RAND Journal of Economics
Vol. 21, No. 1, Spring 1990

How broad should the scope of patent protection be?

Paul Klemperer*

I explore the trade-off between a patent's length *(that is, its lifetime) and its* width *(that is, its scope of coverage). A wider patent generally reduces the distortion of consumers' choices between the patented brand of the product and unpatented, lower-priced varieties sold by competitors but also permits higher prices, which increase (relative to profits) the deadweight losses from consumers switching consumption out of the product class. I show under what conditions infinitely lived but very narrowly focused patents are the socially efficient way to reward innovation and under what conditions very short-lived but very broad patents are optimal.*

1. Introduction

■ The optimal *length* of patent protection has been extensively analyzed.[1] Another important policy question, however, that has not been satisfactorily addressed, is, What is the optimal *width* of patent protection? For example, if a company invents a new drug to alleviate a heart condition, how similar a drug should a competitor be allowed to sell? If a computer software firm markets a new program, how different should any rival product be required to be?[2]

This kind of question has been answered differently in different periods of U.S. history and is today answered differently in different countries. The most famous early patent application is that of Eli Whitney, whose 1794 invention of the cotton gin changed the course of U.S. history.[3] Unfortunately for Whitney, patent width at that time was essentially

---

* St. Catherine's College and Institute of Economics and Statistics, Oxford University.

I am very grateful to Alan Beggs, Jeremy Bulow, and Meg Meyer for numerous helpful comments and suggestions. I have also had useful advice from many other colleagues, including James Mirrlees, John Ockendon, Ivan Png, Richard Romano, Garth Saloner, Suzanne Scotchmer, Graeme Segal, John Vickers, and Michael Waterson, from the referees, and from participants in seminars in Bristol, Cambridge, London, Munich, Oxford, and Paris.

[1] See, for example, Nordhaus (1969) and Scherer (1972).

[2] In this example, copyright law as well as patent law is relevant. (See Menell (1987).) Major computer software copyright suits being litigated at the beginning of 1989 included Apple's suit to prevent Hewlett-Packard, Microsoft, and others from implementing "user-interface technologies" similar to that of Apple's Macintosh (which allows users to choose commands by pointing at pictures, called icons, on a screen), Lotus' suit against Mosaic Software and Paperback Software for producing spreadsheet programs very similar to its Lotus 1-2-3 program, Ashton Tate's suit against Fox Software for producing database management products that understand Ashton Tate's dBase language, and Nintendo's suit against Atari Games about videogame software. (See Wiegner and Heins (1989) and *The Economist* (March 25, 1989, pp. 78–79) for these and other examples.)

[3] See Schneider (1988) for discussion of this patent application.

zero; under the Act of 1793, would-be competitors had simply to file almost identical applications to receive cotton gin patents of their own and, as a result, Whitney's invention earned him very little. The 1836 and subsequent acts allowed for broader patent width, however. For example, in 1847 Winans applied for a patent on coal cars that were circular in cross-section (and so could carry more weight than traditional cars of rectangular cross-section), and the Supreme Court ruled in 1853 that a competitor's octagonally-cross-sectioned car (which could be thought of as halfway between Winans' and the traditional designs) did indeed infringe Winans' patent.[4] More recently, the inventor of the oversized tennis racket (Howard Head of Prince Manufacturing) was granted U.S. patent protection for racket faces ranging from 85 to 130 square inches.[5] Competitors were forced to produce outside this range. (The conventional racket face at the time was 70 square inches.) Other countries enforce narrower patent widths than the U.S., however. For example, it is a commonly expressed view that patent widths are now very much narrower in Japan than in the U.S.,[6] and indeed, Prince failed to obtain useful patent protection in England, Germany, or Japan.[7] The broader construal of patent claims by American than by foreign courts is supported by the "doctrine of equivalents" which states that a product serving the same function as another may infringe its patent, a doctrine which has no counterpart in, for instance, Japan.[8]

In this article I am interested in the trade-off between patent width and patent length. That is, I am interested in designing patents of optimal shape that allocate a given profit reward, V, to the innovator at the least social cost. I do not address the question of what size V should be.[9]

Increasing the lifetime of a patent simply multiplies by the same factor the present values of both the monopoly profit resulting from the patent and its associated deadweight loss. (I assume that the innovator and society use the same discount rate.) To award total monopoly profit, V, with the least deadweight loss, therefore, I choose the patent width that minimizes the deadweight loss per dollar of profit generated by the patent and then "multiply by" the lifetime required to generate the correct total profit. (See Section 2.)[10]

Section 3 sets out my basic model. I assume that all consumers prefer the same variety of a product but differ both in their demands and costs of substituting alternative varieties, costs which I model as "transport costs" in the standard way. The patentholder produces the single variety that all consumers prefer; the width of the patent is the distance from the preferred point to that point beyond which competitors are allowed to produce.

In Section 4 I analyze the welfare losses, relative to the competitive provision of consumers' preferred variety, that are caused by the granting of the patent. These welfare losses are of two kinds: (1) those on consumption that is switched to less-preferred varieties of the product that are unpatented and so are sold competitively (e.g., purchases of 84 square-

---

[4] See Schneider (1988).

[5] See McQuade (1982).

[6] Consider, for example, the view quoted in Melloan (1988). "In Japan," says Mr. Spero [President of Fusion Systems], "it's possible to patent a relatively minor change such as a 'bicycle with red pedals' . . . [An inventor] is vulnerable to an assault of competing applications with minor variations . . . the Japanese patent system [is not designed to] give the inventor's company a chance to obtain some market advantage."

[7] Similarly, English trade sources report that Wilson's patent on its thick-throated "Profile" design of tennis racket was of almost no value in the U.K. Within one season, competitors were able to produce near-copies that did not infringe Wilson's patent.

[8] See Allegretti and Sherry (1988) and Scotchmer and Yoshikawa (1987) for further discussion.

[9] Once I have determined the optimal patent shape for any given V, the problem of choosing V to generate the socially appropriate level of R&D is closely related to standard analyses of optimal patent length when width is not a control variable. (See the references cited in footnote 1 above.) However, this problem, unlike mine, requires information about the supply of innovations.

[10] Section 2 discusses exceptions to this rule.

inch rackets by consumers who would have bought 110 square-inch rackets if the latter had been sold competitively), and (2) those on consumption that is switched out of the product class (e.g., consumers who give up or fail to take up tennis because of the higher prices). Wider patents reduce consumers' freedom to substitute competitively provided, unpatented varieties of the product and so, we shall see, generally reduce losses of the first kind relative to profits. On the other hand, wider patents allow the patentee to charge higher prices (widening Prince's patent to run from 50 to 150 square inches would presumably have raised the profit-maximizing price for 110 square-inch rackets), and higher prices generally raise the second kind of deadweight losses relative to profits. At one extreme, with infinitely wide patents—no competition permitted within the product class (e.g., the patent covering all tennis rackets)—all consumers make the socially correct decision between varieties of the product, so only the second kind of losses matters (Lemma 1). At the other extreme, with arbitrarily narrow patents, only the first kind of deadweight losses matters (Lemma 2). The reason for the latter result is that arbitrarily close competition permits only arbitrarily small price-cost margins that are of first order in the distance from the competitively provided product variety. Profits are therefore of first order, but losses from consumers switching to other product classes are of second order, so this kind of losses is unimportant. However, a nonnegligible fraction of consumers switches to the competitively provided variant of the product, so the first kind of deadweight losses is of first order, and therefore it is only this kind that matters.

I show (Propositions 1 and 2) in Section 5 that if all consumers have the same per-unit transport costs of substituting to a less-preferred variety of the product, then infinitely lived, narrow patents are optimal. The reason is that the patentholder will set prices so that no consumers switch to unpatented brands (in this case, the alternative would be to sell to no consumers), so there are zero deadweight losses of this (first) kind. Therefore, only losses of the second kind arise, and (by Lemma 2) these become arbitrarily small relative to profits as patents become narrower. Thus, long and narrow patents are optimal in this case.

On the other hand (Propositions 3 and 4), if for each consumer, the value of consuming the preferred variety exceeds the value of consuming no variety of the product by the same monetary amount, then short-lived patents that are as wide as possible are optimal. The reason is that the patentholder sets prices so that no consumers substitute out of the product class. (The alternative is for the patentholder to make no sales.) Therefore, only losses of the first kind arise, and (by Lemma 1) these are minimized with wide patents. (In the limit, with an infinitely wide patent, the patentholder has no competitors, so if consumers have a common reservation price, the patentholder charges that reservation price and there are no deadweight losses.)

Section 5 also notes some special cases in which all patent widths are equally efficient.

More generally, the first kind of losses (from substituting the (socially) wrong product variety) is relatively more important, the more elastic the distribution function of consumers' substitution costs between product varieties. The second kind of losses (from consumption switched out of the product class) matters more when the distribution function of consumers' valuations of their preferred variety of the product class is more elastic. I show that when the former (latter) distribution function is more elastic below its monopoly point, the relatively greater concern with minimizing the first (second) kind of losses typically makes the widest (narrowest) possible patent optimal.

