15
Price discrimination and patent policy

Jerry A. Hausman
Massachusetts Institute of Technology and NBER
and
Jeffrey K. MacKie-Mason
University of Michigan and NBER

Chapter 15. *Mathematical Models in Economics*. MOL Bacharach, MAH Dempster, and JL Enos, eds. University of Oxford (1990).

15

Price discrimination and patent policy

Jerry A. Hausman
Massachusetts Institute of Technology
and National Bureau of Economic Research

and

Jeffrey K. MacKie-Mason
University of Michigan
and National Bureau of Economic Research

Abstract

Patent and antitrust policy are often presumed to be in conflict. As an important example, there is ongoing controversy about whether price discrimination by a patent holder is an illegal or socially undesirable exploitation of monopoly power. In this article, we show that no conflict exists in many price discrimination cases. Even ignoring the (dynamic) effects on incentives for innovation, third-degree price discrimination by patent holders can raise (static) social welfare. In fact, Pareto improvements may well occur. Welfare gains occur because price discrimination allows patent holders to open new markets and to achieve economies of scale or learning. Further, even in cases where discrimination incurs static welfare losses, it may be efficient relative to other mechanisms, such as patent life, for rewarding innovators with profits.

1. Introduction

A patent grant establishes a lawful monopoly with exclusive rights to 'make, use, or vend' with the patented good or process during the life of the patent. U.S. courts, however, have viewed the patent grant as a franchise, or privilege 'conditioned by a public purpose' (Neale and Goyder 1980, p.289). Through case law the courts have circumscribed

---

Reprinted with permission from the *Rand Journal of Economics*, Vol. 19, no. 2, Summer 1988.
We thank Oliver Hart, Steve Salop, Dick Schmalensee, Steve Salant, and Jean Tirole for helpful discussions, and three anonymous referees for useful comments. We gratefully acknowledge financial support from the MIT Center for Energy Policy Research.

the ways in which patentees may 'use or vend' their patented innovations. Price discrimination is one practice that has been repeatedly challenged.[1]

Economic analyses of price discrimination emphasize two issues. First, price discrimination raises the patentee's profitability, which is the purpose of the patent grant. Second, price discimination misallocates resources among purchasers and thus causes a decrease in social welfare.[2] Customers facing different prices will have different marginal valuations for the patented good, so that some loss in consumer welfare—relative to a first-best world—is inevitable. On the other hand, increasing the expected reward by allowing price discrimination should increase innovative effort, which presumably benefits society. Given the existing patent laws (which have their genesis in the Constitution), we must presume that monopoly rewards to inventors have some social desirability.

An optimal social policy for patents and monopoly will maximize the net social benefit of encouraging innovation while incurring monopoly misallocations. One can view the case law as a series of attempts to find the boundaries of the optimal trade-off. For a particular behaviour, such as price discrimination, the question is whether the incremental gains from discrimination exceed the allocation costs incurred.[3] This balancing test is a difficult one to implement. There is almost no agreement on how much investment is induced and what the resulting social pay-off is from an increase in expected profit. Further, the courts deal with one case at a time, but an optimal policy should be based on the benefits and costs for all innovative effort.[4]

In this article we show that the difficult balancing of incentives and allocation costs may not be necessary. The trade-off can be viewed

[1] We consider only third-degree discrimination and use 'price discrimination' as an economic term of art, as defined in, say, Pigou (1920). Legal usage may sometimes differ.

[2] Bowman (1973) believes that price discrimination should be encouraged because the net loss in allocative efficiency is small, while the gain in innovation incentives from higher profitability is significant (pp. 56, 112). Sullivan (1977) opposes price discrimination (at least sometimes) because no socially desirable gain is obtained by increasing the patentee's profits *ex post*, while the cost to consumers can be high. Baxter (1966) has opposed patentee price discimination because of inefficiences caused by charging different consumers different prices. See also Kaplow (1984).

[3] This marginalist approach embodies certain assumptions about continuity and concavity of the social welfare function; more generally, the resulting policy should be examined as a whole for its global optimality.

[4] In particular, the *ex post* difference between measuring the benefits of a particular innovation and the monopoly costs created by that patent holder is not the right balance to check. The *ex ante* expected social benefits should be equal to the social costs for the *marginal* innovative effort, not for *each* project.

as one between dynamic welfare benefits (innovation incentives) and static welfare effects. We find, however, that the static effect will in many cases be beneficial, so that no trade-off exists. If the cases in which both static and dynamic effects are beneficial are numerous or important enough, then optimal policy should allow discrimination.

It is well known that discrimination may raise static welfare in some cases, but certain characteristics that are typical of new innovations have been ignored. Price discrimination can provide opportunities to serve new markets and to achieve scale and learning economies, both of which are important for many patented innovations.[5] Opening markets and achieving scale economies increase static welfare, thereby increasing the likelihood that price discrimination for patented goods will yield static welfare gains.

We demonstrate that patentee price discimination can be Pareto improving under some circumstances. This is a strong result, since Pareto improvements can occur without even counting the benefits of increasing innovation incentives. The rarity of Pareto improvements in applied settings emphasizes the importance of new markets and scale economies when analysing the intersection of patent and monopoly policy.

In addition to the marginal social efficiency of allowing price discrimination, *ceteris paribus* there are interesting questions about the optimal mix of different patent characteristics. Optimal policy will use the least-cost means of providing a given reward to innovators. For any given level of patent reward, some amount of price discrimination yields higher welfare than uniform pricing. Thus, even if innovators receive sufficient rewards without price discrimination, an appropriate policy may be to allow discrimination, while, say, reducing the life of a patent.

In the next section we will consider the role of new markets in the welfare analysis of price discrimination. Section 3 addresses scale or learning economies. Then we present as an example some facts from a recent patent case involving Du Pont's aramid fiber, Kevlar®. Aramid fiber is precisely the sort of major innovation that most economists believe patent policy is intended to encourage (Scherer 1980). In §5 we investigate the relative efficiency of price discrimination as one

---

[5] Of course, opening new markets and scale or learning economies are not limited to the production of patented innovations. The results in this paper apply generally to other instances in which price discrimination can occur. We think that the circumstances will most often arise for patents, because patents may establish the necessary monopoly power for price discrimination to take place, and because we expect new markets and scale economies to be more important for innovations than for 'mature' products.

instrument for providing rewards to innovators. Our conclusions are summarized in §6.

