Società italiana di economia pubblica

XV CONFERENZA

DIRITTI, REGOLE, MERCATO
Economia pubblica ed analisi economica del diritto

Pavia, Università, 3 - 4 ottobre 2003

THE CONTRACT THEORY OF PATENTS

VINCENZO DENICOLÒ AND LUIGI ALBERTO FRANZONI

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

DIRITTI, REGOLE, MERCATO
Economia pubblica ed analisi economica del diritto

XV Conferenza SIEP - Pavia, Università, 3 - 4 ottobre 2003

pubblicazione internet realizzata con contributo della COMPAGNIA di San Paolo

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

The contract theory of patents

Vincenzo Denicolò and Luigi Alberto Franzoni[1]
Department of Economics, University of Bologna

Bologna, August 2003

[1]We are grateful to Richard Posner, Frank Stephen, and seminar participants in Athens and Hamburg for helpful comments.

Two distinct theories of patents, the "reward theory" and the "contract theory," are customarily adopted by the courts to justify the patent system. The reward theory maintains that the function of the patent system is to remunerate successful innovators so as to encourage R&D effort. In contrast, the contract theory holds that the function of the patent system is to promote the diffusion of innovative knowledge. Assuming that in the absence of patent protection innovators would rely on trade secrecy, it views patents as a *contract* between innovators and society whereby a property right is granted in exchange for *disclosure*.

This paper develops an economic analysis of the contract theory of patents. To disentangle the disclosure from the reward motive for granting patents, we assume that the innovation process is entirely serendipitous, so that R&D effort is not a concern. Our main finding is that the disclosure motive alone suffices to justify the grant of patents. The optimal patent duration should strike a balance between the incentive to induce disclosure and the aim of limiting the monopoly distortion induced by patents.

1 Introduction

Two distinct theories of patents, the "reward theory" and the "contract theory," are customarily adopted by the courts to justify the patent system. The *reward theory* maintains that the function of the patent system is to remunerate successful innovators so as to encourage R&D effort. In the wording of the US Constitution, intellectual property rights are granted in order to "promote the Progress of Science and useful Arts." This theory– by far the most prominent approach to the economic analysis of patents since the classic work of Nordhaus (1969) – assumes that unpatented innovations are easily imitated, and thus focuses on the "non-exclusive" nature of technological knowledge. In this perspective, in the absence of a patent system, there would be too little investment in R&D.

In contrast, the *contract theory* emphasizes the "non-rival" nature of innovation: once it is created, it can be shared at no cost. On the assumption that absent patent protection firms can practice their innovations secretly, it views patents as a "contract" between innovators and society, whereby a temporary property right is granted in exchange for disclosure. Thus, the contract theory holds that the function of the patent system is to promote the diffusion of innovative knowledge. This theory has a long tradition and is popular with courts. In the landmark case *Universal Oil Products v. Globe Oil & Refining* (1944), for instance, the US Supreme Court couched the view that: "As a reward for inventions and to encourage their disclosure, the United States offers a seventeen-year monopoly to an inventor who refrains from keeping his invention a trade secret. But the quid pro quo is disclosure of a process or device in sufficient detail to enable one skilled in the art to practice the invention once the period of the monopoly has expired; and the same precision of disclosure is likewise essential to warn the industry concerned of the precise scope of the monopoly asserted." Clearly, the reward and the contract theory are complementary, rather than alternative. However, each of them is logically independent of the other.[1]

In this paper, we concentrate on patents as alternatives to trade secrets, as legal devices able to induce firms to disclose to the public their innovation, rather than tools necessary to foster industrial research. In this way, we try and provide a rigorous economic analysis of the contract theory of patents.

We assume that technical complexity combined with trade secret law make secrecy an effective and valuable tool to protect innovations.[2] This assumption is clearly not

[1]See Eisenberg (1989) for an accurate overview of current legal theories and Machlup (1968) for an historical perspective.

[2]On the economics of trade secrets, see Friedmand, Landes and Posner (1991).

appropriate for certain industries, while it is for many others. A large amount of empirical evidence shows that secrecy and lead-time are consistently regarded as better protection mechanisms than patents by most firms, with the notable exception of those active in the pharmaceutical, chemical and mechanical industries (see, for instance, Cohen, Nelson and Walsh 2000 and Arundel 2001). Furthermore, secrecy is shown to have increased in importance over the last decade. This might be partly explained by the strengthening accorded to trade secret protection in national legislations following the TRIPs (Trade Related Aspects of Intellectual Property Rights) chapter of the Uruguay Round Agreement of 1994. In fact, Art. 39 of the TRIPs defines a (minimal) international standard for the legal protection of "undisclosed information" against unfair competition, which has since become compulsory for all countries belonging to the WTO.[3]

In order to disentangle the disclosure from the reward motive for granting patents, we assume that the innovation is the fruit of "serendipity".[4] We thus focus on the costs and benefits of patents v. trade secrecy. We posit that when the innovation is

[3] The definition of the protected "undisclosed information" conforms to the definition found in the US Uniform Trade Secret Act: the information must be secret, have commercial value because it is secret and have been subject to reasonable steps by the person lawfully in control of the information to maintain its secrecy. Moreover, only the disclosure to, acquisition by, or use by others "in a manner contrary to honest commercial practices" is unlawful. Such unlawful practices include breach of contract, breach of confidence and inducement to breach, as well as the acquisition of undisclosed information by third parties who knew, or were grossly negligent in failing to know, that such practices were involved.

[4] According to the Oxford Dictionary On Line: "Walpole formed the word on an old name for Sri Lanka, Serendip. He explained that this name was part of the title of a silly fairy tale, called The Three Princes of Serendip: as their highnesses traveled, they were always making discoveries, by accidents and sagacity, of things which they were not in quest of...." The notion of serendipity in scientific research has been extensively studied by Robert Merton (see Merton and Barber 2004).

secretly used, it can either leak out (by which the innovation becomes public knowledge) or can be replicated by one duplicator. By contrast, patenting provides the innovator with certain but temporary protection from both the risk of leakage and duplication. Innovations will differ by the ease with which they can be duplicated.

