PATENT RACES AND OPTIMAL PATENT BREADTH AND LENGTH*

VINCENZO DENICOLÒ

This paper reexamines the issue of optimal patent breadth in extending the earlier literature to the case where many firms race for a patent. It also discusses several examples that suggest the relevance of the nature of competition prevailing in the product market to explain the diverse results found in the literature. Loosely speaking, the less efficient competition in the product market, the more likely it is that broad and short patents are socially optimal.

I. INTRODUCTION

The patent system promotes research and development (R&D) awarding monopoly power to innovators. To avoid excessive monopoly power, governments usually fix a finite patent duration. Less obviously, but at least as important, the extent of monopoly over the new technology may be limited in a number of ways: through compulsory licensing, allowing other firms to “invent around” the patent etc. All these aspects determine what has been called the “breadth” of a patent. Recently, a few papers have addressed the problem of the optimal patent breadth-length mix.

Gilbert and Shapiro (henceforth G-S, [1990]) claim that the optimal design would call for patents of infinite length, with breadth adjusted so as to provide a pre-specified reward to the patentee. This result parallels a similar one obtained by Tandon [1982], who examined the case of compulsory licensing.

More precisely, G-S find a general sufficient condition for an infinite patent duration to be optimal, namely that social welfare be decreasing and concave in the innovator’s profits (taken as a measure of patent breadth). From their analysis it is clear that if social welfare were convex in the breadth of the patent, then patents of maximum breadth and minimum length would be optimal. Nonetheless, G-S show that if the product is homogeneous and firms compete in prices, their sufficient condition for infinite patent duration generally holds.

By way of contrast, Klemperer [1990] has developed a model with product differentiation and price competition where either infinite or

---

* I wish to thank Jeroen Hinloopen, two anonymous referees and the Editor of this Journal for their comments on an earlier draft.

1 See Nordhaus [1969] and [1972], and Scherer [1972] for a classic analysis of optimal patent life.
minimum patent length may be optimal. Though Klemperer’s model is quite general, there are cases where either a maximum length or a maximum breadth completely eliminate the monopolistic distortions associated with the patent. So, in his example with inelastic individual demand where maximum patent breadth is optimal, social welfare is actually increasing in the breadth of the patent. The example is by no means pathological, but it must be said that a case for maximum patent breadth based on the assumption of a positive relationship between patent breadth and (static) social welfare is not compelling.

Short patent lives are also found to be optimal by Gallini [1992] in a different context. She considers the case where the innovation can be perfectly imitated at a cost whose size depends on the breadth of the patent. With homogeneous product and price competition, clearly no imitation would occur in equilibrium. But, under different assumptions about the nature of competition prevailing in the product market (for instance, Cournot), imitators would enter until their profits are driven to zero. Gallini then proves that broad patents are optimal because they lower the socially wasteful imitation costs.

These authors do not analyse the choice of the level of R&D investment that the patent system should generate. Instead, they take the socially desired R&D investment as pre-specified, and study the efficient way to incentivate firms to invest in R&D exactly that amount. In particular, they assume that the innovator should be provided with a pre-specified reward. One aim of this paper is to show that this assumption is restrictive, and is based on a peculiar view of the strategic interaction between innovating firms. Modelling the patent race explicitly, we analyse in more detail the incentives to invest in R&D. Thus we consider firms which compete in the product market, and also compete for obtaining an innovation. In this context, we show that the innovator’s profits are just one component of the firms’ incentive to innovate.

We then generalise G-S’s condition for maximum patent length, and the dual condition for the optimal patent length to be minimum, to this richer framework. In the framework adopted in the early literature, the social problem is equivalent to minimising the ratio of the deadweight losses associated with the patent to the innovator’s profits. In the more general framework studied in this paper, the denominator of the ratio to be minimised is an expression which measures the firms’ incentive to innovate and involves the profits earned by non-innovators and the profits earned after the patent expires, as well as the patentee’s profits.

This extension allows us to consider examples that shed new light on the debate on the optimal patent design. (We shall also reconsider some

---

2 A similar point is made by Waterson [1990], Proposition 1.
examples already discussed in the literature.) The examples confirm that almost anything could happen. The interesting question is, what are the key determinants.

A tentative answer (which will be further detailed in the final section of the paper) is as follows. Any definition of patent breadth involves the idea that narrowing patent breadth leads to more competition in the product market. This lowers the flow of profits earned by the innovator and may increase those earned by non-innovators, thus reducing the incentive to innovate. The effect on social welfare is, however, ambiguous. It might happen, as in Klemperer’s example, that social welfare does not increase. If that is the case, then clearly maximum breadth is optimal. If instead, as is more likely, social welfare does increase, the point is whether it increases more or less rapidly than the incentive to innovate decreases as the patent is narrowed. This depends on the nature of competition. Competition of the Bertrand variety in a homogeneous product market, which reduces the equilibrium price but preserves production efficiency, is the most efficient type. In this case, the deadweight loss decreases more rapidly than the incentive to innovate and therefore the G-S result applies. But focussing on Bertrand competition may underestimate the value of breadth relative to length, because other forms of competition may not be so efficient. For instance with Cournot competition, narrowing the breadth of the patent tends to increase the output of less efficient firms which may be undesirable. In such cases, maximum patent breadth may turn out to be optimal.