Section 6 extends the basic model to allow the patentholder to produce multiple varieties, to analyze consumers with different most-preferred varieties, and to consider alternative specifications of transport costs. I show that my most important results are robust to these extensions.

Closely related to my work is that of Gilbert and Shapiro (1990), who also analyze the trade-off between patent "breadth" and length. They begin their analysis by interpreting breadth very generally as any aspect of patent policy that affects the patentholder's flow rate of profits, and they isolate a property under which infinitely lived patents are optimal.[11] However, they then specialize to interpret breadth as the price-cost margin that the patent-holder is permitted to set in a homogeneous-product market. What I mean by breadth, on the other hand, is the region of (differentiated) product space covered, and I assume that there are no direct controls on the patentholder's price. Under Gilbert and Shapiro's interpretation of breadth, infinitely lived patents are typically optimal.[12] However, my results show that there is no such presumption when a patent's breadth measures how different competitors' products must be in order not to infringe the patent.

I conclude in Section 7.

2. Efficient patent widths

■ I consider a single patent of width w and lifetime L and wish to choose w and L to minimize the present value of social costs subject to the present discounted value of the patentholder's profits, equalling V. I assume a common social and private discount rate of i.[13] Writing $s(w)$ and $\pi(w)$ for instantaneous social costs (relative to competitive provision) and patentholder profits that are generated by a patent of width w, respectively, the problem of determining the optimal patent shape[14] is

$$\min_{\substack{0 \le w \le \infty \\ 0 \le L \le \infty}} \left\{ \int_0^L s(w) e^{-iT} dT \right\}$$

subject to

$$\int_0^L \pi(w) e^{-iT} dT = V.$$

Now rewriting the constraint as $\int_0^L e^{-iT} dT = V/\pi(w)$ and substituting for $\int_0^L e^{-iT} dT$ in the objective function, I can rewrite the objective function as $s(w)V/\pi(w)$. Therefore, noting that since $0 \le L \le \infty$, $w \ge \pi^{-1}(iV)$ is required ($\pi(\cdot)$ is an increasing function), I can reformulate the problem as

(i) choose

$$w = \underset{\pi^{-1}(iV) \le w \le \infty}{\operatorname{argmin}} \; r(w) \equiv \frac{s(w)}{\pi(w)}$$

---

[11] My analysis of efficient patent widths in Section 2 also does not depend on a specific interpretation of width.

[12] Formally, this result corresponds to that for a special case of my model. The reason is that in a homogeneous-product market, there can be no welfare losses caused by consumers switching to less-preferred but unpatented varieties of the product, i.e., no losses of the first kind I identify, so the analysis is the same as that of the special case of my model (consumers with identical transport costs) in which the patentholder sets price such that no losses of the first kind arise. See footnote 23.

[13] I assume that the problem is feasible, i.e., that a patent of infinite width and infinite lifetime yields profits of at least V. Allowing the profits to exceed V would not affect the results, unless there is a minimum feasible patent lifetime. The reason is that a patent yielding larger profits than V is socially more costly than the patent of the same width but shorter lifetime that yields exactly V in profits. I introduce a minimum patent lifetime only in Proposition 4; allowing profits to exceed V would increase the range for which I find the minimum possible lifetime to be optimal in that proposition.

[14] In my model "rectangular" patents—patents with the same width at every time before expiration—are optimal because I assume that social costs (utility losses) are additive across time and demands are independent across time. (See Section 3.) If consumers could substitute across time, patent lifetime and patent width would play more similar roles than in my model. (Increasing either parameter would then reduce competition for the patentholder's good today, so patents of medium widths and lifetimes would probably be optimal in situations analogous to those in which I find very short-lived patents to be optimal.)

and (ii) choose

$$L = \frac{1}{i} \ln \left(\frac{\pi(w)}{\pi(w) - iV} \right)$$

(since (ii) implies that $\int_0^L e^{-iT} dT = V/\pi(w)$).

In other words, I choose the patent width that generates profits in the most efficient way (smallest social cost per dollar of profit generated). I then compute the lifetime that generates the correct total profits; since increasing the patent life multiplies profits and social costs by the same factor, this does not affect the social cost per unit profit.

The initial development is simplified by concentrating on minimizing $r(w)$ without explicit concern for the constraints on w. I will show, however, that when $r(\cdot)$ is minimized at $w = 0$, the minimum feasible w ($w = \pi^{-1}(iV)$) is typically optimal in the model. (See Proposition 2.)[15] Similarly, if the product class is of only finite size or if there is a minimum feasible patent lifetime, and therefore a maximum feasible patent width, then when $r(\cdot)$ is minimized at $w = \infty$, the widest feasible patent is typically optimal. (See Proposition 4.)

3. A simple model

■ To proceed further, I use a simple model that allows me to analyze the trade-off between the different kinds of social costs discussed in the introduction. The patentholder produces a single good for which the aggregate demand per unit time would be $F(p)$ at price p if no other product varieties were available. However, a patent of width (more precisely, radius) w allows competing firms to produce product varieties a distance w from the patentholder's product. I assume that each consumer has a transport cost of t per unit distance per unit purchased of substituting alternative varieties for the patentholder's good.[16] The cost t is distributed with density $g(t)$ and is independent of consumers' individual demands. It will be convenient to write $G(t) = \int_t^\infty g(\tau)d\tau$ for the probability that a given consumer has a transport cost that exceeds t. Without loss of generality, I normalize demand so that $F(0) = 1$.[17] For simplicity, I assume free entry to the industry subject to the noninfringement of the patent and that knowledge of the innovation allows competitors' products to be produced without fixed costs and at the same constant marginal costs as the patentholder's product. Without further loss of generality, I assume marginal costs to be zero. I discuss extensions to this model in Section 6.

4. Analysis of social costs and profits

■ Under the assumptions above, free entry drives the price to zero at the patent boundary. The boundary is a width w from the variety that consumers prefer and that the patentholder produces, so consumers with transport costs t prefer to buy the most-preferred variety at price p rather than to buy the variety at the boundary at price zero if and only if $tw \geq p$.

---

[15] If my model were to include a maximum feasible patent lifetime, $\bar{L}$, the corresponding minimum feasible width $w = \pi^{-1}(iV/(1 - e^{-i\bar{L}}))$ would typically be optimal. Requiring $\underline{L} \leq L \leq \bar{L}$ is equivalent to requiring the additional constraints on w of $\pi^{-1}(iV/(1 - e^{-i\bar{L}})) \leq w \leq \pi^{-1}(iV/(1 - e^{-i\underline{L}}))$ and has little effect on the results.

[16] I obtain this structure of demand if the jth consumer derives net utility $m_j + \phi_j(q) - (p + t_j w)q$ from purchasing q units at price p per unit of a variety a distance w from the most-preferred variety, where m_j is j's income (so, $m_j - pq$ is the money available for purchase of all other goods), $\phi_j(q)$ is j's gross utility from q units, and t_j is j's transport cost per unit distance per unit purchased. One possibility is that each consumer wishes to buy at most one unit of the product, so $\phi_j(q) = 0$ for $q < 1$ and $\phi_j(q) = r_j$ for $q \geq 1$, where r_j is j's reservation price for the most-preferred variety. In this case, $F(p)$ is the number of consumers for whom $r_j \geq p$.

[17] For simplicity, I assume where necessary that $F(\cdot)$ and $G(\cdot)$ are real-analytic functions (so they are also continuously differentiable).

Since consumers' total demand is $F(p)$ and a fraction $G(p/w)$ have transport costs $t \geq p/w$, the patentholder's demand is

$$H_w(p) = F(p)G(p/w).$$

The patentholder chooses the price, $p^*(w)$,[18] that maximizes its profits

$$\pi(w) = \max_p \{pH_w(p)\}.$$

Total social costs (relative to the case of competitive provision in the absence of any patent) are simply the deadweight loss triangle assessed at the profit-maximizing price $p^*(w)$ on the demand curve $H_w(p)$, that is,

$$s(w) = \int_0^{p^*(w)} H_w(\rho)d\rho - \pi(w).$$

These social costs comprise (i) the deadweight losses, $s_n(w)$, on the "nontravelling" consumers (on the consumers whose transport costs are high enough that they buy from the patentholder, if at all, but who purchase less at price $p^*(w)$ than they would at price zero); (ii) the transport costs, $s_{t1}(w)$, incurred on "travelling" consumers (on the remaining consumers who travel to the patent boundary to purchase there (at price zero) if they purchase at all); and (iii) the additional deadweight losses, $s_{t2}(w)$, incurred on travelling consumers, since these consumers purchase less at their effective marginal cost (their travelling cost) then they would at price zero.[19] Thus, $s_{t1}(w)$ represents the deadweight losses of the first kind described in the introduction (consumption switched to the wrong product variety), while $s_n(w)$ and $s_{t2}(w)$ together comprise the second kind (consumption switched away from this product class).