We take that the duplicator is entitled to a valid patent, but cannot exclude the first inventor from the innovation (that is, being the first inventor is a defense to infringement). This assumption is in line with the patent law of most countries, except the US.[5] Under our assumption, successful duplication leads to duopoly, which can last for the duration of the patent if the duplicator patents, or forever (conditional on the secret not leaking out) if the duplicator also prefers to conceal the innovation. In a companion paper (Denicolò and Franzoni 2003), we analyze the relative merits of the first inventor defense against infringement, and conclude that it is generally not desirable.

Critics of the contract theory argue that the disclosure motive alone cannot justify patents. To have innovators patenting and disclose innovations, so the argument goes, patent life must be at least as long as the expected duration of the secret. But this implies that the expected deadweight loss associated with patents must be at least as large as that associated with secrecy, which means that patents cannot improve social welfare unless they lead to greater innovative effort, as suggested by the reward theory.[6]

---

[5] In 1999, prior user rights were introduced (*American Inventors Protection Act*), but limitedly to inventors of "business methods."

[6] The patent system was fiercely attacked in the middle of the 19th century, when patents were perceived as obstacles to free trade. See Penrose and Machlup (1950). At the meeting of the prestigious Statistical Society of London of 1862, J.E.T. Rogers critically assessed traditional justifications for the patent system. When reviewing the contract theory, he argued that the content of the contract was unfair: what kind of contract is this, where the innovator keeps the best innovations for himself and gives the worse (i.e. those which can hardly be kept secret) to the state? For a (recent) critical perspective on the patent system, see, for instance, Boldrin and Levine (2003).

We contend that this critique is misleading in that it implicitly posits that secrecy involves only the risk of involuntary disclosure of the secret to the public. However, trade secret law does not protect the inventor from independent rediscovery, so concealing an innovation involves also the risk that the innovation is duplicated by somebody else. The possibility of duplication alters the comparison between patents and secrets in two ways. First, all efforts exerted for the replication or independent rediscovery of the innovation are socially wasteful (in view of the non-rival nature of the innovation). Second, duplication leads to an industrial structure where few firms interact on the same market (with monopoly followed by duopoly, triopoly, etc.) Industry evolution under patent, instead, is usually "drastic," as monopoly for the statutory duration of the patent is followed by perfect competition after the patent expires (assuming that the patent is broad enough to prevent imitators from "patenting around"). Thus, even if the innovator gets the same profits from secrecy and patenting, the deadweight loss associated with the two options will generally differ.

Our main finding is, indeed, that the disclosure motive alone suffices to justify patents of positive length: that it to say, early "disclosure" through the patent specification is socially valuable. More precisely, we show that for marginal innovations (i.e., those for which the innovator is just indifferent between patenting and not), patenting is always socially preferable to secret use. There are two effects at work: first, having innovators patent decreases the deadweight loss associated with the exploitation of the innovation if the competition in the market served by the innovation is relatively weak. Second, patenting prevents wasteful duplicative effort. In our model, the latter effect, compounded with the former, makes patenting socially desirable independently of the intensity of market competition. However, patents also entail a cost, as they cannot be tailored to each individual innovation. This means that (infra-marginal) innovations will end up being over-protected, causing an excessive deadweight loss. Thus, an increase in the patent life has two effects on social welfare: the positive effect is that it induces disclosure of marginal innovations, which is socially desirable; the negative one is that it increases the deadweight loss associated with infra-marginal innovations which would have been disclosed anyway. The optimal patent duration has to strike a balance between these two effects. We can show, however, that the optimal patent life is positive. The reason is that, when the patent life is sufficiently short, the rent offered to inframarginal innovations is very small: thus, only the positive effect (disclosure) is at work, making the patent system socially desirable.

Finally, we analyze how the optimal patent policy depends on key economic variables. We find that the optimal patent duration is negatively related to the "size" of the innovation (the optimal patent is shorter, if the market served by the innovation is bigger). We charachterize the optimal amount of disclosure and show that a smaller fraction of firms should rely on the patent system if competition in the product market becomes more intense,

Section 2 introduces the model, section 3 describes the equilibrium behavior of the players. Section 4 is devoted to the optimal patent life, and section 5 concludes.

2 The model

There is a continuum of patentable innovations characterized by the ease with which they can be replicated. Each innovation occurs in a separate homogeneous good industry, with linear demand function: $P(Q) = a - Q$. The innovation is drastic, and we normalize to zero the post-innovation marginal cost.[7]

For each innovation, there is an innovator (a she) and a potential duplicator (a he). The innovator, who has serendipitously discovered the original innovation, can choose which type of proprietary protection to adopt (patent or secret). If she patents, she becomes a temporary monopolist in her industry. By setting $p_m = \frac{1}{2}a$, she reaps monopoly profit $\pi_m = \frac{1}{4}a^2$ for the duration of the patent, T. The patent is granted in exchange for the disclosure of the innovation: upon expiry of the patent, anybody "skilled in the art" can use it, and hence the price and the innovator's profits are driven to zero.

Alternatively, the innovator can rely on trade secrecy. Here the risks are leakage of the secret, which occurs with exogenous probability $1 - z \in (0, 1)$, or independent rediscovery by the duplicator, which occurs with probability $y \in [0, 1]$. Leakage of the secret has the same effects as expiration of the patent: the innovation becomes public and profits are driven to zero. The parameter z ("strength of secrecy") will be taken as exogenous throughout the paper; it may be seen as an index of the ease by which each innovation can be concealed from the public. In contrast, the probability of duplication y is chosen by the duplicator, who aims at maximizing his expected profits and faces a duplication cost equal to $\frac{1}{2}\alpha y^2$. The parameter α measures the difficulty of duplicating the innovation. It is distributed across the population of innovations with cumulative distribution function $F(\alpha)$, density $F'(\alpha) = f(\alpha) > 0$, and support $[0, \infty)$.

