Finnish Economic Papers – Volume 14 – Number 1 – Spring 2001

ON THE OPTIMAL PATENT POLICY*

TUOMAS TAKALO

Department of Economic and FPPE, P. O. Box 54, FIN-00014 University of Helsinki, Finland

Numerous attempts have been made to identify the optimal mix of patent breadth and patent life. Unfortunately, the range of contradictory results reported in literature is rather impressive. The aim of this note is to develop a very stylised model that encompasses a variety of different findings, and to derive a general rule for the optimal patent policy. (JEL: O34, O31)

1. Introduction

A pervasive obstacle in seeking the optimal technology policy is the public good aspect of intellectual property. On the one hand, intellectual property does not wear out and it is thus wasteful to restrict its use. On the other hand, without the protection of intellectual property, inventors cannot fully appropriate the return on their work, and, in consequence, there is too little innovation in the economy. Accepting that market failure in creating intellectual property rights justifies government intervention raises the question of how intellectual property should be protected and how long. The principal policy tool both in theory and practice has traditionally been patent institution.[1]

Nordhaus (1969) was the first to offer a rigorous model explaining the fundamental trade-off between static and dynamic considerations in designing patent policy: if one wants to spur innovative activity, it is possible only at the expense of the competition. Since Nordhaus's seminal works (1969) and (1972) there has been extensive research on patent protection and its consequences for social welfare. A primary objective of this note is to explain a central finding from this previous research in a simplified framework, and generalise the finding further. It should be evident that such an attempt presupposes considerable simplifications. As David puts it in his article on intellectual property protection in economic theory and history:

'There is no settled body of economic theory on the subject that can be stated briefly without

* *I am most grateful to an anonymous referee, Klaus Kultti, Stephen Martin, Mikko Puhakka, Rune Stenbacka and Mihkel Tombak for helpful comments and discussions. Part of the research was conducted when I was visiting the Department of Economics at the MIT. Financial support from the Yrjö Jahnsson Foundation is gratefully acknowledged.*

[1] *In reality, there are myriad devices to appropriate inventive returns. The legal protection of intellectual property has been traditionally divided into two main branches.* Industrial property protection *deals principally with industrial designs, patents, trademarks and service marks, trade secrets, and appellations of origin.* Copyright protection *usually applies to artistic, audiovisual, literary, musical, and photographic works. The discussion in this article follows the main body of the literature and focuses on the patent protection, although it could be extended with proper modifications to apply to other forms of legal protection such as copyright.*

doing serious injustice to the sophisticated insights that have emerged over many decades of debate. Instead, the relevant economic literature is extensive, convoluted, and characterized by subtle points of inconclusive controversy'. David (1993), p. 23.

The view in David (1993) is no doubt justified, but the recent developments in the field give reasons for a more positive assessment of the theory.

In his first seminal work, Nordhaus (1969) simply shows that the policy-makers' problem is to fine-tune the term of patent protection in order to balance static welfare losses and dynamic welfare gains optimally. As a result, the patent monopoly should last a limited time only. Nordhaus's model thus deals with *patent life* or *patent length*, i.e. the number of years that the patent is in force. His model provides a simple description of the patent system in its original purpose, that is, when a patent confers temporary but complete protection over an invention. The pertinence of this view is, however, much in doubt. Since the pioneering study by Mansfield (1961), researchers have reported overwhelming evidence of the inability of patent protection to prevent imitation with a few exceptions such as the pharmaceutical industry.[2] Nordhaus (1972) thus extends his model to allow imperfect patent protection. In other words, Nordhaus (1972) formalises the concept of *patent breadth* or *patent width*.