I am interested in minimizing the ratio of total social costs to profits:

$$r(w) \equiv \frac{s(w)}{\pi(w)} = \left[\frac{\int_0^{p^*(w)} H_w(\rho)d\rho}{p^*(w)H_w(p^*(w))} \right] - 1 = \left[\frac{\underset{[0,p^*(w)]}{\text{average}}\, H_w(\cdot)}{H_w(p^*(w))} \right] - 1.$$

I now show that for sufficiently wide patents this ratio depends only on the demand curve $F(\cdot)$, while for sufficiently narrow patents the ratio depends only on the distribution of transport costs, $G(\cdot)$.

Lemma 1. Assume that a monopolist with zero costs facing demand $F(p)$ has a unique optimal price, p^m. Then,

$$\lim_{w\to\infty} r(w) = \left[\frac{\int_0^{p^m} F(\rho)d\rho}{p^m F(p^m)} \right] - 1$$

(which is the ratio of deadweight losses to profits at the monopoly price on $F(p)$).[20]

[18] To avoid some small but uninteresting complications, I assume henceforward that $pH_w(p)$ has a unique maximizer.

[19] Since a fraction $G(p^*(w)/w)$ of consumers is nontravelling,

$$s_n(w) = G\left(\frac{p^*(w)}{w}\right)\left[\int_0^{p^*(w)} F(\rho)d\rho - p^*(w)F(p^*(w))\right].$$

A density $g(t)$ of consumers faces an effective price equal to their total transport costs of wt per unit, so $s_{t1}(w) = \int_0^{p^*(w)/w} g(t)[wtF(wt)]dt$ and $s_{t2}(w) = \int_0^{p^*(w)/w} g(t)[\int_0^{wt} F(\rho)d\rho - [wtF(wt)]]dt$. (To confirm $s(w) = s_{t1}(w) + s_{t2}(w) + s_n(w)$, integrate $s_{t1}(w) + s_{t2}(w)$ by parts and then use the change of variables $\rho = wt$.)

[20] The proofs of all lemmas and propositions can be found in Appendix A.

The intuition is straightforward. If the patent boundary is far enough, (almost) no consumers prefer to travel to the patent boundary rather than pay even the monopoly price. Therefore, the patentholder faces (almost) the same demand as a monopolist, $F(p)$, and charges the monopoly price; since (almost) no consumers travel, the profits and deadweight losses are the same as those for a monopolist. (Note that even as $w \to \infty$, no consumer incurs total per-unit travelling costs above the monopoly price.)

Lemma 2. Assume that a monopolist with zero costs facing demand $G(\tilde{p})$ has a unique optimal price, $\tilde{p}^m$. Then,

$$\lim_{w \to 0} r(w) = \left[\frac{\int_0^{\tilde{p}^m} G(\rho)\,d\rho}{\tilde{p}^m G(\tilde{p}^m)} \right] - 1$$

(which is the ratio of deadweight losses to profits at the monopoly price on $G(\tilde{p})$).

The intuition for Lemma 2 is that with a very narrow patent, only very low prices attract any customers at all. At these low prices, the demand curve $F(p)$ is effectively vertical at $F(0)$ $(=1)$ relative to the transport-cost curve: a doubling of price has almost no effect on demand, $F(p)$ but makes the large proportion of consumers with transport costs between $\frac{p}{w}$ and $\frac{2p}{w}$ per unit distance switch from buying the patentholder's product to travelling and buying at the patent boundary. The patentholder therefore acts (almost) as a monopolist against the transport-cost curve $G(\cdot)$, taking demand for the product as given at $F(0)$.

For small w, therefore, price-cost margins are of order w, so profits from nontravellers, $\pi(w)$, are of order w, but the deadweight losses incurred by nontravellers, $s_n(w)$, are of order w^2. However, a nonnegligible proportion of consumers incurs transport costs of order w, so total transport costs, $s_{t1}(w)$, are of order w. The additional deadweight losses incurred by these consumers, $s_{t2}(w)$, are of order w^2 (since the effective prices faced by these consumers are of order w). Thus, for narrow patents, the only social costs that matter are the transport costs incurred by consumers substituting away from the patentholder's product.[21]

The results of Section 5 follow from this discussion.

---

[21] To confirm my interpretation of Lemmas 1 and 2, I can integrate by parts the expressions in footnote 19 for $s_{t1}(w)$ and $s_{t2}(w)$ to obtain

$$\left(\frac{s_{t1}(w)}{\pi(w)}\right) = \frac{1}{p^*(w)} \int_0^{p^*(w)} \left(\frac{F(\rho)G\left(\frac{\rho}{w}\right)}{F(p^*(w))G\left(\frac{p^*(w)}{w}\right)} - 1 + \frac{\rho F'(\rho)G\left(\frac{\rho}{w}\right)}{F(p^*(w))G\left(\frac{p^*(w)}{w}\right)} \right) d\rho,$$

$$\left(\frac{s_{t2}(w)}{\pi(w)}\right) = \frac{1}{p^*(w)} \int_0^{p^*(w)} \left(1 - \frac{F(\rho)}{F(p^*(w))} - \frac{\rho F'(\rho)G\left(\frac{\rho}{w}\right)}{F(p^*(w))G\left(\frac{p^*(w)}{w}\right)} \right) d\rho,$$

and

$$\left(\frac{s_n(w)}{\pi(w)}\right) = \frac{1}{p^*(w)} \int_0^{p^*(w)} \left(\frac{F(\rho)}{F(p^*(w))} - 1 \right) d\rho.$$

As $w \to \infty$, $G\left(\frac{p^*(w)}{w}\right) \to G(0) = 1$ and $G\left(\frac{\rho}{w}\right) \to G(0) = 1$ for $\rho \le p^*(w)$, since $p^*(w)$ is bounded above, so the integrands of both $\left(\frac{s_{t1}(w)}{\pi(w)}\right)$ and $\left(\frac{s_{t2}(w)}{\pi(w)}\right)$ integrate to zero. Thus, the only losses that matter are those of the nontravellers, and the ratio of these losses to profits, $\left(\frac{s_n(w)}{\pi(w)}\right)$, depends only on $F(\cdot)$. As $w \to 0$, on the other

5. Optimal patent widths and lengths

■ **Conditions under which infinitely lived, narrow patents are optimal.** Lemma 2 showed that for very narrow patents, the only social costs that matter are those incurred by travelling consumers. When all consumers have the same transport costs, however, the patentholder must set a price such that no consumers want to travel. (If any consumer preferred to travel rather than buy the patentholder's product, then all consumers would prefer to travel.) I therefore obtain the following result.

Proposition 1. If all consumers have identical transport costs, then patents of infinitesimal width are most efficient.

In a practical problem, I need strictly positive patent widths, $w \geq \pi^{-1}(iV)$, to give a positive reward, V. The point of Proposition 1, however, is fairly robust.

Proposition 2. If all consumers have identical transport costs, then infinitely lived patents are optimal ($L = \infty$, with minimum possible width $w = \pi^{-1}(iV)$) if

(i) the elasticity of $F(p)$, that is, $(-pF'(p)/F(p))$, is nondecreasing in p, or

(ii) the value of the patent to be awarded, V, is sufficiently small.

Remark 1. Most standard demand curves, including all concave, linear, and not overly convex demand curves, satisfy condition (i).[22] The optimality of infinitely lived patents when consumers have identical transport costs is therefore fairly general.

The intuition for Propositions 1 and 2 is that when all consumers have the same transport costs, they all have the same preference between the patentholder's product and the competitively supplied variant, so the patentholder sets a price such that none substitute the competitively supplied variant. Therefore, the only social costs are the standard deadweight losses, $s_n(w)$, on the demand curve $F(\cdot)$. Now the ratio of deadweight losses to profits is, in general, increasing in price for given demand. This is always true at sufficiently low prices, since profits are of first order in price, while deadweight losses are of second order. It is also true at all prices if the elasticity of demand is nondecreasing, since a firm selling at a higher elasticity earns lower profits relative to the consumer surplus dissipated. Since narrower patents constrain the patentholder to lower prices, the narrowest possible patent is therefore the most socially cost-effective way of awarding profits to an innovator.[23]

Figure 1 illustrates the result by showing the profits and deadweight losses for patents