Throughout the paper, we consider the case of a single invention. The results can be applied to a group of independent inventions, but when there is a sequence of related innovations, important dynamic problems arise which would considerably complicate the analysis.3

The rest of the paper is organised as follows. In section II, we describe several possible interpretations of what may be meant by patent “breadth”. In section III, the patent race is analysed and the incentives to innovate are identified in terms of post-innovation profits. Section IV extends G-S’s result to this richer setting. Section V illustrates by examples. Section VI interprets the results and concludes the paper.

II. CONCEPTS OF PATENT “BREADTH”

While the concept of length of a patent is clearcut, what may be meant by patent “breadth” is less straightforward. In this section, we briefly

---

3 See Merges and Nelson [1994], Green and Scotchmer [1995] and Chang [1995] for a discussion and analysis of some of these problems.
© Blackwell Publishers Ltd. 1996.
illustrate some possible interpretations—and measures—of the breadth of a patent.4

First, consider the case of a process innovation. To fix ideas, assume that marginal costs are constant and that before the innovation all firms have the same cost \( c \). Then, the innovating firm reduces its own cost to \( c - d \), where \( d \) measures the cost improvement. A wide patent implies that the new production process cannot be imitated and therefore the non-innovating firms will stick to their pre-innovation cost \( c \). But if the patent is more narrowly defined, one can imagine that even the non-innovating firms can develop similar processes without infringing the patent and therefore reduce their costs to a certain extent. The breadth of the patent may be measured by the fraction of the cost reduction that does not spill out as freely available technology to the non-innovating firms. Thus, denoting the breadth of the patent by \( 1 - \alpha \), with \( 0 \leq \alpha \leq 1 \), the non-innovating firms will have a marginal cost equal to \( c - \alpha d \). This is the interpretation suggested by Nordhaus [1972]. The same idea applies to the case of quality improvements.

Second, consider Klemperer’s case of a product innovation with differentiated products. Then one may measure the breadth of a patent by the distance (in some characteristics space) between the patented product and the products that other firms can sell without infringing the patent. In this context, a wider patent implies a higher demand curve for the patentee. The exact way in which the demand curve shifts as the breadth of the patent varies depends on the structure of the market.

Third, a wider patent may mean that it is more costly to imitate the innovation. Then one can measure the breadth of a patent by means of the cost of imitation. This is the route followed by Gallini [1992], who assumes that the cost of imitation is fixed.

Fourth, the breadth of a patent may determine the number of applications of an innovation in independent markets which are reserved for the patentee, as in Matutes et al. [1996].

More generally, our analysis can be applied when there are two instruments available to reward the innovator. For instance, assume as in Tandon [1982] that there is compulsory licensing of the innovation. Then the royalty rate is an additional instrument that can be used along with the patent’s life, analogously to the breadth of the patent.

To encompass all these (and possibly other) interpretations of the patent breadth, we index by \( \alpha \), with \( 0 \leq \alpha \leq 1 \), the degree of dissemination of technological knowledge allowed by the patent: \( \alpha = 0 \) means maximum protection against imitation and \( \alpha = 1 \) means that the patent is so narrowly defined that there is actually free access to the

---

4 Klemperer [1990] discusses some historical examples.

© Blackwell Publishers Ltd. 1996.
new technology. Thus \((1 - \alpha)\) is a measure of the breadth of the patent.\(^5\)

III. THE PATENT RACE AND THE INCENTIVES TO INNOVATE

In the literature cited in the introduction, the incentive to innovate is identified with the prize accruing to the patentee, i.e., the discounted sum of its post-innovation profits.

This identification may be appropriate in some contexts; for instance, when there is only one firm doing R&D. Alternatively, it may be appropriate if the patent race is modelled according to the "winner takes all" assumption as in Loury [1979] and Lee and Wilde [1980]. In these models, the prize to the losers of the patent race is zero.\(^6\)

However, in a more general setting one must take into account the possibility that the losers of the patent race get positive profits in the post-innovation equilibrium (and also, although this turns out to be less relevant to the issue addressed in this paper, that before the innovation firms make positive profits).\(^7\) Moreover, since one must allow the length of the patent to be finite, even if the post-innovation profits of the losers are zero while the innovation is protected by the patent, the prize to the losers may be positive for they may get positive profits after the patent has expired. In these cases, as shown by Beath et al. [1989], the equilibrium level of R&D is determined by the "profit incentive" (i.e., the difference between the patentee's profits and its pre-innovation profits) and the "competitive threat" (i.e., the difference between the profits to the winner and to the losers). Therefore, the incentives to innovate are not simply measured by the flow of profits accruing to the patentee.\(^8\)

To summarise: fixing the incentive to innovate is equivalent to fixing the discounted profits of the innovating firm if there is just one firm doing

\(^5\) G-S use the post-innovation profits of the innovating firm (before the patent expires) as a general measure of the breadth of the patent. We prefer to take the breadth of the patent explicitly as a parameter, because this facilitates the presentation and discussion of the examples. Clearly, though, when the post-innovation profits of the patentee are a strictly increasing function of \(\alpha\), the "reduced form" approach of G-S is equivalent to ours.