While the notion of patent length is indisputable, the meaning of patent breadth, or patent width, is relatively vague. The width of the patent grant measures the degree of the patent protection. If patents are narrow, a patent is easy to 'invent around', that is, it is easy to produce a non-infringing substitute for the patented invention. An extremely narrow patent does not protect even against trivial changes such as changes in colour. This kind of description is too loose to provide an unambiguous ground for the modelling attempts, and the definition of patent breadth in the literature varies from one author to another. Nordhaus's (1972) pioneering model deals with process innovations, and he measures patent breadth by the fraction of the cost reduction not freely spilling over to competitors. In Klemperer's (1990) and Waterson's (1990) product innovation models, patent breadth reflects the distance in the product space between the patented product and the nearest non-infringing substitute. In a similar vein, Matutes, Regibeau and Rockett (1996) define patent breadth by the number of different applications protected by the same patent grant.

The simplest definitions of patent width are provided in Gilbert and Shapiro (1990) and Gallini (1992). In Gallini (1992), the width of the patent is equivalent to an increase in imitation costs caused by patent protection. Such a view is supported by the much-cited queries by Mansfield, Schwartz and Wagner (1981) and Levin, Klevorick, Nelson and Winter (1987). Gilbert and Shapiro (1990) simply identify the patent breadth with the innovator's profit while the patent is in force. In doing so, their analysis also encompasses Tandon's (1982) investigation of the compulsory licensing of patented innovations, because compulsory licensing simply reduces the patentee's profits by facilitating imitation. The compulsory royalty rate, the patent holder's profit with compulsory licensing, can thus be equated with the patent width.

Ambiguous assumptions often lead to ambiguous outcomes, the issue of the socially optimal patent length-breadth mix being no exception That is what David (1993) referred to when he writes about 'inconclusive controversy' in the quotation above. Sometimes the optimal patent has maximum length and minimum breadth, as in Tandon (1982) and Gilbert and Shapiro (1990), sometimes the result is the reverse, as in Gallini (1992), and sometimes the length-breadth mix makes no difference, as in Nordhaus (1972). As if to summarise, Klemperer (1990) provides examples of all these results.

Fortunately, in an excellent article Denicolò (1996) reconciles these seemingly contradictory findings. He demonstrates within a unified framework that the difference in the results reported in the literature is caused by the dissimilar influences of patent breadth on post-inno-

[2] Other empirical studies on the rate of imitation include Mansfield, Schwartz, and Wagner (1981), Mansfield (1985, 1986, 1993), Levin, Klevorick, Nelson and Winter (1987), Harabi (1995), and Arundal and Kabla (1998).

vation profits and social welfare in these models. To be more precise, Denicolò's (1996) theorem predicts that the optimal patent has maximum breadth and minimum length, when both the incentive to innovate and the post-innovation social welfare are convex functions of the patent breadth, the reverse being true if they are concave. Whilst Denicolò's theorem is convenient, it fails to provide policy advice when the second derivatives of these functions take the opposite signs. A major aspect of my assignment here is to advance the theory by deriving a rule for the optimal patent policy that includes also these cases ignored by Denicolò (1996).[3] Another key consideration is to keep the framework as simple and instructive as possible without compromising analytical rigour.

The model is presented in the next section, and the development of the theory and the central concepts are discussed by means of the model. Section 2 also includes the rule for the optimal patent policy that yields the rule established by Denicolò (1996) as a corollary. These rules are then applied to the Nordhaus model of the optimal patent length and breadth in section 3. The concluding remarks are in section 4.

2. *Optimal patent length and breadth*

Nordhaus's (1969) question is simply how long should a patent grant stay in force? The policy-makers' problem is to fine-tune the term of patent protection in order to balance the static and dynamic inefficiencies optimally. To begin the discussion, consider an inventor with a strictly convex cost function

$$C(\alpha) = \frac{1}{2}R\alpha^2, \tag{1}$$

where parameter R reflects the exogenous efficiency of the existing invention technology. It is assumed that R is large enough that in all circumstances $\alpha \le 1$ and, accordingly, α can be regarded as the success probability of the invention. For simplicity, I work directly with α instead of treating investment level as a decision variable. With success the inventor accrues monopoly profits π^m during the life of patent, and some competitive return $\bar{\pi}$ after the patent expires and the innovation becomes available to everyone. If imitation or entry to the industry is costly after the patent expires, $\bar{\pi} > 0$. The legal duration of patent protection, often referred to simply as patent length, is denoted by T. The inventor's return on successful inventive effort is thus