2. Opening markets with price discrimination

We first briefly review the known results on the welfare effects of third-degree price discrimination. Schmalensee (1981) and Varian (1985) have formalized and extended the basic results, which are originally due to Robinson (1933).

With fairly general assumptions one can show that, aside from any incentive externalities, a *necessary* (not *sufficient*) condition for price discrimination to increase static Marshallian welfare (the sum of consumers' and producer's surpluses) is that total output of the product increase. The intuition is straightforward. If different customers are paying different prices for a product, their marginal valuations are driven apart. Thus price discrimination necessarily leads to allocation inefficiences. For welfare to increase, total output must increase sufficiently for the resulting surplus gains to exceed the allocative losses.

Since output increases are not sufficient for welfare increases, it is usually necessary to analyse the specifics of each case. Schmalensee (1981) identifies some extreme cases for which there are definite welfare predictions. Varian (1985) derives bounds on the static welfare effect of discrimination, but general conclusions on the sign of the change are not possible.

The results above assume that all markets have positive demand under both price discrimination and uniform pricing. This assumption may often fail to hold. For example, when an intermediate good has applications in several different uses, reservation prices in the different uses may differ substantially, especially if there are different competing alternatives. The uniform-pricing firm may earn its highest profit at a price that excludes demand from low reservation-price uses.

When new markets may open, price discrimination can lead to Pareto welfare improvements. The result is very strong when there are only two potential markets:[6]

Definition 1. Two demand functions are *non-substitutable* if lowering the price to one market does not reduce the purchaser's surplus in the other market, at an unchanged price in the latter.[7]

[6] Throughout, when we say a demand function exists, we mean that there is positive demand at a price exceeding marginal cost.

[7] Non-substitutability is a 'non-envy' requirement for consumer utility. In the case of derived demands for an innovation by downstream producers of different products, we require that a lower input price to industry B not lower profits in industry A. The non-substitutability condition means that our proposition holds

Proposition 1. Suppose there are two different non-substitutable demand functions for a good, and one market is not served under uniform pricing. If marginal cost is constant or decreasing, then price discrimination will always (weakly) yield a Pareto improvement.

Proof. If one market, say market 2, is not served under uniform pricing, then the uniform price must be the monopoly price for market 1, say p_1^m.

Suppose marginal cost is constant at c; introduce discrimination. The patentee sets marginal revenue equal to c in each market. This yields p_1^m in market 1 again, and some price p_2^m in market 2. Consumer welfare in market 1 is unchanged, consumer surplus in market 2 has increased, and the patentee obtains higher profits, thereby establishing a Pareto improvement.

Now let marginal cost be decreasing. Consider charging (p_1^m, p_2^m) under discrimination. Total output is unambiguously greater than under uniform pricing (p_1^m, p_1^m), so that marginal cost is lower. Thus, to reach the profit-maximizing point ($MR = MC$), output in each market must be increased and price lowered below (p_1^m, p_2^m). There is a strict Pareto improvement with all customers (in both markets) and the patentee is better off. ■

Fig. 1 illustrates the proposition with two linear demand curves.

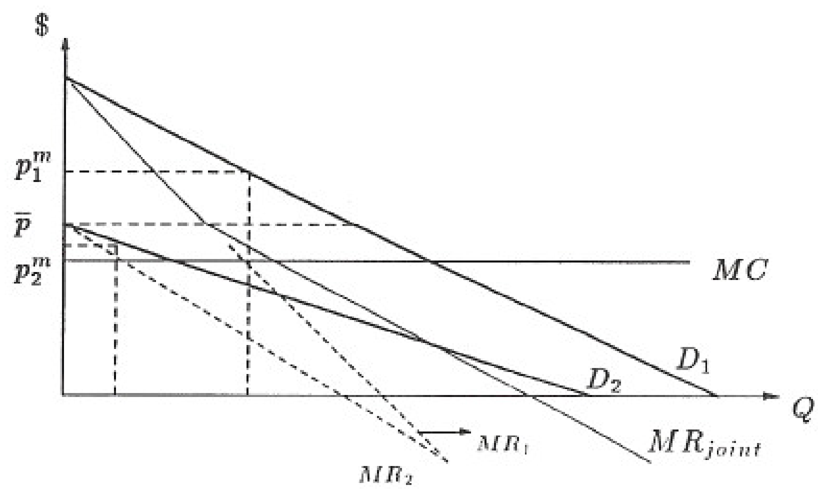

Figure 1
Pareto Improvement from a New Market

both for independent demands, such as those considered by Schmalensee (1981), and for interdependent demands that are complementary.

At the joint demand when price is the reservation price for market 2 ($\overline{p}$), marginal cost exceeds joint marginal revenue, so that the second market is excluded by the optimal uniform price, p_1^m. With discrimination market 1 pays p_1^m, and market 2 pays p_2^m.

When more than one market is served under uniform pricing, the price depends on an appropriately weighted average of demand elasticities. When discriminatory prices are introduced, some prices may rise while others fall, which violates the conditions for a Pareto improvement. But aggregate Marshallian welfare may increase. We restate a result from Schmalensee (1981) and then show that when new markets are introduced, welfare is more likely to increase.

Proposition 2. If there are more than two demand curves, all markets are served under uniform pricing, and marginal cost is non-increasing, then price discrimination may yield an increase in Marshallian welfare.

Proof. See Schmalensee (1981).[8] ■

As a corollary, note that *if marginal cost is constant*, then with more than one market served under uniform pricing, at least one discriminatory price *must* be higher than the uniform price, so that a Pareto improvement is not possible.[9] (We treat the case of decreasing marginal cost in §4.)

We now address the situation with multiple markets, when some are excluded by uniform pricing.