Suppose that the innovator has relied on secrecy. If the duplicator manages to replicate the innovation (by legal means), he in turn has to decide how to protect it. We assume that the second inventor is entitled to a valid patent, but cannot exclude

[7] A drastic innovation does not leave any space to producers relying on the old technology. Production with the old technology is not profitable, even if the innovator charges a monopoly price.

the first inventor from the innovation (that is, being the first inventor is a defense to infringement).[8] It follows that if the duplicator patents, there will be duopoly for the duration of the patent. If the duplicator instead relies on trade secrecy, there will be duopoly with probability z; with probability $1 - z$, the secret leaks out and perfect competition prevails.[9]

The duopoly equilibrium will be characterized as follows. Denoting by $Q_d = ka$ aggregate duopoly output, we let k range from $\frac{1}{2}$ (perfect collusion, $Q_d = Q_m = \frac{1}{2}a$) to 1 (perfect or Betrand competition, $Q_d = a$), as competition increases, without making any specific assumption on the nature of competition in the product market. The equilibrium price will be $p_d = a - Q_d$, and each firm will earn duopoly profits $\pi_d = \frac{1}{2}k(1-k)a^2 \leq \frac{1}{2}\pi_m$. Thus, k is an index of the strength of product market competition.

All future profits are discounted at the common discount rate r. It is convenient to define

$$\text{“\textit{normalized}” \textit{patent length}: } \tau \equiv \left(1 - e^{-rT}\right). \tag{1}$$

τ equals the share of overall discounted profits accruing to the patentee (π_m/r or π_d/r, respectively, depending on whether one or two firms operate in the market). As T goes from 0 to ∞, τ ranges from 0 to 1.

To summarize, the timing of the game played by the firms is as follows. First, the innovator decides whether or not to patent; second, if the innovator has not patented, the duplicator decides his duplication effort; third, upon successful duplication, the duplicator decides whether or not to patent; finally, if neither the innovator nor the

[8] On the first inventor defense, see Denicolò and Franzoni (2003).

[9] For simplicity, we assume that the probability of leakage is independent of whether or not the innovation has been duplicated.

duplicator patent, Nature determines whether the innovation leaks to the public or not.

At the outset, the law-maker sets the patent policy, i.e. a patent length T. The law-maker does not know the marginal duplication cost α of each innovation, but knows the distribution function $F(\alpha)$. So the law-maker chooses T, or, equivalently, τ, so as to maximize expected social welfare, where the expectation is calculated over $F(\alpha)$. By contrast, firms make their decisions after uncertainty has been resolved: each innovator knows how difficult it is to duplicate her own innovation.

3 Firms' behavior

We solve the model proceeding backwards. We start by considering the decision problems of the innovator and the duplicator, for any given level of α (marginal duplication costs). At this stage, the normalized patent duration, τ, is taken as an exogenous parameter. Later we shall analyze the optimal patent life and endogenize τ.

The duplicator's problem. Upon duplication, the second inventor must decide whether or not to patent. If he does not patent, his expected payoff is $\mathcal{V} = z\frac{\pi_d}{r}$: discounted duopoly profits times the probability that the innovation does not leak to the public. If he patents, he has to share the market with the first innovator until the patent expires, obtaining $\mathcal{V} = \int_0^T e^{-rt}\pi_d dt = \tau\frac{\pi_d}{r}$. The duplicator opts for patent or secrecy depending on which provides the highest profits. In particular, patent protection will be chosen when the normalized patent duration (strength of patent) is greater than the probability that the innovation does not leak (strength of secrecy):

$\tau \geq z$.[10] The duplicator's payoff is therefore

$$\mathcal{V} = \begin{cases} z\,\dfrac{\pi_d}{r} & \text{for } \tau < z, \\ \tau\,\dfrac{\pi_d}{r} & \text{for } \tau \geq z. \end{cases}$$

Moving one stage back, let us now consider the duplicator's optimal research effort. The duplicator chooses y so as to maximize

$$\Pi = y\mathcal{V} - \frac{1}{2}\alpha y^2.$$

Provided that $\mathcal{V}/\alpha < 1$, the optimal duplication effort is:

$$y^*(\alpha) = \frac{\mathcal{V}}{\alpha} = \begin{cases} z\,\dfrac{\pi_d}{r\alpha} & \text{for } \tau < z, \\ \tau\,\dfrac{\pi_d}{r\alpha} & \text{for } \tau \geq z, \end{cases}$$

while $y^*(\alpha) = 1$ for $\mathcal{V}/\alpha \geq 1$. Thus, the research effort of the duplicator will increase with the expected profits from successful duplication.

The innovator's problem. The innovator, who already owns the innovation, must decide whether or not to patent. If the innovator patents, she earns monopoly profits until the patent expires:

$$V_P(\tau) = \int_0^T e^{-rt}\pi_m dt = \tau\frac{\pi_m}{r}. \tag{2}$$

If she relies on trade secrecy, her payoff depends on the duplicator's behavior, as described above. Thus,

$$V_{TS}(\alpha) = \begin{cases} [1 - y^*(\alpha)]\, z\dfrac{\pi_m}{r} + y^*(\alpha)\, z\dfrac{\pi_d}{r} & \text{for } \tau < z \;\; (\textit{duplicator opts for secrecy}), \\ [1 - y^*(\alpha)]\, z\dfrac{\pi_m}{r} + y^*(\alpha)\, \tau\dfrac{\pi_d}{r} & \text{for } \tau \geq z \;\; (\textit{duplicator opts for patent}). \end{cases}$$

[10] To fix ideas, we assume that a firm patents if it is indifferent between patenting and not, but our results are independent of this tie-breaking rule.