\(^6\) This is the appropriate approach when firms doing R&D are pure laboratories which patent the innovation and then license the new technology to other firms operating in the downstream product market. By way of contrast, we consider vertically integrated firms which do R&D and compete in the product market.

\(^7\) Think for instance of the case of a non drastic cost reducing innovation, when firms are quantity-setting Cournot players in the product market. For an analysis of this case, see Delbono and Denicolo [1991] and example 1 below.

\(^8\) Another possible justification of the "winner takes all" hypothesis consists in assuming Bertrand competition in a homogeneous product market with constant marginal costs. Then, before the innovation—when they share the same technology—all firms make zero profits, and after the innovation the winner, which has reduced its own cost, will be the only active firm. Moreover, profits will again fall to zero after the patent expires.

© Blackwell Publishers Ltd. 1996.
R&D or if profits are positive for the innovating firm alone before the patent expires. Since these assumptions appear to be quite restrictive, we shall remove them and explicitly consider a race for a patentable innovation between \( n \) competing firms which are symmetrically placed at the outset.

Each firm \( i \) invests in R&D an amount \( x_i \) per unit of time; \( x_i \) is a flow cost that firm \( i \) pays until one player succeeds.\(^9\) Assuming an exponential distribution, the probability of being successful at a date \( \tau \) prior to date \( t \) is \( \Pr(\tau \leq t) = 1 - e^{-h(x_i)t} \), where \( h(x_i) \) is the (twice differentiable) hazard function which gives the instantaneous conditional probability of success by firm \( i \) as a function of its R&D expenditure \( x_i \). There are decreasing returns in the R&D technology, that is, \( h'(x_i) > 0 \) and \( h''(x_i) < 0 \).

The payoff function of firm \( i \) is the present value of expected profits, net of R&D costs:

\[
\Pi_i = \int_{0}^{\infty} e^{-\left[ \sum_{j \neq i} h(x_j) + r \right] t} \left[ h(x_i)V + X_{-i}L + \pi - x_i \right] dt
\]

\[= \frac{h(x_i)V + X_{-i}L + \pi - x_i}{X_{-i} + h(x_i) + r},\]

where \( r \) is the interest rate, \( x_i \) is \( i \)'s R&D expenditure, \( h(x_i) \) is \( i \)'s instantaneous probability of innovating, \( X_{-i} = \sum_{j \neq i} h(x_j) \) is the instantaneous probability that one of the \( (n-1) \) rivals of firm \( i \) innovates, \( \pi \) is the current profit, \( V \) is the present value of the prize accruing to the winner, and \( L \) is the present value of the prize to the losers.

Denoting the length of the patent by \( T \), the prize to the winner is:

\[V = \int_{0}^{T} \pi_{w} e^{-rT} dt + \int_{T}^{\infty} \pi^{**} e^{-rT} dt = \frac{1 - e^{-rT}}{r} \pi_{w} + \frac{e^{-rT}}{r} \pi^{**},\]

where \( \pi_{w} \) is the flow of profits accruing to the patentee, and \( \pi^{**} \) is the profit after the patent has expired (the same for all firms). Similarly, the prize to the losers of the R&D race is:

\[L = \frac{1 - e^{-rT}}{r} \pi_{L}^{*} + \frac{e^{-rT}}{r} \pi^{**},\]

where \( \pi_{L}^{*} \) is the flow of profits accruing to the non-innovating firms when they have no access to the patented innovation. We assume

\(^9\)We are thus following Lee and Wilde [1980]. However, exactly the same results would be arrived at if instead we assumed that R&D costs are paid at the beginning of the patent race as in Loury [1979].
\( \pi^*_w \geq \pi^{**} \geq \pi^*_l \); the first inequality is strict except for the case of zero patent breadth. This implies \( V \geq L \) with a strict inequality for a positive patent breadth.

Each firm chooses its R&D investment to maximise its expected profits (1). The first order condition for a maximum is:\(^{10}\)

\[
(4) \quad h'(x_i)(X_{-i} + r)V - X_{-i}h(x_i) - h'(x_i)X_{-i}L - h'(x_i)\pi + x_ih'(x_i) - r = 0.
\]

Since all firms are identical, we look for a symmetric Nash equilibrium where \( x_i = x \) for all \( i \). Condition (4) then becomes:

\[
(5) \quad (n - 1)h(x)(V - L) + rV - \pi = \frac{1}{h'(x)} [nh(x) + r - xh'(x)].
\]

We assume that the socially desired R&D effort is predetermined. With a fixed number of firms in the R&D race, this means that the equilibrium R&D investment \( x \) must equal a predetermined level \( \bar{x} \).\(^{11}\) Fixing \( x \) at \( \bar{x} \) in equation (5), substituting for \( V \) and \( L \) and rearranging, we get:

\[
(6) \quad z[(n - 1)h(\bar{x})(\pi^*_w - \pi^*_l) + r(\pi^*_w - \pi^{**})] = K,
\]

where \( z = (1 - e^{-rT})/r \) is a discount factor and \( K = [nh(\bar{x}) + r - \bar{x}h'(\bar{x}) + \pi - \pi^{**}/h'(\bar{x})] \) is a constant.\(^{12}\) Equation (6) is the constraint for the social welfare maximisation problem we shall study in the next section.