Proposition 3. If a patentee faces the same conditions as in Proposition 2, but a non-substitutable market can be served that is not served under uniform pricing, then price discrimination yields a strictly better welfare outcome than if the potential new market did not exist.

Proof. With constant marginal cost surplus in all previously served markets is not reduced by the introduction of a new market composed of non-substitutable demands. Consumer surplus increases in the new market, and profits increase. The argument for decreasing cost follows the proof of Proposition 1. ■

---

[8] Schmalensee assumes that marginal cost is constant, but the argument for decreasing marginal cost follows directly from his results and our proof of Proposition 1. For a Marshallian welfare increase, a total output increase is necessary; if marginal cost is decreasing, then necessarily all equilibrium prices will be lower than if cost is constant when output increases. Schmalensee also assumes that demands are independent, but Varian (1985) extends the result to interdependent demands.

[9] If there is constant marginal cost, the uniform price would raise profits from all markets. Lowering the uniform price a little further loses no profit in the highest-price market, but increases profits elsewhere.

Proposition 3 emphasizes one of our main points. When all markets are being served, Marshallian welfare may increase or decrease when price discrimination is allowed.[10] Proposition 3 suggests that when new markets can be served, the probability of a welfare increase is greater. In §4 we discuss reasons for and provide examples of why price discrimination may be especially important in opening new markets for innovations.

3. Discrimination with scale and learning economies

With innovative products one cn expect scale economies of at least two kinds during the early years of production. First, increases in output can lead to conventional scale economies, especially since initial production facilities may be built well below minimum efficient scale due to uncertainty about commercial success. Output increases can follow from outward shifts in the demand curve as information about the product diffuses or new uses are developed. Second, as the patentee accumulates production experience, unit costs may decline as a result of learning. The learning-curve effect is much the same as that of scale economies, with unit costs, say, a decreasing function of cumulative output over time.[11]

Scale and learning economies are important for three reasons. First, when discrimination leads to an output increase and Marshallian welfare gain, scale economies are achieved, which then increase the magnitude of the welfare gain. Second, when scale economies are possible, price discrimination can yield a Pareto improvement *even if no new markets open.* We demonstrate this below. Third, scale economies may lead to new markets' opening under discriminatory pricing when new markets would not open without such economies. Scale economies can be necessary for opening new markets when the marginal cost associated with uniform-pricing output exceeds the reservation price in potential new markets. For example, under price discrimination it may be possible to use a technology with higher fixed costs but lower variable costs owing to greater sales to markets that would not

---

[10] See, e.g. Schmalensee (1981); Scherer (1980). These authors observe that, *a priori*, the demand curve convexity conditions for a welfare improvement are no less likely than the conditions required for a welfare decrease.

[11] The expectation of learning effects is another reason that initial plants may be undersized; thus learning effects can reinforce conventional plant economies. For simplicity, we do not model the dynamics of investment and output, so that we cannot formally treat learning effects. See, e.g. Spence (1981). It can be shown, however, that our results generalize to learning economies since, as Spence has shown, learning economies can be formally described as a decreasing shadow marginal cost.

be served with another, higher-unit-cost technology.[12] Thus, when scale economies are possible, price discrimination may lead to Pareto improvements in two situations where such improvements could not otherwise occur.

Pareto improvements with scale economies: no new markets. We now demonstrate that in the presence of scale economies, price discrimination can be Pareto-improving, even without new markets' opening.[13] Let Q be the total output, $C(Q)$ total cost, p_1 the price to the i^{th} market, and η_i the demand elasticity in market i.

Definition 2. The *virtual elasticity*, $\eta^*(p)$, is the price elasticity of demand that necessarily obtains if price p is the monopoly profit-maximizing price for the market of interest; i.e. $\eta^*(p)$ solves

$$\frac{p - C'(Q)}{p} = -\frac{1}{\eta}.$$

Definition 3. The *uniform price*, p^u, is the profit-maximizing price when all markets are charged the same price.

Definition 4. The *uniform-price profit function*, $\Pi(p)$, is the profit earned if the monopolist charges the same price to all markets:

$$\Pi(p) = p\sum_{i=1}^{N} q_i - C(\sum_{i=1}^{N} q_i).$$

Definition 5. The *maximal discriminatory price*, $\overline{p}$, is the $\max\{p_i^m\}_{i=1}^N$, where there are N possible markets, and p_i^m is the price charged to the i^{th} market by a profit-maximizing, third-degree price discriminator.

Proposition 4. If there are two or more demand curves for a product, marginal cost is decreasing in total output, and the uniform price profit function is monotonic in the closed interval $[\overline{p}, p^u]$, and

$$|\eta^*(p)| \geq |\sum_{i=1}^{N} s_i(\overline{p})\eta_i(\overline{p})| \qquad (1)$$

where s_1 is the output share of the i^{th} market, $q_i/\sum q_i$, then price discrimination will (weakly) yield a Pareto welfare improvement over uniform pricing.

---

[12] We demonstrate this in Appendix 1.

[13] Proposition 4 will state a sufficient condition for Pareto improvement. A numerical example later will demonstrate that the condition is feasible. For simplicity, we assume in the analysis that scale economies follow from declining marginal cost, although in the example, economies are of the plant-scale variety.

Proof. We shall show that condition (1) implies $p^u \geq \overline{p}$, which implies that under price discrimination, the price to every market (weakly) falls. Let $\pi(p_1, \ldots, p_N)$ be the total profit function, so that $\Pi(p) = \pi(p, \ldots, p)$ is the uniform-pricing profit function. By monotonicity over the interval defined above, if $\Pi_p(\overline{p}) \geq 0$, then $p^u \geq \overline{p}$, where subscripts denote differentiation. Writing out the profit function, we have

$$\pi(p_1, \ldots, p_N) = \sum_{i=1}^{N} p_i q_i(p_i) - C\big[\sum_{j=1}^{N} q_j(p_j)\big]\,. \tag{2}$$

We shall prove the result if we can show that (1) is a sufficient condition or

$$\Pi_p(\overline{p}) = \sum_{i=1}^{N} \pi_i(\overline{p}, \ldots, \overline{p}) \geq 0\,. \tag{3}$$