If the innovation is duplicated and does not leak to the public, which occurs with probability $z[1 - y^*(\alpha)]$, the innovator gets monopoly profits forever. If the innovation is duplicated, the innovator either gets duopoly profits forever scaled down by the probability that the innovation does not leak out (if the duplicator opts for secrecy), or duopoly profits for the duration of the patent (if the duplicator opts for patenting). The duplicator patents when patent protection (τ) is stronger than trade secret protection (z). The patent protects both the duplicator and the innovator (who has a prior user right) from leakage, but has a temporary duration.

Who patents in equilibrium? In principle there are three possible outcomes: nobody patents, the innovator patents, or the duplicator patents. We can prove, however, that the latter outcome cannot be part of a subgame perfect equilibrium of the game.

Proposition 1 *Independently of the nature of the innovation, if the first inventor does not patent, neither will the duplicator.*

This result is due to the fact that the innovator has greater incentives to patent than the duplicator. Intuitively, the duplicator patents only if the normalized patent duration is not less than z, the strength of trade secrecy. Since he cannot exclude prior users, he simply compares the expected level of duopoly profits under the two types of protection. Things are different for the innovator. When the normalized patent duration equals the strength of secrecy, the innovator would be indifferent between patenting and secrecy only if there were no risk of duplication (patent and secret would yield the same amount of expected monopoly profits); with a positive probability of duplication, the innovator definitely prefers to patent (as duplication would scale down her profit from monopoly level to duopoly level). Thus, a normalized patent duration less than z is sufficient to induce the innovator to patent. As a result, if the innovator decides to rely on secrecy, it means that patent protection is very week. In this case, the duplicator will also rely on secrecy.

Using the fact that the duplicator never patents in equilibrium, we can conclude that $\mathcal{V} = z\frac{\pi_d}{r}$. Hence, the optimal duplication effort is

$$y^*(\alpha) = \begin{cases} z\dfrac{\pi_d}{r\alpha} & \text{for } z\dfrac{\pi_d}{r\alpha} < 1, \\ \\ 1 & \text{for } z\dfrac{\pi_d}{r\alpha} \geq 1. \end{cases} \tag{3}$$

Duplication is more likely to take place if duopoly profits are larger, the discount rate is lower, marginal duplication costs are lower, and if the innovation can be more easily concealed from the public.

Which innovations are patented? When deciding whether or not to patent, the innovator will compare patent profits $V_P(\tau) = \tau\frac{\pi_m}{r}$ with secrecy profits

$$V_{TS}(\alpha) = z\left[(1 - y^*(\alpha))\frac{\pi_m}{r} + y^*(\alpha)\frac{\pi_d}{r}\right].$$

Note that $V_{TS}(\alpha)$ is non-decreasing in α: the value of secrecy is positively related to the marginal duplication costs. This means that for any given normalized patent duration τ, those innovations will be patented which can be more easily duplicated (small α). We can then define a patenting threshold $\hat{\alpha}(\tau)$, such that innovators patent if, and only if, $\alpha \leq \hat{\alpha}(\tau)$.

The threshold is defined as follows. First, note that $V_{TS}(\alpha)$ is minimal when the innovation is replicated with certainty, $y^*(\alpha) = 1$. This is the case for $\alpha \leq z\frac{\pi_d}{r}$. Here, the innovator patents only if $\tau\frac{\pi_m}{r} \geq z\frac{\pi_d}{r}$, i.e. only if $\tau \geq \tau_0 \equiv z\frac{\pi_d}{\pi_m}$. For $\tau < \tau_0$, no innovation is patented, and we can let $\hat{\alpha}(\tau) = 0$. Second, for all α, we have $V_{TS}(\alpha) \leq z\frac{\pi_m}{r}$ (the latter is the payoff from secrecy when duplication does not occur). Since $V_P(\tau) = \tau\frac{\pi_m}{r}$, all innovations are patented for $\tau > z$. Thus, for $\tau > z$, we can let $\hat{\alpha}(\tau) = \infty$.

For intermediate values of τ, i.e. for $\tau_0 \leq \tau \leq z$, solving $V_{TS}(\alpha) = V_P(\tau)$ yields

$$\hat{\alpha}(\tau) = \frac{z^2}{z-\tau}\frac{\pi_d}{r}\frac{\pi_m - \pi_d}{\pi_m}, \tag{4}$$

with

$$\frac{\partial \hat{\alpha}(\tau)}{\partial \tau} = \frac{\hat{\alpha}(\tau)}{z-\tau} \geq 0. \tag{5}$$

The threshold $\hat{\alpha}(\tau)$ is depicted in Figure 1. Innovators choose to patent for values of α and τ that lie to the right of the $\hat{\alpha}(\tau)$ curve.

[[Figure 1 about here]]

Inspection of (4) reveals that $\frac{\partial \hat{\alpha}}{\partial z} > 0$ if and only if $\tau < \frac{1}{2}z$, $\frac{\partial \hat{\alpha}}{\partial \pi_m} > 0$, $\frac{\partial \hat{\alpha}}{\partial \pi_d} > 0$ and $\frac{\partial \hat{\alpha}}{\partial r} < 0$.

Using the linear demand function, we can obtain a more explicit expression for $\hat{\alpha}(\tau)$:

$$\hat{\alpha}(\tau) = \frac{z^2}{z-\tau}\frac{1}{r}\frac{1}{2}\left[k(1-k)a^2(1-2k(1-k))\right], \quad \text{for } \tau_0 \leq \tau \leq z, \tag{6}$$

with $\tau_0 = 2zk(1-k)$, $\alpha_0 = \frac{1}{2}zk(1-k)a^2$, $\frac{\partial \hat{\alpha}}{\partial a} > 0$ and $\frac{\partial \hat{\alpha}}{\partial k} < 0$. These results can be summarized as follows.