Denote the expression inside square brackets in (6) by \( I \), i.e.

\[
(7) \quad I = (n - 1)h(\bar{x})(\pi^*_w - \pi^*_l) + r(\pi^*_w - \pi^{**}).
\]

Equation (6) says that to obtain a pre-specified level of R&D in the equilibrium of the patent race, the discounted value of \( I \) for the duration of the patent must be kept constant. Notice that \( I \) can be interpreted as a measure of the incentive to innovate of firms engaged in the patent race. It is a weighted average of a modified \textit{“profit incentive”}\(^{13}\) \( (\pi^*_w - \pi^{**}) \) and the competitive threat \( (\pi^*_w - \pi^*_l) \). The weight of the competitive threat is \( (n - 1)h(\bar{x}) \), that is the instantaneous probability that firm \( i \) loses the patent race. This makes intuitive sense: the more likely it is that some other firm wins the race, the more important the competitive threat becomes. If instead \( n = 1 \), so that the only firm doing R&D would be sure to win the race, only the profit incentive would matter.

Generally speaking, constraint (6) differs from that considered by G-S [1990] and Klemperer [1990], who assumed a constant present value of the

\(^{10}\) It may be easily checked that the second order condition holds.

\(^{11}\) Since \( n \) and \( h \) are given, fixing \( x \) also implies that the expected date of innovation is predetermined.

\(^{12}\) We assume \( K > 0 \). This means that if the innovation is not protected, the incentive to innovate falls short of the level required to stimulate the desired amount of R&D investment.

\(^{13}\) Actually, the profit incentive as defined by Beath \textit{et al.} is \( (\pi^*_w - \pi) \).

© Blackwell Publishers Ltd. 1996.
IV. CONDITIONS FOR MAXIMUM OR MINIMUM PATENT LENGTH

In this section we extend the G-S analysis deriving a sufficient condition for the optimal patent length to be maximum (i.e., infinite) or minimum. Recall that \((1 - \alpha)\) is a measure of the breadth of the patent.

Let \(S\) denote the flow of social welfare, i.e., the sum of producers’ and consumers’ surplus. Generally speaking, social welfare \(S\) and the post-innovation profits \(\pi^*_w\) and \(\pi^*_l\) depend upon \(\alpha\). We assume that \(\pi^*_w(\alpha) < 0\) and \(\pi^*_l(\alpha) \geq 0\).\(^{14}\) Recall also that, under our parameterisation, \(\alpha = 1\) describes the situation after time \(T\) when the patent expires.\(^{15}\) Then clearly \(\pi^*_w(1) = \pi^*_l(1) = \pi^*\) which implies \(I(1) = 0\). Moreover, \(I'(\alpha) < 0\).

Regarding \(S(\alpha)\), the most natural assumption would be that instantaneous social welfare decreased with the breadth of the patent, i.e., \(S'(\alpha) > 0\). However, there are non pathological examples where the opposite is true for some \(\alpha\)'s. One example is Klemperer’s [1990] model when all consumers have inelastic demand. They can buy a fixed quantity of the good either from the innovator which produces a high quality good or from imitators which sell inferior brands. A broader patent reduces the market share of imitators and therefore increases social welfare. As another example, consider the case of quantity competition between two firms in the product market: if the cost gap between the high cost firm and the low cost firm is very large, social welfare may increase if the patent is narrowed.\(^{16}\) In both cases, \(S(0) > S(\alpha)\) for \(\alpha\) small; in Klemperer’s example, moreover, \(S(0) = S(1)\). In what follows, we shall explicitly state when condition \(S'(\alpha) > 0\) is assumed to hold. However, we do assume that \(S(\alpha) \leq S(1)\).

The social welfare maximisation problem may be stated as follows: Choose \(\alpha\) and \(T\) so as to maximise total discounted social welfare; that is:

\[
\max_{\alpha, T} \left( \frac{1 - e^{-rT}}{r} S(\alpha) + \frac{e^{-rT}}{r} S(1) \right)
\]

subject to the constraint (6), i.e., \(zI(\alpha) = K\). This problem is equivalent to

\[
\min_{\alpha, T} zD(\alpha)
\]

\(^{14}\)All relevant functions are assumed continuous and twice piecewise differentiable.

\(^{15}\)Since (6) must hold, in the time interval \((0, T)\) the breadth of the patent must be positive, i.e., \(\alpha < 1\), otherwise no firm would have an incentive to invest in R&D.

\(^{16}\)See example 1 below. A similar property is exhibited also by our example 4, which is based upon Gallini [1992].
where \(D(x) = S(1) - S(x)\) denotes the deadweight loss resulting from a patent of breadth \(x\), again subject to (6). Equation (6) defines \(x\) as a function of \(T\). Let \(\tilde{T}\) denote the value of \(T\) that solves equation (6) for \(x = 0\); similarly, let \(\tilde{x} < 1\) denote the value of \(x\) that solves equation (6) when \(T\) tends to \(\infty\).

In words, the social problem is to choose the patent’s length and breadth so as to minimise the discounted deadweight loss over the lifetime of the patent, under the constraint of generating a given incentive to innovate. An increase in the length of a patent (and hence in \(x\)) multiplies by the same factor both the present value of the deadweight loss \(D(x)\) and that of the incentive \(I(x)\). Therefore, if the constraint is binding, the optimal patent’s breadth is the one that minimises the ratio \(\psi(x) = D(x)/I(x)\).