Differentiating (2) (letting primes denote differentiation of demand functions) yields

$$\pi_i(\overline{p}, \ldots, \overline{p}) = q_i(\overline{p}) + \overline{p} q_i'(\overline{p}) - C'\big[\sum_{j=1}^{N} q_j(\overline{p})\big]\, q_i'(\overline{p}) \qquad \text{for } \ i = 1, \ldots, N\,. \tag{4}$$

Rewriting (4), we have

$$\begin{aligned} \pi_i(\overline{p}, \ldots, \overline{p}) &= q_i(\overline{p})\left[1 + \frac{\overline{p}}{q_i(\overline{p})} q_i'(\overline{p}) \left\{ \frac{\overline{p} - C'\big[\sum_{j=1}^{N} q_j(\overline{p})\big]}{\overline{p}} \right\}\right] \\ &= q_i(\overline{p}) - \frac{q_i(\overline{p})\eta_i(\overline{p})}{\eta^*(\overline{p})} \end{aligned} \tag{5}$$

by the definition of the implicit elasticity, $\eta^*(\overline{p})$.

Now sum equations (5):

$$\Pi_p(\overline{p}) = \sum_{i=1}^{N} q_i(\overline{p}) - \frac{1}{\eta^*(\overline{p})} \sum_{i=1}^{N} q_i(\overline{p})\eta_i(\overline{p})\,.$$

After rearrangement, we find that condition (3) is satisfied if and only if

$$\eta^*(\overline{p}) \ \leq \ \frac{\sum_{i=1}^{N} q_i(\overline{p})\eta_i(\overline{p})}{\sum_{i=1}^{N} q_i(\overline{p})} \tag{6}$$

or

$$\eta^*(\overline{p}) \ \leq \ \sum_{i=1}^{N} s_i(\overline{p})\eta_i(\overline{p})\,. \tag{7}$$

Thus, (7) is a sufficient condition for a Pareto improvement, given the other assumptions of the proposition.[14] ■

Remark 1. The proposition has the following intuition. For a Pareto improvement to occur, the discriminatory price in each market must be lower than the uniform price, for which it is sufficient to establish that the uniform price is greater than the maximal discriminatory price ($\overline{p}$). To show that $p^u > \overline{p}$, we need to show that uniform-pricing marginal profit is positive at $\overline{p}$, so that a profit-maximizing firm would set a uniform price above $\overline{p}$. Equation (7) provides a sufficient condition for this requirement. Of course, there may be a welfare increase even if a Pareto improvement is not achieved.[15]

Condition (1) is equivalent to marginal profit being non-negative at a uniform price of $\overline{p}$. The left side of the condition is the aggregate demand elasticity that must obtain if $\overline{p}$ is the optimal uniform price. The right side is an expression for the actual elasticity of demand at $\overline{p}$, $\sum s_i \eta_i = (dQ/dp)(\overline{p}/Q)$. Thus, the condition is that if the actual demand elasticity at $\overline{p}$ is too low for $\overline{p}$ to be the uniform price, the uniform price must be higher than $\overline{p}$ (since demand elasticities are negative, $\eta^* < \sum s_i \eta_i$ means the right-hand side is lower in absolute value). This is related to the monopoly pricing rule: raise price until demand becomes elastic enough that no further profit gains are possible.

Remark 2. We use a numerical example to demonstrate below that condition (1) is feasible. In fact, the example demonstrates that Pareto improvements can occur under a wider variety of conditions than those in Proposition 4.

Remark 3. It is not obvious from condition (1) why scale economies are necessary to get a Pareto improvement. To see that they are, consider the following illustration of how a Pareto improvement can occur. Suppose there are two markets: one is relatively inelastic and profitable under discriminatory pricing, the other is relatively elastic and not very profitable. It is intuitively clear that the optimal uniform price will be set near the monopoly price for the first, inelastic market; this sacrifices the small profits in the second market for the greater

---

[14] Equation (6) requires that $L(\overline{p})$ be non-zero. Since a uniform-pricing monopolist sets marginal revenue of the joint demand curve equal to marginal cost, $L(\overline{p})$ can be zero only if the uniform-pricing aggregate demand is zero ($P = MR$), in which case a weak Pareto improvement from price discrimination follows immediately.

[15] We require local monotonicity here so that we can rely on the first-order conditions. Monotonicity in profits is not necessary for a Pareto improvement, however, as we show by our example below.

profits in the first. At that relatively high uniform price, demand in the elastic, second market is relatively low, so that total output is low and marginal cost is high.

Now let the firm discriminate. The price to the elastic market is dropped substantially, and as a result it elicits a big increase in output. With scale economies marginal cost falls. To a first-order approximation the uniform price was already profit-maximizing for the inelastic market, so that with all else equal the discriminatory price will not be much higher. But the elastic market has a flywheel effect: it drives marginal cost down through a large output increase and thus moves the $MR = MC$ profit-maximizing equilibrium price in the inelastic market *lower* than the uniform price, by moving down the MR curve to reach the new, lower MC.[16]

Using the first-order condition approach of Proposition 4, we can easily demonstrate the necessity of declining marginal cost if we assume that, given a vector of prices in the other markets, profits in each market are single-peaked. (Global concavity of the joint profit function is sufficient to ensure this.) Let market 1 be the market with the maximal discriminatory price, $p_1^m = \overline{p}$. For markets $i = 2, \ldots, N$, marginal profit at a uniform price of $\overline{p}$ (equation (4)) is negative because by definition $\overline{p}$ is above the profit-maximizing price, p_1^m, for these markets. Then, for (3) to hold, it must be that $\pi_1(\overline{p}, \cdots, \overline{p})$ is positive, since all other terms in the sum are negative. Use the price-discrimination, first-order condition for market 1,

$$\pi_i(\overline{p}, p_2^m, \ldots, p_N^m) = q_1(\overline{p}) + \overline{p}q_1'(\overline{p}) - C'\Big[\sum_{j=1}^{N} q_j(p_j^m)\Big]q_1'(\overline{p}) = 0\,, \quad (8)$$

to substitute into condition (4) for $i = 1$, and obtain,

$$\pi_1(\overline{p}, \cdots, \overline{p}) = \pi_1(\overline{p}, p_2^m, \ldots, p_N^m) + q_1'(\overline{p})\Big\{C'\Big[\sum_{j=1}^{N} q_j(p_j^m)\Big] - C'\Big[\sum_{j=1}^{N} q_j(\overline{p})\Big]\Big\}\,. \quad (9)$$