$$P(T) = \int_0^T e^{-rt}\pi^m dt + \int_T^\infty e^{-rt}\bar{\pi}dt, \tag{2}$$

and the inventor's problem is thus to choose α so as to maximise

$$\alpha P - \frac{1}{2}R\alpha^2. \tag{3}$$

The solution is

$$\alpha = \frac{P}{R}. \tag{4}$$

Equation (4) exhibits the classical rationale for intellectual property protection – the investment in innovation increases with the duration of protection, that is, $\partial\alpha/\partial T>0$. Similarly the social return on inventive effort is given by

$$S(T) = \int_0^T e^{-rt}W^m dt + \int_T^\infty e^{-rt}\overline{W}dt, \tag{5}$$

where W^m and $\bar{w}$ depict social welfare as the total of consumer surplus and industry profits when the patent is in force and after it expires. The essential distinction between the private and social return on innovation can be seen by contrasting (5) with (3). The private return P increases while the social return S decreases with the term of protection T. The social planner's task is then to

$$\max_T \alpha S - \frac{1}{2}R\alpha^2, \tag{6}$$

subject to (4). The first-order condition is

$$\alpha_T S = \alpha(R\alpha_T - S_T), \tag{7}$$

in which the subscripts denote the partial derivatives. The trade-off between the static and dynamic considerations facing the policymakers can now clearly be observed in (7), which simply shows that optimal patent life equalises

[3] *There also some links between this note and a recent analysis of optimal patent breadth and life by Wright (1999).*

the marginal dynamic gain of prolonged protection with the marginal static loss. In other words, the left-hand side of equation (7) explains how an increase in patent life encourages inventive endeavours, but after the innovation is made, consumers are worse off because inventor's monopoly lasts longer, as conveyed by the last term on the right-hand side. Notice that the increased R&D expenses due to the accelerated innovative effort must also be counted in the welfare losses; this effect is depicted by the term $\alpha R \alpha_T$ on the right-hand side.

As mentioned in the introduction, Nordhaus's seminal model outlined above is particularly unsatisfactory in so far it excludes the possibility that the patented innovation is imitated when the patent is still in force. It is thus worth to allow for an imperfect patent protection. As there is no consensus in the literature how to render patent protection imperfect, it is here opted for the most general notion by assuming that the inventor's profit and social welfare are functions of patent breadth. Let w denote the width of the patent grant. The innovator's profit after successful innovation $\pi(w)$ then depends on patent breadth so that $\pi(1) = \pi^m$ and $\pi(0) = \overline{\pi}$. Similarly, $W(w)$ denotes static social welfare as a function of patent breadth so that $W(1) = W^m$ and $W(0) = \overline{w}$. The strain caused by the static and dynamic inefficiencies manifests itself in the contrary effects of the patent breadth on social welfare and the innovator's profit, i.e. $W'(w)<0$ and $\pi'(w)>0$.

The private and social returns on innovation can now be rewritten as

(8) $$P(T,w) = \int_0^T e^{-rt}\pi(w)dt + \int_T^\infty e^{-rt}\overline{\pi}dt$$

and

(9) $$S(T,w) = \int_0^T e^{-rt}W(w)dt + \int_T^\infty e^{-rt}\overline{W}dt \,.$$

In designing the optimal patent, both length and width have usually been chosen so as to maximise the social utility from existing innovation, constraining the supply of innovation to a predetermined level. In other words, the social planner's problem is to maximise S with respect to T and w, maintaining α as a constant. The first-order condition for the inventor's problem is now re-expressed as

(10) $$\alpha = \frac{P(T,w)}{R} \,.$$

Let $T(w)$ be the patent length which maintains innovation activity at the required level defined by equation (10), and let the term $\underline{T}$ denote the value of T solving equation (10) for perfect patent protection w=1. Similarly, $\underline{w}$ denotes the value of w solving equation (10) when T approaches infinity. To keep the subsequent discussion interesting, the minimum values $\underline{T}$ and $\underline{w}$ are assumed to exist, and to be positive and finite. Differentiating (10) yields