Proposition 2 *The innovation is more likely to be patented if it can be duplicated at a lower cost, it is less easily protected from public disclosure (if $\tau < \frac{1}{2}z$), it yields greater monopoly profits, and it yields greater duopoly profits (as this makes duplication more likely). A larger discount rate tilts the balance in favor of secrecy, as it slows down duplication. With a linear demand function, innovations are more likely to be patented if they are big (large a) and if competition is soft (small k).*

When the market served by the innovation is large, the duplicator invests greater resources to "catch-up." Duplication becomes more likely and secrecy becomes less valuable. As a result, larger innovations are more likely to be patented.[11] A similar argument applies to the strength of competition: when competition is more intense, duplication is less likely to occur. Thus, even if duopoly profits are lower, secrecy becomes more valuable, and patenting becomes less attractive. In the limit case where competition is nearly perfect, duplication does not take place and innovators patent only if the normalized patent duration τ is greater than the probability z that the secret does not leak to the public.

4 Optimal patent length

We now turn to the characterization of the optimal patent policy. The optimal patent length, τ^*, maximizes expected social welfare, which is defined as the expected value of the discounted social returns from the innovation, less duplication costs.

Social welfare. The social returns from the innovation typically exceed the private returns as they also include consumers' surplus. Let S_m, S_d and S_c denote the instantaneous social returns from the innovation under monopoly, duopoly, and perfect competition, respectively, and let $S_c - S_m = \Delta_m$, and $S_c - S_d = \Delta_d$ denote the usual deadweight losses. With the linear demand function, we have $\Delta_m = \frac{1}{2}\pi_m = \frac{1}{8}a^2$, and $\Delta_d = \frac{1}{2}(a - Q_d)^2 = \frac{1}{2}a^2 (1-k)^2$.

Assuming that the social discount rate equals r, when an innovation is patented,

[11] Recall that each innovation drives the marginal cost to 0, so the size of the market (a) equals the size of the innovation.

expected social welfare is

$$\begin{aligned} W_P(\tau) &= \int_0^T e^{-rt} S_m dt + \int_T^\infty e^{-rt} S_c dt \\ &= \tau \frac{S_m}{r} + (1-\tau) \frac{S_c}{r}, \end{aligned} \tag{7}$$

i.e., society obtains a flow of monopoly surplus S_m for the duration of the patent plus a flow of perfect competition surplus S_c thereafter.

If the innovation is protected by trade secrecy, we have

$$W_{TS}(\alpha) = [1 - y^*(\alpha)]\, z \frac{S_m}{r} + y^*(\alpha)\, z \frac{S_d}{r} + (1-z) \frac{S_c}{r} - \frac{1}{2} \alpha \left[y^*(\alpha)\right]^2. \tag{8}$$

Society obtains a permanent flow of monopoly surplus S_m if the innovation is not duplicated and the secret does not leak out. If the innovation is duplicated and the secret does not leak out, it obtains a permanent flow of duopoly surplus S_d. If the secret leaks out, it obtains a permanent flow of perfect competition surplus S_c. In all cases, society now also bears the duplication cost $\frac{1}{2}\alpha \left[y^*(\alpha)\right]^2$.[12] Clearly, W_P depends on the patent life τ but is independent of marginal duplication costs α; while W_{TS} depends on α but is independent of τ.

We assume that the law-maker chooses a patent life, τ, which applies to all innovations independently of the ease with which they are duplicated. This reflects the fact that duplication costs are typically unverifiable and cannot be used to tailor the patent length to each individual innovation. Given a normalized patent duration τ, we know that innovators will patent if $\alpha \leq \hat{\alpha}(\tau)$. Social welfare can be written as

$$\mathcal{W}(\tau) = \int_0^{\hat{\alpha}(\tau)} W_P(\tau)\, f(\alpha) d\alpha + \int_{\hat{\alpha}(\tau)}^{\infty} W_{TS}(\alpha)\, f(\alpha) dz, \tag{9}$$

[12] Recall that we have assumed that the second inventor invests in duplication before uncertainty over the leakage of the secret is resolved. If second inventor invests after this uncertainty is resolved, the duplication cost is borne with probability z, but this would not affect our results.

where the first integral vanishes for $\tau < \tau_0$, and the second for $\tau > z$.

Le us consider the effect of a change in the normalized patent duration on social welfare. Differentiating (9) , we get:

$$\frac{d\mathcal{W}(\tau)}{d\tau} = [W_P(\tau) - W_{TS}(\hat{\alpha}(\tau))] f(\hat{\alpha}(\tau)) \frac{\partial \hat{\alpha}(\tau)}{\partial \tau} + \int_0^{\hat{\alpha}(\tau)} \frac{\partial W_P(\tau)}{\partial \tau} f(\hat{\alpha}(\tau)) d\alpha. \quad (10)$$

An increase in the (normalized) patent duration has two effects: one the one hand, it leads marginal innovators (i.e., those with $\alpha = \hat{\alpha}(\tau)$) to switch from secrecy to patenting; on the other hand, it affects the social value of infra-marginal innovations (which are patented). These two effects are captured by the first and the second term in eq. (10), respectively.

Marginal innovations. Let us start with the first effect. From eqs. (7) and (8), upon simplification (and omitting the argument of $\hat{\alpha}(\tau)$) we get :

$$W_P(\tau) - W_{TS}(\hat{\alpha}) = -[\tau - (1 - y^*(\hat{\alpha}))z] \frac{\Delta_m}{r} + y^*(\hat{\alpha}) z \frac{\Delta_d}{r} + \frac{1}{2}\alpha [y^*(\hat{\alpha})]^2. \quad (11)$$

When the innovator switches from trade secrecy to patenting, social welfare is affected in three ways: the expected deadweight loss from monopoly changes (it goes up, for marginal innovators), the deadweight loss from duopoly vanishes, and duplication costs are saved.