---

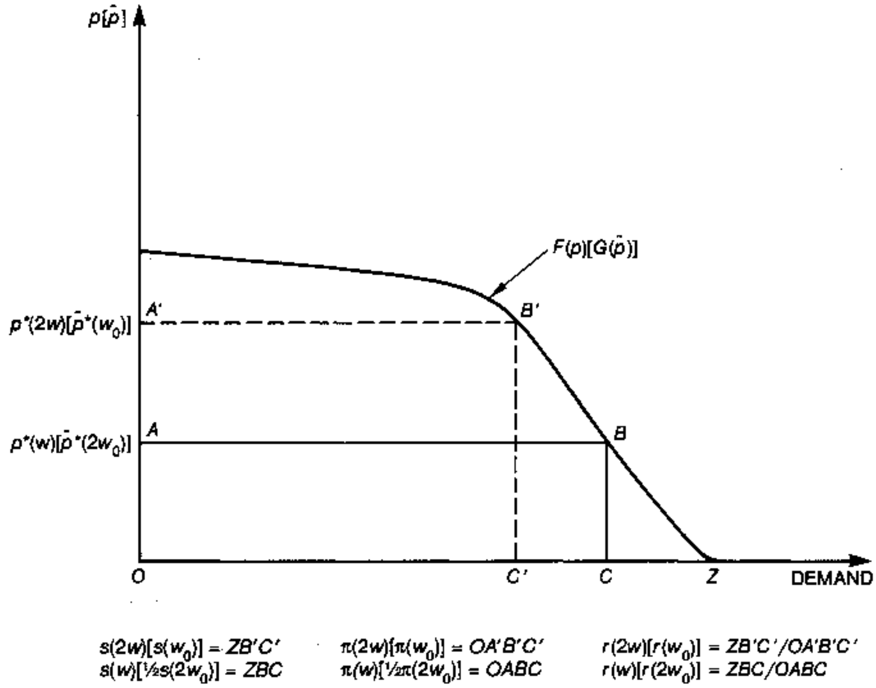
hand $p^*(w) \to 0$, so also $\rho \to 0$, $F(p^*(w)) \to F(0) = 1$, and $F(\rho) \to F(0) = 1$, so the integrands of both $\left(\frac{s_n(w)}{\pi(w)}\right)$ and $\left(\frac{s_{t2}(w)}{\pi(w)}\right)$ vanish (provided F' not $\to -\infty$). The only losses that matter are the actual travelling costs of the travellers, and the ratio of these losses to profits, $\left(\frac{s_{t1}(w)}{\pi(w)}\right)$, depends only on $G(\cdot)$.

[22] Condition (i) is also satisfied by $F(p) = A(p + c)^{\alpha}$, $\alpha < -1$, $A > 0$, $c > 0$; this form of demand, together with our assumption of zero marginal costs, is equivalent to assuming constant elasticity demand, $F(p) = Ap^{\alpha}$, and marginal costs $c > 0$.

[23] Proposition 2 corresponds very closely to Gilbert and Shapiro's (1990) result (their Propositions 2 and 3) that in a homogeneous-product market infinite patent lives and tight price controls are generally optimal. (See also Tandon (1982).) The reason is that in a homogeneous-product market, there is no possibility of consumers substituting related varieties of the product, so only the second of the two kinds of welfare losses that I identified in the introduction, $s_n(w)$, arises. Correspondingly, I have shown here that when consumers have the same transport costs, only this kind of welfare losses arises even in a differentiated-products market. In this special case of my model, therefore, the conditions on $F(\cdot)$ under which infinitely lived patents are optimal (because a narrow patent endogenously forces the patentholder to choose low prices) are exactly the conditions on demand under which Gilbert and Shapiro find that infinitely lived patents (together with exogenously imposed low prices) are optimal. That is, Section 3 of Gilbert and Shapiro's article corresponds closely to this subsection of Section 5 of mine. Of course, in a differentiated-products market, this is only a special case. We shall see in the next subsection that when the first kind of welfare losses are important, infinitely lived patents are often the worst policy.

FIGURE 1

PROFITS AND SOCIAL COSTS WHEN CONSUMERS HAVE IDENTICAL TRANSPORT COSTS [RESERVATION PRICES]



of width w and $2w$ for which prices are set such that consumers just do not want to travel, so $p^*(w) = kw$ and $p^*(2w) = 2kw$ if all consumers have transport costs equal to k. Thus, $r(w) = (ZBC/OABC)$, $r(2w) = (ZB'C'/OA'B'C')$, and $r(w) < r(2w)$ as drawn, since demand is not too convex. (The annotations in brackets correspond to Propositions 3 and 4 and will be discussed below.)

□ **Conditions under which short-lived patents that are as wide as possible are optimal.** Propositions 3 and 4 stand in direct contrast to Propositions 1 and 2.

Proposition 3. If all consumers have identical reservation prices for the most-preferred product variety, then patents of infinite width are most efficient.

The intuition is straightforward. An infinite patent width precludes all competition, so the patentholder charges consumers' common reservation price, yielding profits without either travelling costs or other deadweight losses being incurred.

The conclusion is, of course, unaltered if the product class is of finite radius. (For example, the International Tennis Federation's 1989 rules prohibit the use of a racket face exceeding 178.25 square inches.) In this case, a patent of the maximum possible width precludes any competition and so, is optimal.

A more important constraint on the problem may be that there is a shortest feasible patent lifetime. For example, $L \geq \underline{L}$ may be required to ensure that firms prefer to patent innovations rather than try to keep the details secret. If the reward that we require for the patentholder is large enough, then even a patent that precludes all competition requires a longer lifetime than $\underline{L}$, so the conclusion is again unaltered: patents of maximum width remain optimal. For

$$V < \int_0^{\underline{L}} \pi(\infty) e^{-iT} dT,$$

however (or equivalently, $\underline{L} > \frac{1}{i} \ln (\pi(\infty)/(\pi(\infty) - iV))$), the constraint $L \geq \underline{L}$ would imply $w \leq \pi^{-1}(iV/(1 - e^{-i\underline{L}})) < \infty$. Nevertheless, the point of Proposition 3 remains fairly robust even in this case.

Proposition 4. If all consumers have identical reservation prices for the most-preferred product variety and if there is a shortest allowable patent lifetime $\underline{L} > \frac{1}{i} \ln (\pi(\infty)/(\pi(\infty) - iV))$, then this patent lifetime is optimal (with maximum possible width

$$w = \pi^{-1}(iV/(1 - e^{-i\underline{L}}))) \qquad \text{if}$$

(i) the elasticity of $G(t)$, that is, $(-tG'(t)/G(t))$, is nondecreasing in t, or
(ii) the value of the patent to be awarded, V, is sufficiently large.

Remark 2. All concave, linear, and not overly convex distribution functions $G(\cdot)$ satisfy condition (i). The optimality of the shortest feasible patent lifetimes when consumers have identical reservation prices is therefore fairly general.

The intuition for Proposition 4 is that when all consumers have identical reservation price demands, no deadweight losses are incurred on consumers who purchase from the patentholder. The only social costs are the transport costs of consumers substituting the product at the patent boundary for their preferred product variety. Since the price the patentholder can charge is limited to the consumers' reservation price, fewer consumers incur transport costs as the patent becomes wider. This both increases profits and, in general, reduces total transport costs,[24] so the widest feasible patent is the most socially cost-effective way of awarding profits to an innovator.

Clearly Propositions 1 and 3 are duals of each other, as are Propositions 2 and 4. (See the structures of the proofs in the Appendix.) To see why, consider Proposition 4 for patents of width w_0 and $2w_0$ for which consumers' reservation price, say k, is the binding constraint on the patentholder's price, so $p^*(w_0) = p^*(2w_0) = k$. Social costs and profits could then be measured in the usual way at price k on the demand curves $G\left(\frac{p}{w_0}\right)$ and $G\left(\frac{p}{2w_0}\right)$ that represent consumers' willingness to pay to avoid transport costs when the patent widths are w_0 and $2w_0$, respectively. However, the demand $G\left(\frac{p}{2w_0}\right)$ is just the demand $G\left(\frac{p}{w_0}\right)$ scaled vertically by a factor of two. Thus, I can equivalently compute the ratio of social costs to profits for a patent of width $2w_0$ by halving the price on the original demand curve, $G\left(\frac{p}{w_0}\right)$. I have done exactly this in the brackets in Figure 1 by putting the *relative* price, $\tilde{p} = \frac{p}{w}$, on the vertical axis so that $G(\tilde{p})$ is unaffected by changing w, while $\tilde{p}^*(2w_0) = \frac{k}{2w_0}$ is one-half of $\tilde{p}^*(w_0) = \frac{k}{w_0}$. Here,

$$r(w_0) = (ZB'C'/OA'B'C') > r(2w_0) = (ZBC/OABC)$$

with the areas computed on $G(\tilde{p})$.

In Propositions 1 and 2 (identical transport costs) increasing w loosens the constraint on price, so $p^*(w)$ rises up $F(p)$, generally increasing the ratio of social costs to profits. In

[24] Total transport costs need not be reduced as fewer consumers incur them, since the distribution across consumers of incurred transport costs changes as w changes, even when only consumers with total per-unit transport costs less than consumers' common reservation price, k, are considered. This is why $r(w)$ is not necessarily decreasing in w everywhere and why I need some additional conditions ((i) or (ii) of Proposition 4) to establish the optimality of maximum width patents.

Propositions 3 and 4 (identical reservation prices), on the other hand, increasing w tightens the constraint on relative price, so $\tilde{p}^*(w)$ is pushed down $G(\tilde{p})$, generally reducing the ratio of social costs to profits. The conditions on $G(\cdot)$ that imply that the widest possible patents are optimal when reservation prices are identical are therefore precisely the conditions on $F(\cdot)$ that imply that the narrowest possible patents are optimal when transport costs are identical.