(11) $$\frac{dT}{dw} = -\frac{P_w}{P_T} < 0.$$

According to (11) the policy tools are *substitutes* with regard to innovation, or as Nordhaus (1972), p. 430 says it: 'if breadth is reduced the optimal life must increase to compensate'. The social value of an existing innovation is now $S(w, T(w))$. Totally differentiating $S(w, T(w))$ with respect to w gives

(12) $$\frac{dS}{dw} = -\frac{P_w}{P_T}S_T + S_w \,.$$

Let ε_{ik}, $i \in (P, S)$, $k \in (w, T)$, measure the elasticity of the private and social values of innovation in respect of the policy variables. For example, $\varepsilon_{Pw} = \frac{d \ln P}{d \ln w}$.

PROPOSITION. The optimal patent policy is determined by the following three conditions:

i) If patent length has a relatively large impact on the incentive to innovate, i.e. $\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} < \frac{\varepsilon_{Sw}}{\varepsilon_{ST}}$ holds, the optimal patent has minimum breadth and maximum length, i.e. $w=\underline{w}$ and $T=\infty$.

ii) If patent breadth has relatively large impact on the incentive to innovate, i.e. $\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} > \frac{\varepsilon_{Sw}}{\varepsilon_{ST}}$ holds, the optimal patent has maximum breadth and minimum length, i.e. w=1 and $T=\underline{T}$.

iii) If the relative impacts of patent breadth and length are equal, i.e. $\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} = \frac{\varepsilon_{Sw}}{\varepsilon_{ST}}$ holds, social welfare is independent of the combination of patent breadth and length.

Proof: When (12) is positive, the optimal patent should have maximum breadth and minimum length. When (12) is negative, the opposite holds. If (12) is equal to nought, social welfare is independent of the breadth-length mix. It is easy to demonstrate that

$$-\frac{P_w}{P_T}S_T + S_w \gtreqless 0 \qquad (13)$$

equals

$$\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} \gtreqless \frac{\varepsilon_{Sw}}{\varepsilon_{ST}}. \qquad (14)$$

QED

This outcome is easy to explain. When an increase in patent width curbs post-innovation social welfare relatively more and accelerates innovative activity relatively less than an increase in patent life, it is desirable to make patents as narrow as possible by prolonging patent life correspondingly. This leaves the incentive to innovate unaltered but expands static social welfare. However, if patent width stimulates investment in innovation relatively more than patent length while reducing the post-innovation welfare relatively less, as short a patent life as possible is socially optimal.

It is possible to show that in Gilbert and Shapiro (1990) it holds that $\varepsilon_{Pw} / \varepsilon_{PT} < \varepsilon_{Sw} / \varepsilon_{ST}$, and in Gallini (1992) $\varepsilon_{Pw} / \varepsilon_{PT} > \varepsilon_{Sw} / \varepsilon_{ST}$, but this is an unchallenging exercise, the work having been done by Denicolò (1996), who demonstrates how the findings in different models such as Tandon (1982), Gilbert and Shapiro (1990), Klemperer (1990), and Gallini (1992) follow from his theorem. It is thus reasonable to try to establish a link between Proposition and Denicolò's theorem. Such a link quickly follows upon the introduction of $D(w)= \overline{w}-W(w)$ as the static dead-weight loss assigned to the patent protection, and $I(w)=\pi(w)-\overline{\pi}$ as a measure of the relative incentive to innovate.

COROLLARY. (Denicolò, 1996). The optimal patent policy is determined by the following three conditions:

i) If both static social welfare $S(w)$ and relative incentive to innovate $I(w)$ are convex in patent breadth, with at least one being strictly so, the optimal patent has maximum breadth and minimum length, i.e. w=1 and $T=\underline{T}$.

ii) If both $S(w)$ and $I(w)$ are concave in patent breadth, with at least one being strictly so, the optimal patent has minimum breadth and maximum length, i.e. $w=\underline{w}$ and $T=\infty$.

iii) If both S and I are linear in patent breadth, social welfare is independent of the combination of patent breadth and length.