Since at $\alpha = \hat{\alpha}(\tau)$, we have $\tau = z\left[1 - y^*(\hat{\alpha}) + y^*(\hat{\alpha})\frac{\pi_d}{\pi_m}\right]$, eq. (11) rewrites as

$$W_P(\tau) - W_{TS}(\hat{\alpha}) = y^*(\hat{\alpha}) z \frac{\pi_d}{r}\left(\frac{\Delta_d}{\pi_d} - \frac{\Delta_m}{\pi_m}\right) + \frac{1}{2}\hat{\alpha}[y^*(\hat{\alpha})]^2 \quad (12)$$

(this equation holds also for $\tau = \tau_0$). The first term results from the comparison of the deadweight losses associated with patenting and secrecy. It is positive if $\frac{\Delta_d}{\pi_d} > \frac{\Delta_m}{\pi_m}$, that is, if the ratio of deadweight loss to the innovator's profit is greater under duopoly than under monopoly.

To better understand the meaning of the term $\frac{\Delta_d}{\pi_d} - \frac{\Delta_m}{\pi_m}$, let us consider two polar cases. Suppose first that competition is very weak and firms collude in the product market ($Q_d = Q_m$). Upon duplication, consumer surplus remains the same ($\Delta_m = \Delta_d$), but the innovator's profits drop to a half (as she has to share the market). It follows that $\frac{\Delta_d}{\pi_d} = \frac{\Delta_m}{1/2\,\pi_m} > \frac{\Delta_m}{\pi_m}$. In contrast, with tough competition the situation is reversed. The ratio of deadweight loss to the innovator's profit under duopoly is very low (it gets close to zero as the quantity produced converges to that of perfect competition) and $\frac{\Delta_d}{\pi_d} < \frac{\Delta_m}{\pi_m}$, even if duopoly profits are substantially lower that monopoly profits.

With the linear demand function, we have $\frac{\Delta_d}{\pi_d} > \frac{\Delta_m}{\pi_m}$ if, and only if, $Q_d < \frac{2}{3}a$. Recalling that $Q_d = \frac{2}{3}a$ under Cournot competition, it follows that $\frac{\Delta_d}{\pi_d} > \frac{\Delta_m}{\pi_m}$ if, and only if, competition is softer than under Cournot (i.e. $k < \frac{2}{3}$).

The second term in eq. (12) captures the duplication expenditure, and is always positive. It represents the cost associated with the replication of a non-rival good (the innovation). Our result is that, when both terms of eq. (12) are considered, patents are always socially preferable for marginal innovations.

Proposition 3 *The social value of marginal innovations is greater under patenting than under secrecy.*

To prove the Proposition, note that since $y^*(\hat{\alpha}) = \frac{z\pi_d}{\hat{\alpha}r}$ and $\Delta_m = \frac{1}{2}\pi_m$, we have (upon simplification):

$$W_P(\tau) - W_{TS}(\hat{\alpha}) = \frac{z^2}{r^2}\frac{\pi_d}{\alpha}\Delta_d, \tag{13}$$

which is always positive.[13] Due to the duplication costs and the variations in the

[13] With Bertrand competition, duopoly profits as well as duopoly deadweight loss are zero. In this limiting case, patent and secret use are equivalent in terms of deadweight loss, while duplication costs are also zero (there is no incentive to duplicate). Here, patents only protect from inadvertent leak.

innovation's surplus , the increase in social welfare associated with disclosure (i.e. patenting) is larger if competition in the product market is softer (large π_d and Δ_d).

Infra-marginal innovations. Let us now consider the impact of an increase in the patent duration on the social value of patented innovations. In view of eqs. (7) and (8), the second term in eq. (10) reduces to

$$\int_0^{\hat{\alpha}(\tau)} \frac{\partial W_P(\tau)}{\partial \tau} f(\hat{\alpha}(\tau)) d\alpha = -\int_0^{\hat{\alpha}(\tau)} \frac{S_c - S_m}{r} f(\hat{\alpha}(\tau)) d\alpha = -\frac{\Delta_m}{r} F(\hat{\alpha}(\tau)) < 0.$$

An increase in the patent length prolongs monopoly and increases the discounted deadweight loss associated with infra-marginal innovations.

The contract theory of patents. We are now ready to prove the main result of this section.

Note that for $\tau \geq z$ all innovations are patented and hence $F(\hat{\alpha}(\tau)) = 1$. It follows that in this interval, an increase in the patent length can only increase monopoly deadweight loss:

$$\frac{d\mathcal{W}(\tau)}{d\tau} = -\frac{\Delta_m}{r} < 0.$$

Thus, the optimal patent length cannot exceed the strength of secrecy: $\tau^* < z$.

Substituting (5) and (12) into (10), we get

$$\frac{d\mathcal{W}(\tau)}{d\tau} = \frac{\pi_m}{\pi_m - \pi_d} \frac{\Delta_d}{r} \hat{\alpha} f(\hat{\alpha}) - \frac{\Delta_m}{r} F(\hat{\alpha}). \tag{14}$$

The optimal patent length must thus strike a balance between the incentive to induce disclosure (captured by the first term in eq. 14) and the aim of limiting the monopoly distortion induced by patents (the second term in eq. 14). This trade-off differs from the classic Nordhaus trade-off, where the benefit from increased patent protection is faster technical progress. Here, the benefit from increased patent protection is the reduction in secret use of innovations. The cost is the same: prolonged monopoly by (inframarginal) patent holders.

Theorem 4 *The optimal patent length must thus strike a balance between the incentive to induce disclosure and the aim of limiting the monopoly distortion induced by patents. The optimal patent life is not shorter than* $\tau_0 > 0$, *so that at least the weakest innovations are disclosed (i.e. those which are duplicated for sure).*

To prove the second part of the Theorem, we calculate the change in social welfare associated with an increase in the normalized patent length at the point $\tau = \tau_0$, i.e. when patent protection is so weak that only innovations duplicated for sure are patented. Since nobody patents for $\tau < \tau_0$, we have $F(\hat{\alpha}(\tau_0 - \varepsilon)) = 0$ for ε arbitrarily small. Thus, the net effect of increasing patent life from $\tau_0 - \varepsilon$ to τ_0 is, upon simplification,

$$[W_P(\tau_0) - W_{TS}(\hat{\alpha}(\tau_0))]\, F(\hat{\alpha}(\tau_0)) = \frac{z}{r}\Delta_d\, F(\frac{z}{r}\pi_d) > 0.$$

Hence, it is socially desirable to set the patent life so as to induce disclosure at least of the innovations that can be more easily duplicated.