Clearly, this presupposes that \(S'(x) > 0\). If social welfare were not decreasing in the breadth of the patent, the constraint (6) would not bind and maximum patent breadth would be optimal. More generally, there cannot be an interior solution \(x^*\) to the social problem with \(S'(x^*) < 0\).

Assume that \(S'(x^*) < 0\) so that the constraint (6) is binding. We have

\[
\psi(x) = \frac{D'(x)I(x) - D(x)I'(x)}{[I(x)]^2}
\]

The sign of \(\psi(x)\) equals the sign of the numerator of the r.h.s. of (10). Notice that the numerator vanishes at \(x = 1\), and that its derivative is \(D'(x)I(x) - I'(x)D(x)\). It follows that if \(D''(x) > 0\) and \(I''(x) < 0\) (actually, it suffices that one inequality be strict) the numerator is increasing in \(x\); since it is zero at \(x = 1\), it must be negative for \(0 \leq x < 1\). Similarly, if \(D''(x) < 0\) and \(I''(x) > 0\) the numerator is decreasing in \(x\) and therefore it must be positive for \(0 \leq x < 1\).

The above discussion may be summarised as follows.

**Proposition 1.** Assume \(S'(x) > 0\). If \(S''(x) \geq 0\) and \(I''(x) \geq 0\), with at least one strict inequality, the optimal patent has maximum breadth and minimum length, i.e., \(x = 0\) and \(T = \tilde{T}\). If \(S''(x) \leq 0\) and \(I''(x) \leq 0\), with at least one strict inequality, the optimal patent has minimum breadth and maximum length, i.e., \(x = \tilde{x}\) and \(T = \infty\). Finally, if \(S''(x) = 0\) and \(I''(x) = 0\) for all \(x\), social welfare does not depend on the breadth-length mix.

A special case of this proposition arises when the losers of the R&D race have zero profits until the patent expires, i.e., \(\pi_L = 0\), or when \(n = 1\). In these cases \(I'(x) = H\pi_w''(x)\), with \(H\) constant, and we have the G-S result.

Notice that Proposition 1 provides sufficient conditions for maximum

\footnote{We assume that this value exists and is finite; that is, that in case of maximum patent protection the innovation is sufficiently valuable to induce the competing firms to invest in R&D at least the predetermined amount \(n\).}

\footnote{We are here interpreting the constraint as saying that at least a specified R&D effort must be guaranteed, i.e., as a weak inequality.}

© Blackwell Publishers Ltd. 1996.
or minimum patent breadth. In specific applications, a direct approach (i.e., explicit minimisation of \( \psi(x) \)) may prove more powerful. The advantage of the result stated in Proposition 1 is that it is more general and easy to apply, as we shall see in the next section.

V. EXAMPLES

In this section we study some examples which illustrate the application of Proposition 1.

Example 1. The first example is a homogeneous Cournot duopoly with a linear demand function and constant marginal costs. The market demand function is: \( p = a - Q \), where \( p \) is price and \( Q \) is total output. Before the innovation, the two firms produce at constant marginal and average cost, \( c, 0 < c < a \). Then, in the Cournot equilibrium, profits are

\[
\pi = s^2/9 \quad \text{where} \quad s = (a - c).
\]

The innovation reduces the marginal cost to \( c - d \). Assume that the innovation is non drastic so that it does not give monopoly power to the winner; in the present framework this means that \( s > d \). This assumption implies that the loser of the R&D race, even if it continues to produce at cost \( c \), will remain active in the post-innovation Cournot equilibrium.

However, we assume also that the loser of the R&D race may reduce its marginal cost if the patent is narrowly defined, for in this case it may imitate the innovation without infringing the patent. Let \( \alpha \) be the fraction of the cost reduction that spills out to the loser. Then before the patent expires, the winner of the R&D race will produce at cost \( c - d \) and the loser will have costs \( c - \alpha d \). When the patent expires, \( \alpha = 1 \) and both firms will produce at cost \( c - d \).

Routine calculations show that in the post-innovation equilibrium profits are \( \pi_w^* = \frac{1}{9} [s + (2 - \alpha)d]^2 \) and \( \pi_L^* = \frac{1}{9} [s - (1 - 2\alpha)d]^2 \), and consumers’ surplus is \( CS = \frac{1}{18} [2s + (1 + \alpha)d]^2 \).

Thus one gets:

\[
(11) \quad S(\alpha) = CS + \pi_w^* + \pi_L^* = \frac{4}{9} \left[ s + \frac{(1 + \alpha)}{2} d \right]^2 + \frac{(1 - \alpha)^2}{2} d^2,
\]

whence it follows that \( S'(\alpha) = \frac{4}{9} d(4s - 7d + 11ad) \) and \( S''(\alpha) = \frac{4}{9} d^2 > 0 \). Notice that \( S'(0) > 0 \) if \( s > 7/4d \). If this inequality is reversed, for \( \alpha \) low enough social welfare decreases with \( \alpha \). Thus a cost reduction of the high cost firm may be socially disadvantageous if its market share is very low.\(^{19}\)

\(^{19}\)When the cost of the high cost firm decreases, its market share becomes larger and therefore average industry cost may increase. This negative effect may outweigh the positive welfare effect of the increase in output that the cost reduction brings about: see Katz and Shapiro [1985] and also Tirole [1988, ch. 10].