Proof: It is easy to see that (14) is equivalent to $-\frac{P_w}{S_w} \gtreqless -\frac{P_T}{S_T}$. Differentiating (8) and (9), it almost immediately follows that $-P_w/S_w$ is equivalent to I_w/D_w, and $-P_T/S_T$ is equivalent to I/D. By rearranging (12) and substituting I_w/D_w for $-P_w/S_w$ and I/D for $-P_T/S_T$, one can verify that the sign of dS/dw is determined by the sign of $\psi(w)=I_wD-D_wI$. The rest goes as in Denicolò (1996). Taking the derivative of $\psi(w)$ with respect to w yields $\psi_w=I_{ww}D-D_{ww}I$. Clearly, if I_{ww} and S_{ww} are positive, $\psi_w>0$, and if I_{ww} and S_{ww} are negative, $\psi_w<0$. Because $\psi(0)=0$, the sign of ψ_w determines the sign of dS/dw.

QED

Denicolò's theorem is illustrated in Figure 1 which depicts the case in which both the incentive to innovate $I(w)$ and the post-innovation social welfare $S(w)$ are convex functions of patent breadth w. It is clear from the figure that increasing the scope of the protection from $\underline{w}$ exponentially boosts the innovative activity but only slightly diminishes the post-innovation social welfare, suggesting that the socially optimal patent should be as broad as possible.

Whilst Denicolò's theorem is convenient, it fails to predict the optimal patent design when the second derivatives of functions $S(w)$ and $I(w)$ take the opposite signs. In such circumstances one must rely on Proposition. It should be pointed out that such circumstances are by no means exceptional. An example in which the incentive to innovate is concave but the post-innovation social welfare is convex in patent breadth is discussed in detail in Takalo (1998), and displayed in Figure 2. As it stands, Figure 2 supports the optimality of the maximum patent breadth as the incentive to innovate is more elastic with respect to patent breadth than the post-innovation social welfare.

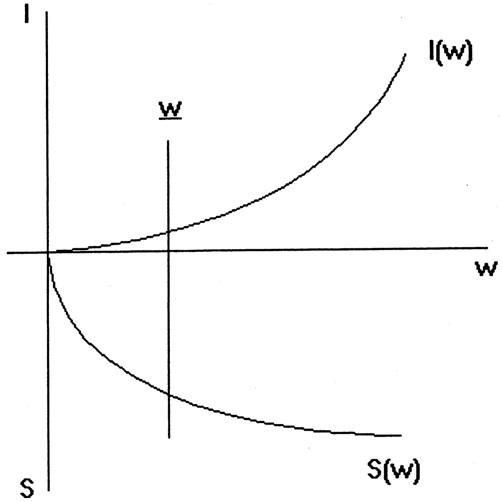

Figure 1. The incentive to innovate $I(w)$ and the post-innovation social welfare $S(w)$ as convex functions of patent breadth w.

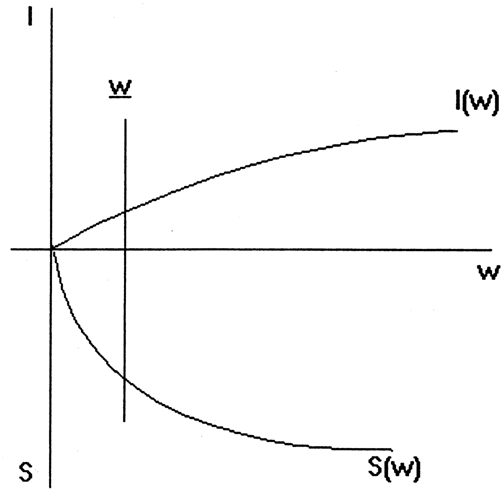

Figure 2. The incentive to innovate $I(w)$ as a concave function of patent breadth w and the post-innovation social welfare $S(w)$ as a convex function of w. The minimum patent breadth making patenting profitable is denoted by $\underline{w}$.