From (8), the first term on the right of (9) is zero. The second term is positive if and only if marginal cost is decreasing, because

[16] Note that scale economies generalize the results about Pareto improvements from opening new markets that were described in the previous section. The flywheel market is more-or-less 'closed' at the uniform price, but 'opens' with discrimination. Introducing scale economies, however, makes Pareto improvements possible with more than two markets; with constant costs Pareto improvements could only occur in the special two-market case.

total output is greater at $\{p_i^m\}$ than at $\overline{p}$. Thus, decreasing marginal costs are necessary for the result to hold.

In fact, to satisfy equation (1), the scale economy effect, $C'(p^m) - C'(\overline{p})$ must be sufficiently large for $\pi_1(\overline{p}) > -\sum_{i=2}^{N} \pi_i(\overline{p})$. This clarifies the flywheel notion: if there are markets for which demand increases rapidly enough as the uniform price is lowered, then marginal cost can be driven down sufficiently to obtain a Pareto improvement.

Condition (1) for Pareto improvement is intuitive and simple. With knowledge of the separate market demand curves and the cost function, it can be calculated for actual cases.[17] The following numerical example will demonstrate that Pareto improvements are possible under even more general conditions.

Numerical example. In Proposition 4 we derived sufficient conditions for a Pareto improvement when the uniform-price profit function is monotonic on $[\overline{p}, p^u]$. To demonstrate that welfare improvements are feasible and can occur under much more general conditions, we now present an example that violates the monotonicity conditions of Proposition 4.

Suppose the patented innovation can be produced in one of two different scale plants with costs

$$C_x = 10Q$$
$$C_y = 65 + 5Q$$

for plants X and Y respectively. Scale economies could be achieved at sufficiently high output by switching from plant X to plant Y, which has lower marginal but higher fixed costs. The global profit function has two peaks because of the switch from plant Y to plant X at higher prices.

Let the demand curves be

$$Q_1 = 1,000\, p_1^{-1.3}$$
$$Q_2 = 1,000,000\, p_2^{-5}$$

for markets 1 and 2 respectively. Under uniform pricing the patent holder builds plant X and charges a price of \$42.97. If the patentee can price discriminate, plant Y will be more profitable, and the optimal prices will be $p_1 = \$21.67$ to market 1 and $p_2 = \$6.25$ to market 2.

[17] We are not recommending that a finding of a Pareto, or even a Marshallian, welfare gain should be necessary for the permissibility of price discrimination in specific cases. Spillovers from the effect of profits on future innovative efforts by all firms also need to be taken into account in a general evaluation.

Consumers in both markets are made better off from the substantially lower prices made possible by the scale economies associated with plant Y; price falls by about 50 per cent in market 1 and by over 50 per cent in market 2. The firm's profits rise from \$249 to \$372.

4. Kevlar®: an example

In this section we illustrate our results with some facts from a patent case concerning Du Pont's aramid fiber, Kevlar®. Du Pont's adversary charged that third-degree price discrimination practices were a misuse of the patent and thus an antitrust violation.[18] Kevlar® is a superstrength synthetic fiber that has a strength-to-weight ratio five times as great as steel. It is used in end-use applications such as tyre belts, bullet-resistant vests, undersea cables, aircraft components, and missile casings. Du Pont charges different prices for Kevlar® in unrelated end uses.

Kevlar® is a typical example. Third-degree price discrimination is often associated with new materials and chemicals that are intermediate inputs with several potential uses. The industry segment of chemicals and allied products is one of the most important for R&D, as such firms hold 22 per cent of all patents held by the U.S. manufacturing sector.[19]

Consider first the role of price discrimination in opening new markets. Aramid fibers can be used in friction products and gaskets, markets previously served predominantly by asbestos. Du Pont prices Kevlar® sold to the friction/sealer market to compete with asbestos. This price has been substantially below the quantity-weighted average price for all Kevlar® sold. Our calculations suggest that the optimal uniform Kevlar® price is not much different from the current average price. Thus, requiring uniform pricing would dramatically increase the Kevlar® price to the friction/sealer market and result in the loss of that use for aramid fiber to asbestos.[20] A similar outcome would

[18] *Akzo N.V.* v. *USITC*, 808 Fed 2d 1471, 1488–89 (Fed. Circ. 1986), cert. denied, __ US __ 1987. For Akzo's (the respondent firm) position on price discrimination, see, for example, Respondent's Prehearing Statement, pp. 80–93. We provided economic testimony on Du Pont's behalf in this proceeding. Much of our knowledge of the facts is based on information submitted to the USITC under confidentiality procedures that preclude public disclosure of such facts as actual prices and costs. Nevertheless, our qualitative discussion is supported by detailed analysis of corporate documents from both Du Pont and Akzo.

[19] According to COMPUSTAT data, chemicals and allied products accounted for over 20 per cent of corporate R&D spending in the United States in 1976. See Bound *et al.* (1984).

[20] Positive social externalities may also be lost when new markets are excluded. For instance, aside from the loss in consumer surplus from forcing customers

occur in the market for Kevlar® as a tyre-belting material, where the competitive substitutes (primarily steel and fiberglass) have low costs.

In general, price discrimination may open new markets for innovations because it helps to recover the high cost of discovering new uses and adopting innovations for new uses.[21] The incremental profits from opening a new market under uniform pricing may not provide sufficient incentive to sink the necessary opening costs; the market-specific profit gains from price discrimination could make these incremental investments more attractive.