□ Special cases in which all patent widths are equally efficient.

Proposition 5. If (i) all consumers have infinite transport costs, or

(ii) all consumers have infinite reservation prices (that is, perfectly inelastic demand) for the most-preferred product variety, or

(iii) all consumers have both identical transport costs and identical reservation prices for the most-preferred product variety,

then all patent widths are equally efficient.

With infinite transport costs (condition (i)), the patentholder sets the monopoly price, p^m, on $F(\cdot)$ for all patent widths (since no consumers will ever travel), so the ratio of social costs to profits is always $\left[\frac{\int_0^{p^m} F(\rho)d\rho}{p^m F(p^m)}\right] - 1.$[25] In the dual, with infinite reservation prices (condition (ii)), the patentholder sets the relative price, $\frac{p(w)}{w}$, at the monopoly price, $\tilde{p}^m$, on $G(\cdot)$ for all patent widths (since doubling the width doubles the amount every consumer is prepared to pay rather than travel and so, doubles the optimal price), so the ratio of social costs to profits is always $\left[\frac{\int_0^{\tilde{p}^m} G(\rho)d\rho}{\tilde{p}^m G(\tilde{p}^m)}\right] - 1$. Under condition (iii), there are no social costs at any patent width.

□ The general case. This subsection shows that corner solutions are very likely in the general case, as well as under the conditions presented earlier in Section 5. Which corner is optimal depends on the relative elasticities of $F(\cdot)$ and $G(\cdot)$.

Observe that if $H_w(\cdot)$ were made everywhere more convex below its monopoly point without altering the monopoly point, this would increase the area of the deadweight-loss triangle, $s(w)$, under $H_w(\cdot)$, without affecting the profit rectangle, $\pi(w)$. That is, for a given monopoly price, greater convexity of $H_w(\cdot)$ raises $r(w)$. Now, under mild conditions on $F(\cdot)$ and $G(\cdot)$ (see the Appendix), the convexity of $H_w(\cdot)$ (as measured by $(-pH''_w(\cdot)/H'_w(\cdot))$) is that of $G(\cdot)$ at $w = 0$, is that of $F(\cdot)$ at $w = \infty$, and exceeds the minimum of these convexities elsewhere. Therefore, I conjecture that in general, $r(w)$ will be minimized at either $w = 0$ or $w = \infty$.

Recall from Lemma 2 that $r(0)$ equals the monopoly social-cost-to-profit ratio on $G(\cdot)$, and from Lemma 1 that $r(\infty)$ equals the monopoly social-cost-to-profit ratio on $F(\cdot)$. Thus, if $G(\cdot)$ is less (more) elastic than $F(\cdot)$ (this is a weaker condition than $G(\cdot)$ being less (more) convex than $F(\cdot)$) below their respective monopoly prices, then $r(0)$ is less (greater) than $r(\infty)$, and I expect that $r(w)$ is minimized at $w = 0$ ($w = \infty$).

As an illustration of this intuition, consider the case of linear demand with a uniform distribution of transport costs from zero to an upper bound—perhaps the most natural example in which consumers differ in both demands and transport costs. Here, both $G(\cdot)$ and $F(\cdot)$ are linear, so since deadweight losses are always half of the profits on linear

---

[25] Of course, the assumption of an infinite transport cost means that we cannot really think of this as representing a differentiated-products market.

demand, both $r(0) = .5$ and $r(\infty) = .5$. However, for all $0 < w < \infty$, $H_w(\cdot) = F(\cdot)G(\cdot)$ is strictly convex, so $r(w) > .5$.[26]

Appendix B gives further details but also shows that it is possible to construct an example with an interior optimum.

Summarizing, interior optima are possible, but I expect that the narrowest (widest) possible patents are typically optimal when $G(\cdot)$ is less (more) elastic than $F(\cdot)$ below their respective monopoly prices.[27]

6. Extensions

■ My model seems a simple one in which to address the trade-off between the social costs of consumption switched to the wrong product variety and those of consumption switched out of the product class. It also clearly exposes a duality between the role of consumers' demand for their preferred product and the role of their transport costs. I show in this section that my main results are robust to several extensions that may make the model a more natural one.

□ **Multiple patented brands.** By assuming free entry without fixed costs subject to non-infringement of the patent, I avoided solving an asymmetric differentiated-products oligopoly problem and computing the number of fixed costs incurred by entrants.[28] If entrants have no fixed costs, however, it is natural to allow the patentholder to sell a range of product varieties at no additional cost.[29] It will often wish to do this in order to price discriminate between consumers with different transport costs. Whereas in my main analysis, consumers either buy the most-preferred variety (do not travel at all) or travel the entire distance to the patent boundary, consumers may now travel intermediate distances, with lower-transport-cost consumers travelling greater distances to obtain lower prices. For the varieties it produces, the patentholder will set a price schedule that is decreasing and convex in distance from the preferred brand. (The convexity follows from the linearity of transport costs, since each consumer would choose either the next further-out or the next closer-in variety in preference to any variety on a concave part of the price schedule, so such a variety would have no purchasers.) It seems very hard to analyze how the patentholder's optimal price schedule, and hence the ratio of social costs to profits, varies with the patent width. Furthermore, since consumers differ in both their demands and their transport costs, but the patentholder can discriminate only on the latter, the duality between the roles of these attributes is no longer perfect. Nevertheless, the main results of Section 5 still apply.

---

[26] The intuition developed in this subsection applies to Propositions 1, 2, 3, and 4 but not to the very special cases of conditions (i) and (ii) in Proposition 5 (for the same reason that Lemmas 1 and 2 do not apply in those cases—the patentholder's price (relative price) remains finite even as $w \to 0$ ($w \to \infty$) if transport costs (reservation prices) are infinite).

[27] Interior optima might be more likely if the model incorporated risk aversion and uncertainty. For example, if agents were risk averse, uncertain lead times between being granted patents and being able to exploit them would favor longer patent lives, but less certain market sizes in the more distant future would favor shorter lives. Waterson (1988) and Schmitz (1989) develop models that emphasize legal uncertainty about the extent of protection a patent provides.

[28] I conjecture that incorporating fixed costs into the model would make narrower patents relatively less attractive both because they allow more competitors than wide patents and because even narrow patents may now permit nontrivial prices.

[29] If (perhaps because of the previously sunk costs of establishing a brand identity) firms sell one brand only, two price-competing entrants without the patent are clearly sufficient for all of my analysis to apply. In fact, my main results hold even with only a single competitor to the patentholder. Propositions 1 and 2 hold because with identical transport costs, the competitor sets a price of zero but attracts no sales in any equilibrium. Proposition 3 holds because the patentholder's profits (weakly) exceed $kG(k/w)$ (setting the price at consumers' common reservation price, k, guarantees this) and because total social costs plus profits cannot exceed k. Proposition 5 applies because

When consumers have identical transport costs, the patentholder has no ability to price discriminate between them (all consumers will choose the same product variety), so the patentholder sells only the most-preferred variety. Thus, Propositions 1 and 2 apply exactly as before.

When all consumers have reservation prices of k, selling only the most-preferred variety at price k yields total profits of $kG(k/w)$. So the patentholder's optimal price schedule yields profits of at least $kG(k/w)$ and, because total social benefits with the patent must be at least as large as profits, these too must be at least $kG(k/w)$. Since total social benefits without the patent are k, total social costs relative to competitive provision of the product are no greater than $k - kG(k/w)$. Thus, $r(w) = \dfrac{s(w)}{\pi(w)} \leq \dfrac{k - kG(k/w)}{kG(k/w)} \to 0$ as $w \to \infty$, so an infinitely wide patent is optimal, and Proposition 3 applies. It follows, as before, that even with a minimum feasible patent lifetime $\underline{L}$, if the required reward is large enough ($V \geq \int_0^{\underline{L}} \pi(\infty)e^{-iT}dT$), then a patent of the maximum width is optimal.

Thus, as in the main model, the widest (narrowest) possible patents are generally optimal when consumers have identical reservation prices (transport costs).[30]

□ **Consumers with different most-preferred varieties.** Another natural extension of the model is to assume that different consumers have different most-preferred varieties within the product class. I can model this by assuming that for each interval of product varieties $[x, x + dx]$ along a "product line," there are consumers "located at x" whose most-preferred varieties lie in this interval and whose aggregate demand per unit time would be $F(p)\,dx$ if no other product varieties were available. I assume, as above, that the patentholder can sell the full range of varieties.[31] For expositional simplicity, I assume here that there is free entry at just one end of a line segment of length w. (It would make no important difference to the results to allow free entry at both ends of the line or to generalize to a multidimensional product space.) Again, the main results of Section 5 apply.