In assessing the reliability of the observations here some caveats should be borne in mind. First, Proposition provides only sufficient conditions for optimal policy. It does not cover cases where the expression $\varepsilon_{Sw} / \varepsilon_{ST} - \varepsilon_{Pw} / \varepsilon_{PT}$ changes its sign as w changes. Second, the incentive to innovate in Denicolò (1996) arises from the equilibrium of a stochastic patent race, whereas here it is determined by a much simpler maximisation problem in which the competitive pressure from other innovators is ignored. As the discussion concerns elasticities, however, adapting more instructive formulation of innovative activity involves no loss of generality. Finally, the underlying assumption in deriving Proposition and Corollary is that an increase in patent length or width invariably stimulates the incentive to innovate and diminishes static social welfare. This assumption covers the most usual cases, and the model above satisfies it. In some special circumstances, however, as in Klemperer (1990) and Waterson (1990), S_w may be positive, and the signs of the inequalities in Proposition should be the reverse, because the proof of the proposition requires dividing by S_w. Changing the signs immediately shows that the optimal patent should have maximum breadth and minimum length when $S_w > 0$ – a heuristic finding indeed.

3. *An example: Nordhaus's model of optimal patent life and breadth*

In this section I further illustrate Proposition and Denicolò's theorem by reconsidering Nordhaus's model of the optimal patent life and breadth (see Nordhaus, 1969, chapter 5, and Nordhaus, 1972). Nordhaus (1969) and (1972) consider a homogenous good industry with the demand function $Q = a-\eta p$, where p and Q denote price and output, and η measures the price elasticity of demand. By employing the technology in (1) the innovator can now reduce the marginal cost of production c so that the size of the cost reduction θ is an increasing and concave function of the investment in invention α. The inventor's post-invention marginal cost is thus $c-\theta(\alpha)$. The invention is non-drastic, that is, the innovating firm cannot drive its competitors out of the market. The competitors' marginal cost in the post-innovation market equilibrium given by c-(1-w)θ depends on patent width. The invention is assumed to be licensed to all firms in the industry with a royalty rate equalling the cost reduction not freely spilling over. The royalty rate is thus θw.

After the patent expires there is free entry, which entirely dilutes the inventor's profit, that is, $\bar{\pi}= 0$. Normalising the level of output before invention to unity,[4] the return on innovation can be written as

$$P(T,w) = I(T,w) = \int_0^T e^{-rt}\eta w dt \,. \tag{15}$$

The static social welfare S is given by

$$S = \int_0^T e^{-rt}\eta w dt + \int_T^\infty e^{-rt}\left(\theta w + \frac{\eta\theta^2 w}{2}\right)dt + \int_0^\infty e^{-rt}(1-w)\left(\theta + \frac{\eta\theta^2}{2}\right)dt \,. \tag{16}$$

The first integral in equation (16) represents the inventor's profit when the patent is in force, the second integral captures the increase in consumer surplus after the patent expires, and the last depicts the effect of the spillover on consumer surplus.

Though not explicitly shown, it is apparent that the optimal patent policy in Nordhaus (1972) is independent of the exact combination of the patent breadth and length. From (15) and (16) the linearity of I and S in w is obvious so that Denicolò's theorem implies the independence of social welfare from the width-length mix. In resorting to Proposition we must calculate the elasticities, which is slightly more involved. Clearly, $\varepsilon_{Sw} / \varepsilon_{ST} - \varepsilon_{Pw} / \varepsilon_{PT}\ \varepsilon_{Pw} = 1$ and $\varepsilon_{PT} = \frac{rTe^{-rT}}{1-e^{-rT}}$, and solving for ε_{Sw} and ε_{ST} yields $-\frac{\theta^2\eta(1-e^{-rT})w}{2rS}$ and $-\frac{\theta^2\eta e^{-rT}wT}{2S}$. Therefore, $\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} = \frac{1-e^{-rT}}{rTe^{-rT}} = \frac{\varepsilon_{Sw}}{\varepsilon_{ST}}$, and by Proposition, the policy variable mix is irrelevant for social welfare.