There are several costs of finding new uses for an innovation. First, the product or process may not be well understood; new properties, or new combinations of properties, can make surprising applications possible. In the case of Kevlar® strength-related uses (such as in bullet-resistant vests and tyre belts) were obvious from the first, but the use of pulped aramid fiber in friction/sealer products was discovered only several years later, after considerable research. New uses may require further R&D to modify the properties of the product or to learn how to combine its properties with other goods.[22] Some important applications of Kevlar® in the tyre and mechanical rubber goods industry took more than ten years to emerge.[23]

The costs of adopting an innovation can also be important. Production processes often involve a complex series of stages, each calibrated to the specific characteristics of the raw materials to maximize production efficiency. In many instances machines are specifically designed to work with a particular combination of inputs. When new inputs are considered, a 'proving out' process is necessary, during which the performance of the new material is tested and production runs are redesigned to reoptimize the over-all process.

---

to use asbestos, when with discriminatory pricing they preferred Kevlar® (for price and/or quality reasons), society would also lose because of higher output of asbestos, a known carcinogen.

[21] In fact, for the Kevlar® programme through about 1984, Du Pont has spent approximately five to seven times as much on R&D and market development *after* the patent was awarded than it spent before.

[22] ITC Investigation No. 337–TA–194, Findings of Fact, p.381. 'When Kevlar® was first offered, potential customers were not familiar with the product and ready markets did not exist. Du Pont has had to create markets for Kevlar®, and continues to do so. Du Pont personnel have identified potential end uses ... developed the technical expertise required to utilize Kevlar® for those end uses, and educated manufacturers and consumers involved in the end-use markets to persuade them to use Kevlar®.'

[23] Findings of Fact, pp. 438–9. Kevlar® was patented in 1973. 'The development of an adhesion activated Kevlar® product had been a technical objective at Du Pont as early as 1977', but the new product '[was not] field tested [until 1985].'

Adoption of Kevlar® has been slowed in several markets by the 'proving out' process, sometimes by as much as two years or more. For example, some tyre manufacturers have switched from steel to fiberglass to Kevlar® for high-performance tyres. The properties, working requirements, and handling characteristics are remarkably dissimilar for inputs that ultimately serve similar ends. Likewise, in aircraft fuselage applications, very detailed testing, design, and performance specification is required to evaluate and choose among aramid fibers, graphite composites, fiberglass, and various plastics.[24]

Plant scale economies and learning curve effects are also often important, especially in the chemical and materials industries. For instance, Lieberman (1984) has found that typical learning economies yield a 78 per cent decrease in marginal costs for each doubling of output. Learning-curve calculations for Kevlar® indicated that unit costs fell by some 60 per cent for each doubling in cumulative output during the first ten or twelve years of production. Scale economies are also a major phenomenon in chemical and related industries.[25]

Since obtaining the aramid fiber patents, Du Pont built a series of three successively larger plants for the spinning of Kevlar®. The first two were called 'market development facilities', the first of which was gradually retired as the next plant in the series came on line. The last plant produces at substantially lower unit cost. Du Pont also achieved a very large one-time drop in unit costs by building a new, larger raw materials plant.[26]

5. Efficient provision of the rewards to innovation

Even if price discrimination sometimes incurs net static welfare losses, policy discussions should be concerned with the efficiency of the trade-off between innovation incentives and static welfare losses. For instance, patent rewards could be decreased either by shortening the life of patents, or by disallowing price discrimination. What is the most efficient pricing scheme the patentee can employ to obtain a given level

[24] Findings of Fact, pp. 389–90. 'Most of the ... manufacturers who purchase Kevlar® require that the Kevlar® "qualify", that is, meet their specification, which is usually a time-consuming and costly process. Any modification in the properties or process of making Kevlar® ordinarily requires requalification ... Work in [passenger tyre, aircraft composites, and other] end uses took at least five and usually closer to ten years before commercial sales were achieved.'

[25] Lieberman (1984) estimates an 11 per cent reduction in marginal costs as plant scale is doubled. Scherer (1980) discusses such economies and has numerous references to empirical validation.

[26] The Pontchartrain plant, which produces the raw materials for Kevlar®, has a capacity nearly double the level of demand at the time the facility was built. The new Spruance spinning facility has twice the output rate of the prior facility.

Table 1. Optimal patent lives and welfare indices*

	Price discrimination		Uniform pricing	
η_1	T^*	Index	T^*	Index
0.25	13	0.744	40	0.486
0.50	12	0.729	28	0.461
0.75	12	1.716	22	0.441
1.00	11	1.705	19	0.425
1.25	11	1.700	17	0.411
1.50	10	1.687	16	0.399
1.75	10	1.680	14	0.389

* T^* is the optimal patent life, index is the ratio of obtained social welfare to the first-best, and η_1 is the final demand elasticity in market 1.

The final demand elasticity in market 2 (η_2) is 2.00. The social discount rate is 3 per cent. The elasticity of the innovation production function with respect to investment is 0.1. The pre-innovation cost in market 1 is 1.5 times as great as in market 2. Appendix 2 gives further details and the derivation of the calculations.

of profit? The problem is that of Ramsey (1927) pricing, which is most often applied to analyse the regulation of natural monopolies (Boiteux 1971). The solution when demands are independent is to charge different prices to different groups of customers, with prices inversely proportional to the demand elasticity.[27] If there is no minimum profit, price should equal marginal cost in each market. To maximize the patentee's profit, however, the optimal prices are precisely those that an unconstrained third-degree price discriminator would charge.

Some amount of price discrimination thus appears to be an efficient way to provide an innovator with a profit reward. Even if discrimination does not always yield the static welfare gains discussed earlier, it might be more efficient, for example, to let the patentee earn $20 a year for five years with discrimination rather than $12 a year for a ten-year patent life with uniform pricing.[28]

In Table 1 we have calculated some examples to illustrate the trade-off between different mechanisms for rewarding innovators, based on Nordhaus' (1972) work on optimal patent lives. Nordhaus calculated patent lives that balanced efficiency costs against innovation incentives for various parameter values; we recalculate with two derived demand curves, rather than one. We first calculate optimal patent lives when

[27] A qualitatively similar condition holds for non-independent demands.