© Blackwell Publishers Ltd. 1996.
Nonetheless, since it is not optimal to set $x$ in the interval where $S'(x) \leq 0$, we can proceed assuming that constraint (6) is binding.

Twice differentiating the expressions for post-innovation profits yields $\pi_{w}''' = \frac{3}{2} d^2$ and $\pi_{w}''' = \frac{3}{2} d^2$. This implies that $I''(x) > 0$ if $r > 3h(x)$.

The above discussion may be summarised in the following proposition.

**Proposition 2.** In the case of a cost reducing innovation in a linear homogeneous Cournot duopoly with constant marginal costs, the optimal patent length is minimum and the optimal patent breadth is maximum if $r > 3h(x)$.

**Example 2.** Consider a product innovation in a vertically differentiated industry. We assume that each consumer buys one unit of a good obtaining utility.

\begin{equation}
U = m\theta - p,
\end{equation}

where $p$ is price, $\theta$ is the quality level and $m$ is a parameter measuring the willingness to pay for quality. This parameter is distributed over the unit interval $[0, 1]$ with density $f(m)$ and distribution function $F(m)$. The number of consumers is normalised to 1, i.e., $F(1) = 1$. Initially, the good cannot be produced. The innovating firm can produce a good of quality $\theta$ with zero costs. Competitive imitators can produce (also at zero costs) a good of quality $x\theta$. Competition amongst imitators implies that quality $x\theta$ will be offered at zero price.\(^20\) Obviously, $\pi_{t} = \pi_{**} = 0$.

The patentee's output is $Q = 1 - F(\bar{m})$, where

\begin{equation}
\bar{m} = \frac{p}{(1 - x)\theta}
\end{equation}

denotes the consumer who is indifferent between buying the low quality and the high quality good. Then choosing $p$ is equivalent to choosing $m$. Denote by $m^*$ the profit maximising value of $m$; it can be easily verified that $m^*$ is independent of $x$. This implies that the patentee's profits $\pi_{w} = (1 - x)\theta m^*[1 - F(m^*)]$ and social welfare

\begin{equation}
S = \int_{0}^{m^*} x\theta dm + \int_{m^*}^{1} \theta dm = \frac{\theta}{2} [1 - m^*(1 - x)]
\end{equation}

are linear in $x$. Thus we may conclude:

**Proposition 3.** In the case of product innovation in a vertically differentiated industry where consumers have utility functions given by

\(^20\)This example is clearly related to Klemperer's [1990] model. Unlike in Klemperer, however, individual demand and transport costs are correlated since both depend on $m$. © Blackwell Publishers Ltd. 1996.
(12), the patent breadth-length mix does not affect discounted overall social welfare.

Example 3. Tandon [1982] studies the optimal patent length in the presence of compulsory licensing of the innovation. He shows that, with a linear demand function, the optimal patent length is infinite. We provide a new derivation and a generalisation of this result.21 Let \( Q(p) \) be the demand function and assume that the innovation makes it possible to produce the good at a constant marginal cost \( c \). The innovation must be licensed at a royalty fee which is determined by the regulator. Here we take the fee as a measure of the patent's breadth \( \alpha \). More precisely, let \( p_M \) denote the monopoly price corresponding to the constant marginal cost \( c \). Clearly, an unregulated innovator will find it optimal to set the royalty fee at \((p_M - c)\). Let \((1 - \alpha)\) be the fraction of the optimal fee that the patentee is allowed by the regulator to charge. Then the good will be sold competitively at an equilibrium price equal to the production cost plus the royalty fee, i.e., \( p = c + (1 - \alpha)(p_M - c) \). In this case, \( \alpha = 0 \) means that the innovator is actually unregulated, whereas \( \alpha = 1 \) means that there is complete dissemination of technological knowledge.

The profits of the innovating firm before the patent expires are \( \pi_w = (1 - \alpha)(p_M - c)Q(p) \), where \( p \) is the equilibrium price. Differentiating twice one gets \( \pi_w''(\alpha) = 2(p_M - c)^2 Q'(p) + (1 - \alpha)(p_M - c)^3 Q''(p) \). Since \( I(\alpha) \) is equal to \( \pi_w'' \) times a constant, it follows that \( I'(\alpha) < 0 \) if \( Q'(p) \leq 0 \).

Social welfare is \( S = \int_0^p pdQ - cQ \). Clearly, \( S'(\alpha) = -(1 - \alpha)(p_M - c) \times Q'(p) > 0 \) and \( S''(\alpha) = (p_M - c)^2 Q'(p) + (1 - \alpha)(p_M - c)^3 Q''(p) \). Proposition 1 then yields:

Proposition 4. In the case of compulsory licensing of a cost reducing innovation with constant marginal costs, the optimal length of the patent is infinite if the demand function is linear or concave \((Q'(p) \leq 0)\).