The intuition behind this result should now be clear. Increasing the patent length induces marginal innovators to switch to patenting, which by Proposition 3 is always desirable, but increases the deadweight losses associated with infra-marginal innovations that would have been patented anyway. When $\tau = \tau_0 - \varepsilon$ there are no infra-marginal innovations and so the only effect of increasing patent life to $\tau = \tau_0$ is that those innovations which can be very easily duplicated (with $\alpha \leq \hat{\alpha}(\tau_0)$), will switch from secrecy to patenting. This is welfare improving by Proposition 3 . Hence, the disclosure motive alone is sufficient to justify patents of positive length. Theorem 4 vindicates the contract theory of patents.

Comparative statics. On the assumption that the optimal duration τ^* lies in the interior of the interval (τ_0, z), we can derive some comparative statics results.[14] At the optimum, we have

$$\frac{\partial \mathcal{W}(\tau)}{\partial \tau} = \frac{\pi_m}{\pi_m - \pi_d} \frac{\Delta_d}{r} \hat{\alpha}\, f(\hat{\alpha}) - \frac{\Delta_m}{r} F(\hat{\alpha}) = 0,$$

that is, using $\Delta_m = \frac{1}{2}\pi_m$,

$$\Psi(\hat{\alpha}) \equiv \hat{\alpha} f(\hat{\alpha})\, H - F(\hat{\alpha}) = 0, \tag{15}$$

where $H \equiv \dfrac{2\Delta_d}{\pi_m - \pi_d} = 4\frac{(1-k)^2}{1-2k(1-k)}$, with $\dfrac{\partial H}{\partial k} < 0$.

Equation (15) identifies the optimal patenting threshold, say α^*. This threshold depends on the distribution of the marginal duplication cost as well as on the degree of competition in the product markets. By implicit differentiation, we get $\frac{\partial \alpha^*}{\partial k} < 0$. As competition becomes more intense, the gain from disclosure becomes smaller. It is then desirable that a smaller fraction of firms patent. Since greater competition induces firms to patent less, it is not clear whether the reduction in the fraction of patenting firms requires a shorter patent duration or a longer one.

Note that the optimal amount of disclosure (the threshold α^*) does not depend on a and r . This means that the same fraction of innovations should be patented, independently of the size of the innovations and the level the discount rate. However, these variables affect innovators' propensity to patent, and thus have an impact on the optimal patent duration. We know from Proposition 2 that if innovations become larger, firms are more keen on patenting. The same amount of disclosure can then

[14]Technically, the solution in on the interior if: i) $z\Delta_d f(z\pi_d/r^+) > \Delta_m F(z\pi_d/r^+)$, and ii) $\lim_{\alpha\to\infty} [\alpha f(\alpha) H - F(\alpha)] < 0$.

be obtained with a shorter patent duration. Similarly, it can be established that the optimal patent is longer is discount rate is higher (see the appendix). These remarks can be summarized as follows.

Proposition 5 *If market competition becomes more intense, a smaller fraction of innovations should be protected by the patent system. This may be obtained by either a shorter or a longer patent duration (as firms become less keen on patenting). The size of the innovations and the discount rate do not affect the optimal share of patenting firms. However, since these variables affect firms' propensity to patent, we have that the optimal patent is longer if the discount rate is higher and the innovations are smaller.*[15]

As explained above, the intuition for these results is straightforward. Big innovations are more likely to be duplicated. As a consequence, innovators are more keen on patenting and a shorter patent suffices to induce the optimal amount of disclosure. When the discount rate is higher, duplication becomes less profitable, and innovators are less keen on patenting. Longer patents are therefore required.

When competition is more intense, we have two effects. On the one hand, it is socially preferable that a smaller share of innovations are patented (since disclosure is socially less valuable); on the other, firms are less keen on patenting. The first effect pushes, caeteris paribus, for a shorter patent; the second for a longer patent. The net effect on the optimal patent duration cannot be signed.

[15] For the same reason, one can show that the optimal patent decreases with the strenght of secrecy z if and only if $\tau < \frac{1}{2}z$.

5 Conclusion

It is an historical fact that patents grew as alternatives to trade secrets (see, for instance, David 1993).[16] Many of the first patents (*privilegi*) were granted to people who had not invented the technology at hand, but just ferreted it out from (foreign) trades. [17] In this perspective, they were instrumental the diffusion of jealously held technology.

In this paper we have tried to show that patents may serve a valuable function as alternative to trade secrets. To disentangle the disclosure motive from the traditional reward motive, we have assumed that the innovation is the outcome of "serendipity," so that stimulating R&D effort is not a concern. In such a framework, it was proved that optimal patent length must strike a balance between the benefit of inducing additional firms to disclose their innovations and the increase in deadweight loss associated with patentees' monopoly power.

We have shown that the optimal patent life is positively related to the intensity of competition in the market for the innovation and the discount rate, while it is negatively related to the size of the marker for the innovation.

Our results provide a perspective on the optimal patent policy which is quite different from that based on the unrealistic assumption that innovations are invariably patented. [18] In fact, our model emphasizes those factors which affect innovations *after* they have been created, like the ease of duplication and the degree of market

[16] Even the duration of early English patents was calibrated with reference to the expected duration of trade secrets: 14 years, that is two terms of service of an apprentice.