Write $s(x)$, $\pi(x)$, and $r(x) = \dfrac{s(x)}{\pi(x)}$, as before, for the social costs, profits, and ratio of social costs to profits generated by consumers whose most-preferred variety is at a distance x from the patent boundary. Then, total social costs and patentholder's profits are $S(w) = \int_0^w s(x)dx$ and $\Pi(w) = \int_0^w \pi(x)dx$, and I now wish to minimize $R(w) = \dfrac{S(w)}{\Pi(w)}$. Note that if $r(x)$ is monotonic increasing (decreasing), then $R(w)$ is also monotonic increasing (decreasing).

To see that Propositions 1 and 2 apply, recall from the previous subsection that when all consumers located at x have the same transport cost, k, then they all choose the same product variety.[32] Therefore, any price schedule that maximizes profits from (only) the

then prices, profits, and social costs are all linear in w. (With quantity competition and a single competitor to the patentholder, Proposition 3 always holds—because setting quantity at $G(k/w)$ guarantees the patentholder profits of $kG(k/w)$—and Proposition 1 holds for a wide class of demand curves, $F(p)$, including linear demand.)

[30] Proposition 5 of the main model also applies. (With infinite reservation prices, the patentholder sells only the preferred variety for any w—this result does not generalize to identical, finite reservation prices—and sets the relative price at $\hat{p}^m$. However, I do not know whether Proposition 4 applies, since $F(\cdot)$ and $G(\cdot)$ are no longer duals of one another.

[31] The assumption that the patentholder can sell the full range of varieties is particularly natural when consumers have different preferred goods, and it here yields a model much more closely related to the main model than that which the alternative assumption yields. In Waterson (1988), consumers prefer different goods, but the patentholder produces only a single brand. Waterson's model emphasizes uncertainty about the extent of protection a patent provides but ignores the effects of patent width on prices. These effects are crucial in my models.

[32] Of course, the minimum feasible width is now $\Pi^{-1}(iV)$ rather than $\pi^{-1}(iV)$ in the statement of Proposition 2.

consumers located at x results in all of these consumers buying at x at the price $p^*(x)$ that would be optimal if, as in the main model, they could buy only at either x or the patent boundary. Now if the elasticity of $F(p)$ is nondecreasing, then $pF(p)$ is quasi-concave, so $p^*(x) = \min\{p^m, kx\}$, where p^m is the monopoly price on $F(\cdot)$ and kx is the total per-unit transport cost of travelling from x to the boundary.[33] Since consumers located at x with transport costs k facing the price schedule $p^*(x) = \min\{p^m, kx\}$, do in fact all buy at x even when they are free to choose any variety, this price schedule is the patentholder's optimal price schedule in the current model. It follows that the function $r(\cdot)$ is exactly that of the proof of Proposition 2 of the main model so $r'(x) \geq 0$. Therefore, $R'(w) \geq 0$, which demonstrates condition (i) of Proposition 2 for the current version of the model. Similarly, for any $F(\cdot)$, $pF(p)$ is increasing for sufficiently small p, so for sufficiently small w the patentholder's optimal price schedule is $p^*(x) = kx$. (As above, this would be optimal at each x if consumers located at x were unable to buy at intermediate points and consumers located at x who face this price schedule do in fact all buy at x.) Again, then, $r(\cdot)$ is the same function as that of the analysis of the main model, so $r(0) = 0$ and $r'(x) \geq 0$ for sufficiently small x. Hence, $R(0) = 0$ and $R'(w) \geq 0$ for sufficiently small w, which establishes Proposition 1 and condition (ii) of Proposition 2 for the current version of the model by arguments paralleling those for the main model.

To check that Proposition 3 applies, observe that if all consumers have reservation prices of k, then $\Pi(w) \geq \int_0^w kG(k/x)dx$, since selling all varieties at price k would yield the patentholder exactly this profit. Also, since total social benefits with the patent are at least $\Pi(w)$ and social benefits without the patent are kw, $S(w) \leq kw - \Pi(w)$. So, $R(w) = \dfrac{S(w)}{\Pi(w)} \leq \dfrac{kw - \int_0^w kG(k/x)dx}{\int_0^w kG(k/x)dx} \rightarrow 0$ as $w \rightarrow \infty$, using l'Hôpital's rule, so Proposition 3 holds for this version of the model. Of course, it is now natural for the product class to be of finite radius $\bar{w}$ or for all consumers' most-preferred varieties to lie in the range $[0, \bar{w}]$. (Otherwise, profits are unbounded as $w \rightarrow \infty$.) In either case, a patent of the maximum possible width is optimal (the patentholder charges k for all varieties) and is, of course, also optimal when there is a minimum feasible patent lifetime $\underline{L}$, provided that the required reward, V, is large enough.

Again, therefore, the widest (narrowest) possible patents are generally optimal when consumers have identical reservation prices (transport costs).[34]

□ **Alternative specifications of transport costs.** The working paper version of this article shows that my main conclusions are robust to generalizing the model to allow transport costs (i) to be correlated with demands for the most-preferred variety, (ii) to depend on distance in different, nonlinear ways for different consumers, or (iii) to be a fixed cost independent of the quantity purchased, rather than a per-unit cost, for each consumer.

7. Conclusion

■ This article has addressed the question of optimal patent shape by asking what form of patent yields patentholder's profits with least social costs. My model has separated out two kinds of social costs: those of consumers not purchasing any member of the product class and those of consumers inefficiently substituting a less-preferred member of the class. A wider patent reduces the distortion of consumers' choices between the patented brand of the product and unpatented, lower-priced varieties of the product sold by competitors.

[33] As above, I assume that indifferent consumers buy their ideal variety, but this assumption makes no difference to the results.

[34] Again, Proposition 5 of the main model applies. (With infinite reservation prices, the patentholder's optimal price schedule is $p^*(x) = x\hat{p}^m$, since this schedule maximizes profits from consumers at each x; see footnote 30.) Again, I do not know whether Proposition 4 applies.

However, a wider patent also permits higher prices and so increases (relative to profits) the deadweight losses from consumers switching consumption out of the product class. Thus, if demand is relatively more elastic in reservation price than in substitution cost, society should generally be more concerned with nonconsumption and so should narrow patent scope to ensure low prices, while if the converse is true, society should generally be more concerned with substitution within the product class and hence should broaden patents' scope.

In particular, infinitely lived, narrow patents are typically desirable when substitution costs between varieties of the product are similar across consumers, but very short-lived, wide patents are desirable when valuations of the preferred variety relative to not buying the product at all are similar across consumers. Thus, for example, if potential consumers have varying levels of need for a computer program (because they would use it with varying frequencies) but have similar strengths of preferences between a program that is easy to learn and harder-to-learn copies, a narrow but very long-lived patent (or copyright) is probably called for. If, on the other hand, potential users have similar needs for a drug to cure a serious disease but alternative possible formulations of the drug produce side effects which are of different severities for different patients, a broader but probably shorter-lived patent is likely to be preferable.

One way to understand these results is to note that if a social planner could directly control the patentholder's price, he would typically choose both an infinite patent lifetime and an infinite patent width together with a price just high enough to allow the patentholder the correct total discounted profits—just as in the standard optimal tax problem in which the widest possible scope of taxation is generally the efficient way to raise a given revenue. In my model, price cannot be controlled directly, but it can be controlled indirectly by narrowing the scope of patent protection, thereby allowing competition to the patentholder's product. The cost of this is that some consumers inefficiently substitute to competitively provided varieties; indirectly controlling price in this way works best when few consumers actually substitute. In the extreme case in which all consumers have the same costs of substituting between varieties, the patentholder sets a price such that no consumers substitute and a narrow patent achieves the same welfare level as if price could be controlled directly. More generally, a narrow patent is desirable when it causes relatively few consumers to substitute, that is, when demand is relatively inelastic in substitution cost. When demand is relatively inelastic in reservation price, on the other hand, controlling price is relatively less important, and so broader patents are optimal.

These results suggest how optimal patent policies vary across different classes of products. However, we must be extremely cautious about drawing policy conclusions from this simple model. In particular, further research should relax the implicit assumption that innovative activity ceases after a single patent is awarded. For example, a wider patent may prevent duplicative R&D designed merely to closely mimic the patentholder's product but, on the other hand, a narrower patent may provide greater incentives to refine and improve the original innovation.[35] Since subsequent innovations may bring in product varieties superior to that originally patented, an additional policy question is how much of an improvement must a new variety be to not infinge the original patent—that is, we need a theory of optimal patent *height* in addition to our theory of optimal width.[36] Furthermore, patent width may affect R&D efforts even before the first innovation, since narrower patents allow the possibility of patents being granted to more than one winner.[37] These issues are left for future research.

[35] Merges and Nelson (1989) address this issue. See also Gallini (1989).

[36] In this case, I would work in vertical rather than horizontal product space and my assumption that a competitive industry has free access to all qualities outside the range protected by the patent would not make much sense. Furthermore, the innovating firm's choice of product(s) will be affected in a space in which more products are patentable. Green and Scotchmer (1989) and Scotchmer and Green (1989) address the issue of how much of an improvement an innovation should be in order to be patentable.

Appendix A.