[28] These hypothetical numbers assume a discount rate of about 10 per cent.

price discrimination is permissible and compare those with optimal lives when pricing is uniform. Then we calculate social welfare indices that indicate the percentage of the first-best level of welfare that can be obtained with price discrimination or uniform pricing.[29]

Allowing patentees to price discriminate increases the Marshallian efficiency of the patent system. For instance, with derived demand elasticities of $\eta_1 = -0.75$ in market 1, and $\eta_2 = -2.0$ in market 2, the optimal patent life with discrimination is 12 years, while with uniform pricing, $T^* = 22$ years.[30] The optimal life with price discrimination is not shortened to reduce increased static monopoly costs. Rather, with price discrimination the net present value index of social welfare is *higher* than with uniform pricing. In the case just mentioned, price discrimination achieves about 72 per cent of the first-best welfare level, while uniform pricing achieves only 44 per cent of that level. These results hold across a wide range of price elasticities.

6. Conclusion

The casual notion that third-degree price discrimination is good for the monopolist but bad for the public is not true as a general proposition. We have discussed reasons why it may be even less likely that the proposition is true when the monopolist has obtained dominant power through a valid patent on an innovation. We have shown not only that price discrimination by a patentee may often increase Marshallian welfare, but that it could even lead to Pareto welfare improvements.

Two special circumstances associated with patented innovations are important for the welfare effects of price discrimination. First, patented innovations for which discrimination is feasible are often intermediate inputs with applications in widely differing markets. With uniform pricing some markets may not be served (because, say, a competitive substitute is priced lower than the patented good's uniform price), and thus total output of the patented good and welfare may be lower.

---

[29] The assumptions and model used to generate the results are summarized in Appendix 2.

[30] Final demand elasticities less than unity (for example, $\eta_1 = -0.75$) are not inconsistent with monopoly profit maximization in this setting. The 'upward' elasticity—the demand elasticity if price is raised—of the *derived* demand curve facing the monopolist will be greater than unity, as usual. In fact, in the single-demand-curve case studied by Nordhaus, the derived demand upward elasticity is infinite for 'run-of-the-mill' innovations, regardless of the final demand elasticity. If the price of the patented good is raised a little, final good producers will switch all demand to the competing substitute, even if the final demand they face has elasticity less than unity.

Further, and quite important for a new product, declining marginal costs from scale and learning economies may be possible with increasing output. If discimination opens new markets, such economies can increase the welfare gains. Scale economies also make it more likely that new markets will open with discrimination, thereby leading to welfare gains. Moreover, price discrimination with scale economies can yield Pareto improvements in multiple market situations, when new markets alone cannot.

Finally, we have shown that some price discrimination is a relatively efficient way to obtain a given level of profit. The profits a patentee would earn with uniform pricing could also be earned with discriminatory prices and less static welfare loss. Thus, an optimal policy that trades off monopoly costs against the incentive effects of the patent reward should not disallow all discrimination.

In many legal cases the fact of discrimination has not been found to be a patent misuse, but the practices used by the patentee to implement the pricing schedule have sometimes been judged to be monopolistic misuses of the patent grant. The use of price discrimination often involves what appear to be 'vertical restraints', which are often frowned on by antitrust law.[31] For instance, Du Pont's adversaries argued before the ITC that one Du Pont sales arrangement was an illegal vertical restriction.[32] Du Pont sells Kevlar® on the following basis: the buyer can purchase at the price associated with a particular end use of Kevlar®, but if the buyer resells the Kevlar® to another company for a different use (or applies Kevlar® to a different use itself), then the buyer is required to pay Du Pont the difference between the list prices for the two uses.

If price discrimination is expected to be socially desirable overall, it makes little sense to proscribe otherwise harmless means of implementing it. In fact, by ruling out certain implementation schemes, the courts may lead patent holders to devise more costly methods for price discriminating, thus dissipating some of the surplus and reducing welfare gains.[33] Using vertical restraint doctrine to prevent third-degree price discrimination may be socially unproductive.

---

[31] Third-degree discrimination is not feasible unless arbitrage across markets is prevented. Patent holders frequently impose conditions on sales intended to prevent such arbitrage.

[32] In this case, the argument was unsuccessful. See *Akzo N.V.* v. *USITC*, 808 Fed 2d 1471, 1488–89 (Fed. Circ. 1986), cert. denied, __ US __ 1987.

[33] Williamson (1979) is concerned with the reduction in the gains from discrimination due to transaction costs of implementation.

Appendix 1

Pareto improvements with scale economies: new markets

In the text, we demonstrated that price discrimination can improve welfare through output increases and unit cost decreases. Output can increase most obviously by opening new markets. We have observed that the welfare gain from new markets is increased when scale economies are achieved as well. In this appendix we show that scale economies may be necessary for new markets to be opened under price discrimination. That is, even with discriminatory prices, some markets may be shut out unless increased output lowers marginal cost.

A typical case would be one in which a new product can be produced in one of two plants sizes; one plant has lower fixed but higher variable cost. If only the plant with lower fixed/higher variable cost is technologically feasible, then some markets may not be served with either uniform or discriminatory pricing. Those markets may also not be served under uniform pricing by the other, higher fixed/lower variable cost plant, if it is feasible. However, discriminatory pricing may allow the firm to build the more efficient plant, which can reduce marginal cost, open new markets, and thus increase social welfare.

We find these considerations to be particularly relevant for new products, when a patent monopolist is faced by initial plant scale and technology investment decisions. For example, as mentioned earlier, because of large-volume, low-price sales of *K* to the tire and friction-sealant markets, DuPont was able to build large facilities for commercial production of the raw ingredients and for the spinning of aramid fibers.

We will use an example to demonstrate the possibility of a Pareto improvement when new markets are opened only if scale economies are possible. Suppose a patentee has a choice of constructing plants X and Y, with production costs given by

$$C_x = d_x + c_x Q$$
$$C_y = d_y + c_y Q$$

for plants X and Y respectively, with d_x, d_y, c_x, c_y non-negative. Suppose $c_x > c_y$ and $d_x < d_y$, so that plant Y has lower variable but higher fixed cost.

