have the same solution $\theta_0(d_1)$ over the interval $[0, 1[$, which means that the curves $f = \alpha(d_1, \theta)$, $f = \beta(d_1, \theta)$ and $f = \gamma(\theta)$ meet at the same point.

Some straightforward computations show that the equation $\alpha(d_1, \theta) = \beta(d_1, \theta)$ is equivalent to the equation:

$$A(\theta) = \frac{1}{9} \frac{d_1 - 2c + a}{a - d_1}$$

Therefore, equations $\alpha(d_1, \theta) = \beta(d_1, \theta)$ and $\alpha(d_1, \theta) = \gamma(\theta)$ have the same solution in θ over the interval $[0, 1[$ if (and only if) $\frac{1}{9} \frac{d_1 - 2c + a}{a - d_1}$ is a solution of equation $\frac{4(a-d_1)}{c-d_1} X = \frac{8}{X-9}$. It is easy to check that this is satisfied. QED

A3. Proof of lemma 6

It is easy to see that $d_2^{int}(d_1, f, \theta)$ depends on the parameters (f, θ) only through the parameter $f \frac{(3-\theta)^2}{1-\theta}$. With a slight modification of notations, we can write $d_2^{int}(d_1, f, \theta) = d_2^{int}(d_1, \frac{f}{A(\theta)})$. It is also clear that the imitation level d_2^{int} is increasing in $\frac{f}{A(\theta)}$. Hence, lemma 6 simply derives from $d_2^S(d_1) = d_2^{int}(d_1, \frac{f}{A(\theta)} = 9)$.

A4. Proof of lemma 8.

Consider first the case $\theta < \tilde{\theta}(d_1)$. Let us show that equation $\Pi_1^P(d_1, d_2^{int}(d_1, f, \theta), \theta) = \Pi_1^S(d_1, 9c - 4a - 4d_1)$ which expresses that the innovator is indifferent between patenting and keeping secrecy (in this subcase) has a unique solution in f over the interval $[\alpha(d_1, \theta), 9A(\theta)]$. Since the function $f \rightarrow d_2^{int}(d_1, f, \theta)$ is strictly increasing over the interval $[\alpha(d_1, \theta), 9A(\theta)]$, this is equivalent to state that equation $\Pi_1^P(d_1, d_2, \theta) = \Pi_1^S(d_1, 9c - 4a - 4d_1)$ has a unique solution in d_2 over the interval $[d_2^{int}(d_1, \alpha(d_1, \theta), \theta), d_2^{int}(d_1, 9A(\theta), \theta)]$. The last interval can simply be written as $[d_1, 9c - 4a - 4d_1]$. Note that function $F_\theta : d_2 \rightarrow \Pi_1^P(d_1, d_2, \theta) - \Pi_1^S(d_1, 9c - 4a - 4d_1)$ is a convex parabolic function then it is either i/ increasing over $[d_1, 9c - 4a - 4d_1]$ or ii/ U-shaped over $[d_1, 9c - 4a - 4d_1]$. Since $\theta < \tilde{\theta}(d_1)$, we have $F_\theta(d_1) < 0$ (see subcase 2.1). We also know from lemma 4 that $F(9c - 4a - 4d_1) \geq 0$. It follows that in both cases i/ and ii/ equation $F_\theta(d_2) = 0$ has a unique solution over $[d_1, 9c - 4a - 4d_1]$. Hence, equation $\Pi_1^P(d_1, d_2^{int}(d_1, f, \theta), \theta) = \Pi_1^S(d_1, 9c - 4a - 4d_1)$ has a unique solution in f that we denote by $\lambda(d_1, \theta)$. Note that $d_2^{int}(d_1, 1, 0) = d_2^S(d_1) = 9c - 4a - 4d_1$ which leads to $\lambda(d_1, 1) = 0$. Note also that :

$$\Pi_1^P(d_1, d_2^{int}(d_1, f, \theta), \theta) > \Pi_1^S(d_1, 9c - 4a - 4d_1) \text{ if and only if } f > \lambda(d_1, \theta) \quad (18)$$

Furthermore, we know that $d_2^{int}(d_1, \alpha(d_1, \tilde{\theta}(d_1)), \tilde{\theta}(d_1)) = d_1$ and $\Pi_1^P(d_1, d_1, \theta) = \Pi_1^S(d_1, 9c -$

$4a-4d_1$) (see subcase 2.1) so $\Pi_1^P(d_1, d_2^{int}(d_1, \alpha(d_1, \tilde{\theta}(d_1)), \tilde{\theta}(d_1)), \tilde{\theta}(d_1)) = \Pi_1^S(d_1, 9c-4a-4d_1)$ which leads to $\lambda(d_1, \tilde{\theta}(d_1)) = \alpha(d_1, \tilde{\theta}(d_1))$.

Consider now the case $\theta > \tilde{\theta}(d_1)$. Let $\theta_0 > \tilde{\theta}(d_1)$ and $f_0 \in]\alpha(d_1, \theta), 9A(\theta)[$. Point (θ_0, f_0) belongs to the curve $f = \frac{f_0}{A(\theta_0)}A(\theta)$. It is easy to see (graphically or analytically) that curve $f = \frac{f_0}{A(\theta_0)}A(\theta)$ necessarily meets either the curve defined by $\theta = \tilde{\theta}(d_1)$ and $f \leq \alpha(d_1, \theta)$ or the curve defined by $f = \lambda(d_1, \theta)$ and $\theta \leq \tilde{\theta}(d_1)$ at a point (θ_1, f_1) such that $\theta_1 < \theta_0$. Since in any point of the last two curves the innovator's profit under regime P is equal to its profit under regime S, and $\theta_1 < \theta_0$, we derive from lemma 7 that the innovator's profit under regime P is greater than its profit under regime S when $(\theta, f) = (\theta_0, f_0)$. Hence, the innovator chooses to patent its innovation.

Let us now show that $\lambda(d_1, \theta)$ is strictly decreasing in θ . Consider θ_1 and θ_2 such that $\theta_1 < \theta_2 \leq \tilde{\theta}(d_1)$. Points (θ_1, d_1) and $(\theta_2, \frac{\lambda(d_1, \theta_1)}{A(\theta_1)}A(\theta_2))$ belong to the curve defined by $f = \frac{\lambda(d_1, \theta_1)}{A(\theta_1)}A(\theta)$. We derive from lemma 7 that $\Pi_1^P(d_1, d_2^{int}(d_1, \frac{\lambda(d_1, \theta_1)}{A(\theta_1)}A(\theta_2), \theta_2), \theta_2) < \Pi_1^P(d_1, d_2^{int}(d_1, \lambda(d_1, \theta_1), \theta_1), \theta_1) = \Pi_1^S(d_1, 9c-4a-4d_1)$ which leads, according to (19), to $\frac{\lambda(d_1, \theta_1)}{A(\theta_1)}A(\theta_2) > \lambda(d_1, \theta_2)$. Furthermore we know that $A(\theta)$ is positive and decreasing, so $\frac{A(\theta_2)}{A(\theta_1)} < 1$. This allows us to state that $\lambda(d_1, \theta_1) > \lambda(d_1, \theta_2)$. QED.