Example 4. We consider now an example which is a simplified version of the Gallini [1992] model. Gallini [1992] assumes that there is only one firm doing R&D and that either the innovation creates a new product or, equivalently, that the cost innovation is drastic. However, the innovation can be imitated at a fixed cost \( h \), and there is free entry of imitators. Therefore, in the post-innovation equilibrium the profit of the innovator will equal \( h \) until the patent expires. Then imitation becomes costless and profits are driven to zero. In this framework, the breadth of the patent may be taken to influence the cost of imitation \( h \). To be more precise,

---

21 Tandon [1982] works with a linear demand function, though he claims that his result extends to more general demand functions.

© Blackwell Publishers Ltd. 1996.
assume a linear demand function \( Q = a - p \) and a constant marginal cost \( c \). Again let \( s = (a - c) \). Then one may set:

\[
(15) \quad h = \left(1 - \alpha\right) \frac{s^2}{4}
\]

that is, the imitation cost is a fraction of the monopoly profit. If \( \alpha = 0 \) no imitator will enter the market so that there is perfect patent protection\(^{22}\); if \( \alpha = 1 \) imitation is free, the post-innovation price falls to \( c \) and the innovator’s profits are competed away.

Clearly, in the post-innovation equilibrium \( \pi^*_{w} = h \) and \( \pi^*_L = \pi^{**} = 0 \). Since the imitators gain zero profit, social welfare equals the sum of consumers’ surplus and the patentee’s profit, i.e., \( S = CS + h \). By standard calculations, social welfare turns out to be:

\[
S(h) = \frac{1}{2} s^2 - s\sqrt{h} + \frac{3}{2} h. \tag{23}
\]

This implies:

\[
S'(h) = -s/(2\sqrt{h}) + \frac{3}{2} \text{ and } S''(h) = s^2/2 h^{1/2} > 0. \tag{24}
\]

Since \( h \) is a linear function of \( \alpha \), it also follows that \( S'(\alpha) > 0 \). Hence:

**Proposition 5.** In the case of a cost reducing innovation in a market with linear demand function, constant marginal costs and a fixed imitation cost with free entry, the optimal breadth of the patent is maximum and the optimal length is minimum.

**Example 5.** Consider a market with horizontal differentiation à la Hotelling. Consumers are uniformly distributed along the unit line \((0, 1)\); the transport cost is linear and the unit transport cost is denoted by \( t \). Two firms, 1 and 2, are located at 0 and 1, respectively. Their location is fixed. Initially, they produce two goods of the same quality \( \theta \) at zero costs. They are also engaged in a patent race to obtain an innovation which raises the quality level to \( \theta + \tilde{\theta} \) (again at zero productive costs). The loser of the race, however, can imitate and produce a good of quality \( \theta + \alpha \tilde{\theta} \), so that \( \alpha \) is, as usual, an inverse measure of the breadth of the patent.

Each consumer buys at most one unit of the good. We assume that a consumer buying one unit of good \( i \) \((i = 1, 2)\) obtains utility:

\[
U = \theta_i - td_i - p_i
\]

\(^{22}\) Actually, entry is blockaded even if the imitation cost is slightly greater than duopoly (not monopoly) profit; however, since we treat the number of firms as a continuous variable it is easier to work with the parameterisation given by (15).

\(^{23}\) The equilibrium number of firms, price and output level are \( n = s/\sqrt{h} - 1, p = (a + nc)/(1 + n) \) and \( Q = n\sqrt{h} \), respectively.

\(^{24}\) To see this, note that \( S'(h) \) may be positive, and hence \( S'(\alpha) \) may be negative, for values of \( h \) high enough. But notice that \( S'(h) \) is positive only if \( h > s'/9 \), that is only if the fixed cost is higher than duopoly profits, so that entry is actually blockaded: see footnote 22.

\(^{22}\) Gallini [1992] takes a direct approach, i.e., she minimises \( \psi(\alpha) \), and proves the same result under more general conditions. Our indirect approach clarifies the formal relationship between Gallini’s findings and the G-S general result: at least under our simplifying assumptions, in Gallini’s model social welfare is convex in the patentee’s profit.
where \( p \) is price, \( \theta \) is the quality level, and \( d \) is the distance that he has to travel.\(^{26}\)

Since before the patent race the two firms are symmetric, we can assume without loss of generality that firm 1 innovates. Then, \( \theta_1 = \theta + \theta \) and \( \theta_2 = \theta + a\theta \). We assume that \( \theta \) is high enough for the market to be covered. The two firms compete in prices.

In the post-innovation equilibrium, prices are \( p_1 = t + \frac{1}{3}(1 - x)\theta \) and \( p_2 = t - \frac{1}{3}(1 - x)\theta \). The market share of the innovator (or, equivalently, the consumer who is indifferent between buying the high quality and the low quality good) is\(^{27}\)

\[
(17) \quad x = \frac{1}{2} + \frac{1}{6t}(1 - \alpha)\theta
\]

The post-innovation profits are \( \pi_1 = [t + \frac{1}{3}(1 - x)\theta]x \) and \( \pi_2 = [t + \frac{1}{3}(1 - x)\theta](1 - x) \). Social welfare is \( S = \theta_1 x + \theta_2 (1 - x) - \frac{1}{2} t x^2 - \frac{1}{2} t (1 - x)^2 \).