[4] *Nordhaus (1969, 1972) also normalises the marginal cost before invention to unity. As a result, the size of the cost reduction θ becomes equivalent to the size of the relative cost reduction $B=\theta/c$. He then employs B through both his studies. I think, however, this latter normalisation only confuses the reader.*

4. Conclusion

Intellectual property lies at the heart of a modern economy. The performance of the institutions protecting this property is a matter of deep concern in determining future economic well-being. These observations have generated a spectacular growth of theoretical literature on the economics of intellectual property protection over the past 30 years. Economists have for a long time intuitively understood that optimal patent policy depends on the relative effects of patent life and breadth on the incentive to innovate and social welfare. The intuition is formalised in this note. It is shown that optimal patent policy is determined by three conditions. If the marginal rate of substitution of patent life for breadth is larger on the incentive to innovate than on social welfare, the optimal patent has maximum breadth and minimum length. If the same marginal rate of substitution is smaller on the incentive to innovate than on social welfare, the optimal patent has minimum breadth and maximum life. For the special case when patent life and breadth have equal impacts on the incentive to innovate and social welfare, the mix of policy variables does not matter.

References

Arundel, A., and I. Kabla (1998). "What percentage of innovations are patented?" Empirical estimates for European firms. *Research Policy* 27, 127–141.

David, P.A. (1993). "Intellectual Property Institutions and the Panda's Thumb: Patents, Copyrights, and Trade Secrets in Economic Theory and History." In *Global Dimensions of Intellectual Property Rights in Science and Technology*, 19–61. Eds. M.B. Wallerstein, M.E. Mooge, and R.A. Schoen. Washington, D.C.: National Academy Press.

Denicolò, V. (1996). "Patent races and optimal patent breadth and length." *Journal of Industrial Economics* 44, 249–266.

Gallini, N. (1992). "Patent policy and costly imitation." *RAND Journal of Economics* 23, 52–63.

Gilbert, R., and C. Shapiro (1990). "Optimal patent length and breadth." *RAND Journal of Economics* 21, 106–112.

Harabi, N. (1995). "Approbriability of technical innovation: an empirical analysis." *Research Policy* 24, 981–992.

Klemperer, P. (1990). "How broad should the scope of a patent be?" *RAND Journal of Economics* 21, 113–130.

Levin, R., A. Klevorick, R. Nelson, and S. Winter (1987). "Appropriating the returns from industrial research and development." *Brooking Papers on Economic Activity* 3, 783–820.

Mansfield, E. (1961). "Technical change and the rate of imitation." *Econometrica* 29, 741–766.

– (1985). "How rapidly does new industrial technology leak out?" *Journal of Industrial Economics*, 34, 217–223.

– (1986). "Patents and innovation: an empirical study." *Management Science* 32, 217–223.

– (1993). "Unauthorized use of intellectual property: effects on investment, technology transfer, and innovation." In *Global Dimensions of Intellectual Property Rights in Science and Technology*, 107–145. Eds. M.B. Wallerstein, M.E. Mooge, and R.A. Schoen. Washington, D.C.: National Academy Press.

Mansfield, E., M. Schwartz, and S. Wagner (1981). "Imitation costs and patents: an empirical study." *Economic Journal* 91, 907–918.

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

Nordhaus, W. (1969). *Invention, Growth and Welfare.* Cambridge, Mass.: MIT Press.

– (1972). "The optimal life of the patent: Reply." *American Economic Review* 62, 428-431.

Takalo, T. (1998). "Innovation and imitation under imperfect patent protection." *Journal of Economics* 67, 229–241.

Tandon, P. (1982). Optimal patents with compulsory licensing. *Journal of Political Economy* 90, 470–486.

Waterson, M. (1990). "The economics of product patents." *American Economic Review* 80, 860–869.

Wright, D. J. (1999). "Optimal patent breadth and length with costly imitation." *International Journal of Industrial Organization* 17, 299–454.