■ The proofs of Lemmas 1 and 2 and of Propositions 1, 2, 3, 4, and 5 follow.

Proof of Lemma 1. Since $G(\cdot)$ is a decreasing function, $pF(p)G\left(\frac{p}{w}\right) \leq p^m F(p^m)G\left(\frac{p^m}{w}\right)$, $\forall p \geq p^m$, so $p^*(w) \leq p^m$. Therefore, $p^*(w)$ is a continuous function by the theorem of the maximum, and so $\lim_{w\to\infty} p^*(w) = \operatorname{argmax}_p\, (pF(p)G(0)) = p^m$. Now, $\int_0^{p^*(w)} F(\rho)G\left(\frac{\rho}{w}\right)d\rho$, and hence, $r(w)$ is also continuous in w. So, since $\lim_{w\to\infty} G\left(\frac{\rho}{w}\right) = \lim_{w\to\infty} G\left(\frac{p^*(w)}{w}\right) = G(0) = 1, \forall\rho, p^*(w) \leq p^m$, I obtain

$$\lim_{w\to\infty} r(w) = \left[\frac{\int_0^{p^m} F(\rho)d\rho}{p^m F(p^m)}\right] - 1. \quad Q.E.D.$$

Proof of Lemma 2. The proof parallels that of Lemma 1 using the changes of variables $\tilde{p} = p/w$ and $\tilde{p}^*(w) = p^*(w)/w$. So, $\tilde{p}^*(w) = \operatorname{argmax}_{\tilde{p}}\, (\tilde{p}F(w\tilde{p})G(\tilde{p}))$. Since $F(\cdot)$ is decreasing,

$$\tilde{p}F(w\tilde{p})G(\tilde{p}) \leq \tilde{p}^m F(w\tilde{p}^m)G(\tilde{p}^m), \forall\tilde{p} \geq \tilde{p}^m.$$

So, $\tilde{p}^*(w) \leq \tilde{p}^m$, so $\tilde{p}^*(w)$ is continuous in w, and so $\lim_{w\to 0} \tilde{p}^*(w) = \operatorname{argmax}_{\tilde{p}}\, (\tilde{p}F(0)G(\tilde{p})) = \tilde{p}^m$. Now,

$$r(w) = \left[\frac{\int_0^{\tilde{p}^*(w)} F(w\tilde{\rho})G(\tilde{\rho})d\tilde{\rho}}{\tilde{p}^*(w)F(w\tilde{p}^*(w))G(\tilde{p}^*(w))}\right] - 1 \text{ (using the change of variables } \tilde{\rho} = \rho/w\text{);}$$

so since $\lim_{w\to 0} F(w\tilde{\rho}) = \lim_{w\to 0} F(w\tilde{p}^*(w)) = F(0) = 1, \forall\tilde{\rho}, \tilde{p}^*(w) \leq \tilde{p}^m$, I obtain

$$\lim_{w\to 0} r(w) = \left[\frac{\int_0^{\tilde{p}^m} G(\tilde{\rho})d\tilde{\rho}}{\tilde{p}^m G(\tilde{p}^m)}\right] - 1. \quad Q.E.D.$$

Proof of Proposition 1. Let consumers' transport costs be $k < \infty$ per unit per unit distance, so $G(t) = 0$ for $t > k$ and $G(t) = 1$ for $t \leq k$. ($g(t)$ has a mass point at k.) Thus, the monopoly price on $G(\cdot)$ is k; so, using Lemma 2, $r(0) = 0$. (The case $k = \infty$ is covered by Proposition 5.) *Q.E.D.*

Proof of Proposition 2. If all consumers have transport costs $k < \infty$, the patentholder's sales are $F(p)$ if $p \leq kw$, and zero otherwise. Therefore, $p^*(w)$ is monotonic weakly increasing, and since also $p^*(w) \leq kw$, no consumers travel. So, $s_{t1}(w) = s_{t2}(w) = 0$ and so $r(w) = s_n(w)/\pi(w) = (\int_0^{p^*(w)} F(\rho)d\rho/p^*(w)F(p^*(w))) - 1$ and (slightly abusing notation) $r'(w) = \frac{dp^*(w)}{dw}\frac{d}{dp^*}\left(\frac{\int_0^{p^*} F(\rho)d\rho}{p^*F(p^*)}\right)$. To show Part (i) is true, note that

$$\operatorname{sign}\left[\frac{d}{dp^*}\left(\frac{\int_0^{p^*} F(\rho)d\rho}{p^*F(p^*)}\right)\right] = \operatorname{sign}\left[\frac{S(p^*)}{p^{*2}F(p^*)}\right],$$

in which $S(p^*) = \left[p^*F(p^*) - (\int_0^{p^*} F(\rho)d\rho)\left(\frac{p^*F'(p^*)}{F(p^*)} + 1\right)\right]$. Now, $S(0) = 0$ and

$$\frac{dS}{dp^*} = \left(\int_0^{p^*} F(\rho)d\rho\right)\left(\frac{d}{dp^*}\left(\frac{-p^*F'(p^*)}{F(p^*)}\right)\right) \geq 0$$

---

[37] With narrow patents, the original innovator may wish to take out multiple patents. In practice, this may be costly and not improve protection much in a multidimensional product space—my model is not restricted to a single dimension. (For example, Whitney could not have patented all possible trivial modifications to his cotton gin design, nor can a software firm copyright all programs that differ from an original program in only one function.) Note, however, that there may be a socially beneficial role for sleeping patents that widen effective protection without requiring the patentholder to actually produce all varieties. On the other hand, Klemperer (1988) and the fuller analysis of La Manna, MacLeod, and de Meza (1989) show that awarding patents for equivalent technologies to multiple entrants (but lengthening patent lives) generally lowers the social costs of providing a given total reward to innovation. The reason is that, as I have shown in this article, a lower equilibrium price on a given demand curve typically lowers the ratio of social costs to profits.

if the elasticity of $F(p)$, i.e., $\left(\frac{-pF'(p)}{F(p)}\right)$, is nondecreasing, so $S(\cdot) \geq 0$. Therefore, $r'(w) \geq 0$, so the minimum feasible width ($w = \pi^{-1}(iV)$) is optimal if (i) holds. Furthermore, since $r(0) = 0$ ($p^*(0) = 0$) and $r(w) \geq 0$, I can find $\epsilon > 0$ such that $r'(w) \geq 0$ for $w \in [0, \epsilon]$. (My assumption that $F(\cdot)$ is real-analytic ensures that $r'(w)$ does not change sign infinitely often in every neighbourhood of zero.) For all $w > \epsilon$,

$$r(w) \geq \delta = s(\epsilon)/\max_p (pF(p)),$$

since $s(w)$ is increasing because $p^*(w)$ is increasing. (I assume as usual that $\max_p (pF(p))$ exists). Now, I can choose $\alpha > 0$ such that $r(\pi^{-1}(\alpha)) \leq \delta$ (and so also $\pi^{-1}(\alpha) \leq \epsilon$). So, for $V < \alpha/i$, $\pi^{-1}(iV) < \pi^{-1}(\alpha)$, and I have $r(\pi^{-1}(iV)) \leq r(w)$ for all $w > \pi^{-1}(iV)$. Therefore, the minimum feasible width ($w = \pi^{-1}(iV)$) is optimal if $V < \alpha/i$, which demonstrates Part (ii) of the proposition. *Q.E.D.*

Proof of Proposition 3. Let consumer's reservation prices be $k < \infty$ per unit, so $F(p) = 0$ for $p > k$ and $F(p) = 1$ for $p \leq k$. Thus, the monopoly price on $F(\cdot)$ is k, so using Lemma 1, $r(\infty) = 0$. (The case $k = \infty$ is covered by Proposition 5.) *Q.E.D.*

Proof of Proposition 4. If all consumers have reservation prices $k < \infty$, the patentholder's sales are $G\left(\frac{p}{w}\right)$ if $p \leq k$ and zero otherwise. Write $\tilde{p} = \frac{p}{w}$. Then, $\tilde{p}^*(w) = \frac{p^*(w)}{w}$ maximizes $\tilde{p}G(\tilde{p})$ subject to $\tilde{p} \leq \frac{k}{w}$. Therefore, $\tilde{p}^*(w)$ is monotonic weakly decreasing, and (since $p^*(w) \leq k$) consumers purchase the same total number of units of the product class that they would with competitive provision. Therefore, the only social costs are transport costs: $s_\pi(w) = s_\Omega(w) = 0$, so $r(w) = s_t(w)/\pi(w) = (\int_0^{\tilde{p}^*(w)} G(\hat{p})d\hat{p}/\tilde{p}^*(w)G(\tilde{p}^*(w))) - 1$. An argument parallel to that used for Part (i) of Proposition 2 now shows that the maximum feasible width is optimal if condition (i) holds. The proof of Part (ii) is directly analogous to that of Part (ii) of Proposition 2 using the facts that $r(\infty) = 0$ ($\tilde{p}^*(\infty) = 0$) and $r'(w) \leq 0$ for w large enough, hence $\tilde{p}^*$ small enough $\left(\tilde{p}^* \leq \frac{k}{w}\right)$, since $\tilde{p}^*(w)$ is decreasing. *Q.E.D.*

Proof of Proposition 5. If condition (i) holds, $H_w(p) = F(p)\ \forall w$, and if condition (ii) holds, $H_w(p) = G(\tilde{p})\ \forall w$; so, in either case, $r(w)$ is constant. (In case (i) ((ii)), Lemma 2 (Lemma 1) does not apply, since its conditions are not satisfied.) If condition (iii) holds and consumers have transport costs k_t and reservation prices k_r, $H_w(p) = 1$ if $p \leq \min\{k_r, wk_t\}$ and $H_w(p) = 0$, otherwise; so, $r(w) = 0\ \forall w$. *Q.E.D.*

Appendix B

■ Here, I give the details for the final subsection of Section 5.