[17] In England, the system of the privileges was brought into order by the Statute of Monopolies in 1624, which stated that patents could only be granted to the "true and first inventor" of new manifacturs. But again, inventors could just be importers of foreign skills and know-how: "whether they learned by travel or by study, it is the same thing" declared an English court in 1693 (Machlup 1968).

[18] In 1993, the average propensity to patent in the EU was 35.9% for product innovation and 24.8% for process innovations (Arundel and Kabla 1998).

competition. While we have focussed on the classic issue of the optimal patent life, the perspective of the contract theory of patents may shed new light on many other aspects of patent policy. ... This is left for future work.

6 Appendix

■ *Proof of Lemma 1.*

The innovator will choose to patent if $V_P(\tau) \geq V_{TS}(\alpha)$, which means that she patents if patent protection is sufficiently strong. Let $\hat{\tau}(\alpha)$ be the patent duration that makes the innovator indifferent between patenting and not; it is implicitly determined as the solution to $V_P(\hat{\tau}) = V_{TS}(\alpha)$. The innovator will patent if $\tau \geq \hat{\tau}(\alpha)$. Thus, the duplicator patents only if $\tau < \hat{\tau}(\alpha)$ and $\tau \geq z$. We want to show that $\hat{\tau}(\alpha) \leq z$, implying that $\tau < \hat{\tau}(\alpha)$ and $\tau \geq z$ cannot hold simultaneously.

To show that inequality $\hat{\tau}(\alpha) \leq z$, indeed, holds, note that at $\tau = z$ we have (irrespective of whether the duplicator patents or not)

$$V_{TS}(\alpha) = (1 - y^*(\alpha))z\frac{\pi_m}{r} + y^*(\alpha)\, z\frac{\pi_d}{r}$$

whereas $V_P(z) = z\frac{\pi_m}{r}$. Since

$$V_P(z) = z\frac{\pi_m}{r} \geq z\left[(1 - y^*(\alpha))\frac{\pi_m}{r} + y^*(\alpha)\,\frac{\pi_d}{r}\right] = V_{TS}(\alpha),$$

the innovator strictly prefers to patent at $\tau = z$. This means that $\hat{\tau}(\alpha) < z$. Thus, the duplicator will never patent in equilibrium.

■ *Comparative statics*

At the optimum, the second order condition has to be met:

$$\frac{\partial^2 \mathcal{W}(\tau)}{\partial \tau^2} = \frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau} < 0.$$

Since $\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau} \geq 0$, we must have

$$\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha} < 0.$$

Let us consider the relationship between k and the optimal patenting threshold α^* (characterized by equation 12). We have

$$\frac{\partial \alpha^*}{\partial k} = -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial k}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}} < 0,$$

since $\frac{\partial \Psi(\hat{\alpha})}{\partial k} < 0$.

Let us now consider the optimal patent length. We have

$$\begin{aligned}
\frac{\partial \tau^*}{\partial z} &= -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial z}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \tau}} = -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial z}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}} = -\frac{\frac{\partial \hat{\alpha}(\tau^*)}{\partial z}}{\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}} < 0, \quad \text{if and only if } \tau < \frac{1}{2}z, \\
\frac{\partial \tau^*}{\partial r} &= -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial r}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \tau}} = -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial r}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}} = -\frac{\frac{\partial \hat{\alpha}(\tau^*)}{\partial r}}{\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}} > 0, \\
\frac{\partial \tau^*}{\partial a} &= -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial a}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \tau}} = -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial a}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}} = -\frac{\frac{\partial \hat{\alpha}(\tau^*)}{\partial a}}{\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}} < 0,
\end{aligned}$$

where we have used the results of Proposition 2. Furthermore,

$$\frac{\partial \tau^*}{\partial k} = -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial k}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \tau}} = -\frac{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial k} + \frac{\partial \Psi(\hat{\alpha})}{\partial k}}{\frac{\partial \Psi(\hat{\alpha})}{\partial \alpha}\frac{\partial \hat{\alpha}(\tau^*)}{\partial \tau}},$$

which cannot be unambiguously signed.

References

Arundel, A., 2001. The relative effectiveness of patents and secrecy for appropriation. Research policy 30, 611 - 624.

Arundel, A., Kable, I, 1998. What percentage of innovations are patented? Empirical estimates for European firms. Research policy 27, 611 - 624.

Boldrin, M., Levine, D., 2003. The case against intellectual monopoly. Chapters 1 and 2. http://levine.sscnet.ucla.edu/general/intellectual/intellectual.htm

Cohen W., Nelson, R., Walsh J., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). NBER Working Paper W7552.

David, P., 1993. Intellectual property institutions and the Panda's thumb: patents, copyrights, and trade secrets in economic theory and history. in: Wallerstein, M., Mogee M.E., Schoen R. (Eds.), Global dimensions of intellectual property rights in science and technology. National Academy Press, Washington D.C.

Denicolò, V., Franzoni L.A., 2003. Patent, Secrets, and the First Inventor Defense. working paper, www.ssrn.com.

Eisenberg, R., 1989. Patents and the Progress of Science. The University of Chicago Law Review 56, 1017-1086.

Friedman, D., Landes, W., Posner, R., 1991. Some Economics of Trade Secret Law. Journal of Economic Prospectives 5/1, 61-72.

Machlup, F., 1968. Patents, in the International Encyclopedia of the Social Sciences. Collier-MacMillan, London.

Machlup, F., Penrose, E., 1950. The patent controversy in the Nineteenth century. The Journal of Economic History 10/1, 1-29.

Merton, R., Barber E., 2004. The Travels and Adventures of Serendipity: A Study in Sociological Semantics and the Sociology of Science. Princeton University Press.

Nordhaus, W., 1969. Invention, Growth and Welfare: A Theoretical Treatment of Technological Change. Cambridge, MIT Press.

Rogers, J.E.T., 1863. On the rational and working of the patent laws. Journal of the Statistical Society of London 26, 125-.