One can then easily verify that \( \pi''(x) = \pi''_2(x) = \theta^2/18 \), so that \( I''(\alpha) > 0 \) and \( S''(\alpha) = 5/36(\theta^2/t) > 0 \). We can therefore state:

**Proposition 6.** In a market with linear transport costs and horizontal differentiation, the optimal breadth of the patent is maximum and the optimal length is minimum.

A variant of this example assumes that the two firms produce the same quality but can change their location. More precisely, suppose the locations are initially fixed at 0 and 1, respectively, but the two firms compete for an innovation which gives the capability to move towards the centre.\(^{28}\) The innovating firm can move to \( b \) (or \( 1 - b \)); however, the innovation can be imitated to a certain extent, so that the non-innovating firm can also move towards the centre to \( 1 - ab \) (or \( ab \)). In this example it turns out that \( I''(\alpha) > 0 \) and \( S''(\alpha) < 0 \) so that neither of the sufficient conditions of Proposition 1 hold; indeed, the social problem has an interior solution.

---

\(^{26}\)In all the previous examples, narrowing the patent does not enlarge the technological possibilities of the industry as a whole, though it may enlarge those of particular firms. In other words, if a social planner could choose the first best allocation, this would not be affected by the breadth of the patent. The positive effect on social welfare (if any) is brought about by the more intense competition associated with a narrower patent. This example features a direct technological benefit of narrowing the patent breadth, as well as the indirect benefit associated with more intense competition. The reason is that when \( \theta < t \), it is socially efficient that some consumers continue to buy the low quality good.

\(^{27}\)We assume that the innovation is not drastic, that is \( \theta < 3t \).

\(^{28}\)As is well known, in the Hotelling model with linear transport cost both firms have an incentive to move towards the centre of the market. One has to assume \( b < 1/4 \) to avoid problems of non-existence of a pure strategy equilibrium.
VI. CONCLUDING REMARKS

In this paper we have extended the analysis of the optimal patent breadth-length mix to the case of a patent race where the "winner takes all" assumption may not hold.\textsuperscript{29} We have also analysed a series of examples which show that there is no presumption that either infinite or minimum patent length is most likely to be optimal.

Our discussion has clarified the relationship between G-S's "general" result on the optimality of an infinite patent duration and Klemperer and Gallini's counterexamples. In Klemperer's example, social welfare is (locally) \textit{increasing} in the breadth of the patent;\textsuperscript{30} in Gallini's model (at least under certain simplifying assumptions) social welfare is \textit{convex} in the patentee's profit. By way of contrast, G-S's result requires that social welfare be decreasing and concave in the patent breadth.

We have shown that the patent breadth-length optimal mix depends in a subtle way (involving second derivatives) on the relationship between social welfare and post-innovation profits, on the one hand, and the breadth of the patent, on the other hand. And economic theory places no restriction on the concavity of these functions. Thus it should not be surprising that different models and examples yield seemingly contradictory conclusions. But what is the economic intuition underlying these diverse results? That is, what are the economic forces which in any particular example determine the optimal shape of the patent?

We suggest the following answer. Generally speaking, reducing the breadth of a patent leads to more competition in the product market after the innovation. We know that more competition is not always socially desirable. Whatever "more competition" exactly means, it may involve social costs, like duplication of entry costs, inefficient production, and so on. We also know that different forms of competition exhibit various degrees of efficiency; for instance, in a homogeneous market Bertrand competition is more efficient than Cournot competition.

Clearly, if the additional competition brought about by narrowing the patent is on balance socially costly, it is optimal to award patents of maximum breadth. And, for a reduction in the patent breadth to be socially optimal it does not suffice that more competition increases social welfare: it must increase social welfare more than it reduces the incentive to innovate of the firms participating in the patent race.

G-S show that this is indeed the case with Bertrand competition and a

\textsuperscript{29} Another paper where the "winner takes all" assumption is not made is Waterson [1990]. However, his analysis focuses on other issues, like the choice whether to patent the innovation or not or the possibility of litigation over the scope of the patent.

\textsuperscript{30} This is not to say that Klemperer's model, which is quite general, cannot exhibit cases where a maximum breadth is optimal even if social welfare is decreasing in the breadth of the patent. Our comment applies to his example with inelastic individual demand only.

\textcopyright{} Blackwell Publishers Ltd. 1996.
homogeneous product. But this is the case most favourable to the G-S thesis, as in this case competition reduces the equilibrium price while preserving production efficiency. Since competition is not always so efficient, this result cannot be deemed a general one. Loosely speaking, the less efficient is the type of competition prevailing in the product market, the more likely it is that broad and short patents are socially optimal. Broad patents reduce the output of less efficient firms with Cournot competition and avoid wasteful duplication of entry costs when imitation is costly. With differentiated products and price competition, broad patents generally involve social costs but may be very effective in widening the difference between the winners' and losers' rewards, thus increasing the incentive to innovate at a relatively low cost.

VINCENZO DENICOLÒ, 
Dipartimento di Scienze Economiche, 
Università di Bologna, 
Strada Maggiore 45, 
I-40125 Bologna (BO), 
Italy

REFERENCES


© Blackwell Publishers Ltd. 1996.


© Blackwell Publishers Ltd. 1996.