If $\phi(\cdot)$ is more convex than $\psi(\cdot)$, that is, $\left(\frac{-p\phi''(p)}{\phi'(p)}\right) \geq \left(\frac{-p\psi''(p)}{\psi'(p)}\right)$, and the monopoly price on both $\phi(\cdot)$ and $\psi(\cdot)$ is p^m $\left(\text{so, } \frac{-p^m\phi'(p^m)}{\phi(p^m)} = \frac{-p^m\psi'(p^m)}{\psi(p^m)} = 1\right)$, then $\frac{\phi(p)}{\phi(p^m)} \geq \frac{\psi(p)}{\psi(p^m)}$ for all $p < p^m$; so,

$$\left[\int_0^{p^m} \phi(\rho)d\rho \Big/ p^m\phi(p^m)\right] - 1 \geq \left[\int_0^{p^m} \psi(\rho)d\rho \Big/ p^m\psi(p^m)\right] - 1.$$

Thus, for a given monopoly price, greater convexity raises $r(w)$. Now,

$$\left(\frac{-pH''_w}{H'_w}\right) = \frac{-p\left(F''G + \frac{2F'G'}{w} + \frac{FG''}{w^2}\right)}{\left(F'G + \frac{FG'}{w}\right)} = \lambda\left(\frac{-pF''}{F'}\right) + (1-\lambda)\left(\frac{-\tilde{p}G''}{G'}\right) + \mu,$$

where

$$\tilde{p} = \frac{p}{w}, \qquad \lambda = \left(\frac{wF'G}{wF'G + FG'}\right) \in [0, 1], \qquad \mu = \left(\frac{-2pF'G'}{wF'G + FG'}\right) \geq 0,$$

so $\left(\frac{-pH''_w}{H'_w}\right) \geq \min\left\{\left(\frac{-pF''}{F'}\right), \left(\frac{-\tilde{p}G''}{G'}\right)\right\}$. Also, as $w \to \infty$, I have $\lambda \to 1$, $\mu \to 0$ (provided that $\tilde{p} \to 0$, F' not $\to 0$, and G' not $\to -\infty$), so $\left(\frac{-pH''_w}{H'_w}\right) \to \left(\frac{-pF''}{F'}\right)$ (see Lemma 1), while as $w \to 0$, I have $\lambda \to 0$ and $\mu \to 0$ (provided that $p \to 0$, G' not $\to 0$, and F' not $\to -\infty$), so $\left(\frac{-pH''_w}{H'_w}\right) \to \left(\frac{-\tilde{p}G''}{G'}\right)$. (See Lemma 2.)

Example 1: Linear demand, $F(p) = 1 - \beta p$, and uniform density of transport costs on $[0, \gamma]$. Both $F(p)$ and $G(t) = \left(1 - \frac{t}{\gamma}\right)$, $t \leq \gamma$, are linear; so, $r(\infty) = .5$ by Lemma 1, and $r(0) = .5$ by Lemma 2. For all $0 < w < \infty$, $H_w(p) = (1 - \beta p)\left(1 - \frac{p}{w\gamma}\right)$ is a strictly convex function, so $r(w) > .5$. Without loss of generality, I can set $\beta = \gamma = 1$ and compute $r(w) = \frac{p^*(w)(3w + 3 - 4p^*(w))}{6(1 - p^*(w))(w - p^*(w))}$, where $p^*(w) = \frac{(1 + w) - \sqrt{(1 + w)^2 - 3w}}{3}$. The maximum value of $r(w)$ is $7/12$ (at $w = 1$, at which $p^*(w) = 1/3$ and the two kinds of social costs are equal: $s_{t1}(1) = s_{t t}(1) + s_{t2}(1) = 7/54$); so, in this example, the worst patent width yields social costs only one-sixth greater than the socially optimal widths ($w = 0$ and $w = \infty$).

Example 2: An example with an interior optimum. The intuition developed above is that the convexity of $H_w(\cdot)$ always exceeds the minimum of those of $F(\cdot)$ and $G(\cdot)$; so I expect a corner optimum. An important caveat is that the monopoly price on $H_w(\cdot)$, and hence the range of $H_w(\cdot)$ in which we are interested, varies with w. Since the monopoly price on $H_w(\cdot) = F(\cdot)G(\cdot)$ is lower than the monopoly price on either $F(\cdot)$ or $G(\cdot)$ (both functions are decreasing), I may obtain an interior optimum if $F(\cdot)$ and $G(\cdot)$ are (below their monopoly prices) sufficiently more concave at lower prices than at higher prices.

An extreme example has $F(\cdot)$ and $G(\cdot)$ of constant elasticity arbitrarily close to but slightly less in magnitude than -1, truncated at both high prices and at a quantity of unity. Then the monopoly price on either $F(\cdot)$ or $G(\cdot)$ is at the arbitrarily high price at which the curve is truncated, so $r(0)$ and $r(\infty)$ are arbitrarily high. However, the monopoly price on $H_1(\cdot) = F(\cdot)G(\cdot)$ is at one, below which $H_1(\cdot)$ is vertical, so $r(1) = 0$. Therefore, the optimal patent width is the interior width, $w = 1$. Although $H_1(\cdot)$ is more convex than either $F(\cdot)$ or $G(\cdot)$, $H_1(\cdot)$ is nevertheless less convex below its low monopoly price than are $F(\cdot)$ and $G(\cdot)$ on average below their high monopoly prices. (Formal details of this example are in the working paper version of this article.)

References

ALLEGRETTI, D.D. AND SHERRY, S.F. "The Doctrine of Equivalents: the Evolution Continues." Working Paper, Allegretti, Newitt, Witcoff, and McAndrews, Ltd., 1988.

GALLINI, N.T. "Patent Policy and the Disclosure Decision." Working Paper, University of Toronto, 1989.

GILBERT, R. AND SHAPIRO, C. "Optimal Patent Length and Breadth." *The RAND Journal of Economics*, Vol. 21 (1990), pp. 106–112.

GREEN, J. AND SCOTCHMER, S. "Technological Licensing and the Novelty Requirement in Patent Law." Working Paper, Harvard University and University of California, Berkeley, 1989.

KLEMPERER, P.D. "How Broad Should the Scope of Patent Protection Be?" Working Paper, Oxford University, 1988.

LA MANNA, M., MACLEOD, R., AND DE MEZA, D. "The Case for Permissive Patents." *European Economic Review*, Vol. 33 (1989), pp. 1427–1443.

MCQUADE, W. "Prince Triumphant." *Fortune* (February 22, 1982), pp. 84–90.

MELLOAN, G. "An American Views Japan's Copycat Culture." *Wall Street Journal* (July 12, 1988), pp. 29.

MENELL, P.S. "Tailoring Legal Protection for Computer Software." *Stanford Law Review*, Vol. 39 (1987), pp. 1329–1372.

MERGES, R.P. AND NELSON, R. "On The Complex Economics of Patent Scope." Working Paper, Boston University and Columbia University, 1989.

NORDHAUS, W. "Invention Growth and Welfare." Cambridge, Mass.: MIT Press, 1969.

SCHERER, F.M. "Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation." *American Economic Review*, Vol. 62 (1972), pp. 422–427.

SCHMITZ, J.A. "On the Breadth of Patent Protection." Working Paper, SUNY, Stony Brook, 1989.

SCHNEIDER, H.J. "Claims to Fame." Working Paper, Leydig, Voit, and Mayer, Ltd., 1988.

SCOTCHMER, S. AND GREEN, J. "Novelty and Disclosure in Patent Law." Working Paper, University of California, Berkeley and Harvard University, 1989.

——— AND YOSHIKAWA, A. "Protection of Intellectual Property Rights in Biotechnology." Working Paper, University of California, Berkeley, 1987.

TANDON, P. "Optimal Patents with Compulsory Licensing." *Journal of Political Economy*, Vol. 90 (1982), pp. 470–486.

WATERSON, M. "The Economics of Product Patents." Working Paper, University of Reading, 1988.

WIEGNER, K. AND HEINS, J. "Can Las Vegas Sue Atlantic City?" *Forbes* (March 6, 1989), pp. 130–137.