Now consider two potential markets. Let market 1 have a reservation price which is greater than c_x per unit (and thus greater than c_y); let market 2 have a reservation price less than c_x, and greater than c_y. Under this assumption, market 2 will never be served under any rational pricing scheme if the patentee builds plant X (marginal revenue is always less than marginal cost). For example, suppose demands are linear:

$$p_1 = a_1 - b_1 q_1$$
$$p_2 = a_2 - b_2 q_2$$

for markets 1 and 2 respectively, with $a_1 > c_x$, $a_1 > c_y$, $a_1 > a_2$, $c_x > a_2 > c_y$ to satisfy the assumptions above. Since $a_2 < c_x$, market 2 will not be served by plant X under either uniform or discriminatory pricing. If only plant X is available—i.e. there are no plant scale economies—then moving from uniform to discriminatory prices leaves welfare unchanged because in both cases the firm charges the monopoly profit-maximizing price for market 1, and sells the corresponding quantity only to market 1 purchasers.

We can derive the condition under which market 2 would also not be served by uniform pricing even if plant Y is built; the condition requires that marginal cost be greater than marginal revenue (from the combined markets) at a price of $p = a_2$. This condition ($MC > MR$ at $p = a_2$) can be manipulated to obtain:

$$c_y > \frac{2a_2 b_2 + a_2 b_1 - a_1 b_2}{b_1 + b_2}.$$

If we further impose the following condition on plant fixed costs:

$$d_y > \frac{c_y^2 - 2a_1(c_x - c_y) - c_x^2}{4b_1} + d_x,$$

then under uniform pricing, with only market 1 served regardless of plant choice, the patentee will choose to build plant X because $\Pi_x > \Pi_y$ (using obvious notation for profits).

We now have an example in which, if plant economies are unavailable (plant Y technology doesn't exist), discriminatory pricing won't affect welfare. Further, if plant economies are feasible (plant Y), they will not be achieved under uniform pricing. It is easy to show, however, that if plant Y is available, a patentee charging discriminatory prices will choose to build plant Y if

$$d_y < \frac{c_y^2 - 2a_1(c_x - c_y) - c_x^2}{4b_1} + d_x + \frac{(a_2 - c_y)^2}{4b_2},$$

because then $\Pi_y > \Pi_x$. Discriminatory pricing in this case yields a Pareto improvement, by opening new markets and achieving plant economies.

The intuition extends the ideas from the text: when price discrimination is possible, some markets may be served. With the scale economies typically possible for patented innovations, it is more likely that new markets will be served, leading to welfare gains in many cases.

Appendix 2

Calculation of optimal patent lives

Our calculations of optimal patent lives under different assumptions closely follow Nordhaus (1972). We extend his analysis to consider a patented innovation which can be sold to two distinct markets. Each market is assumed to consist of downstream producers of a final good. The innovation reduces the marginal production cost for the two final goods. Thus, the potential purchasers of the innovation can either pay the license fee to use the innovation, or can continue to produce with existing technology.

Normalize units so that unit production costs in each downstream market are given by $\tilde{c}$ when the innovation is utilized. We will in fact be concerned with the percentage cost reduction in each market, given by

$$B_i = \frac{c_i - \tilde{c}}{c_i} \qquad \text{for } i = 1, 2. \tag{A1}$$

We normalize c_1 to unity, and for tractability assume that $c_2 = c_1/k = 1/k$.

Let the innovation production function (IPF) be parametrized so that

$$B_1 = \beta R^{\alpha} \tag{A2}$$

where R is the level of research effort undertaken by the innovator. We assume that the innovation is 'run-of-the-mill' for both downstream markets, following Nordhaus. That is, the value-added due to the innovation is sufficiently small that an optimally-pricing patent monopolist will charge a license fee in each market equal to the unit cost savings times pre-innovation output in that market. The innovation is not drastic enough to warrant lowering the production costs of the downstream producers and raising total output with a lower license fee. Technically, the condition for 'run-of-the-mill' innovations when downstream demands are linear is that

$$\eta_i \beta_i < 1 \qquad \text{for } i = 1, 2. \tag{A3}$$

When an innovation is non-drastic, the optimal license fees are given by

$$\pi_1 = (c_1 - \bar{c})X_0^1 = B_1 X_0^1 \tag{A4}$$

$$\pi_2 = (c_2 - \bar{c})X_0^2 = \frac{1}{k} B_2 X_0^2 \tag{A5}$$

where X_0^j is the pre-innovation output in the jth market. Below we will use the relationship that $B_2 = 1 - k\bar{c} = (1-k) + kB_1$.

Optimal patent life with price discrimination

The value of a T-year patent to the innovator is calculated as

$$V = \int_0^T \left[B_1(R)X_0^1 + \frac{1}{k} B_2 X_0^2 \right] e^{-rt} \mathrm{d}t - sR \tag{A6}$$

where s is the unit price of research effort. The firm's necessary condition for optimal research is

$$\mathrm{D}VR = \int_0^T \left[B_1'(R)X_0^1 + \frac{1}{k} B_2'(R)X_0^2 \right] e^{-rt} \mathrm{d}t - s \ = \ 0. \tag{A7}$$

Manipulation of (A7) and normalization of X_0^1 to unity yields

$$R^* = \left(\frac{\alpha\beta\phi[1 + X_0^2]}{rs} \right)^{\frac{1}{1-\alpha}} \tag{A8}$$

and

$$B_1(R^*) = \beta \left(\frac{\alpha\beta\phi[1 + X_0^2]}{rs} \right)^{\frac{\alpha}{1-\alpha}} \tag{A9}$$

where $\phi \equiv 1 - e^{-rT}$.

Rather than estimate all of the parameters of the IPF, Nordhaus calculates the social indifference curve between patent life and given levels of equilibrium average innovation size, B. To do this, he finds the patent life which maximizes social welfare for any given equilibrium B. We, however, wish to actually calculate optimal patent lives under two different pricing rules (discrimination and uniform), and compare the attained levels of welfare. We follow an approach similar to Nordhaus, by assuming that the equilibrium level of innovation (in market 1 terms) is some particular B_1 for a patent life of 10 years, and then use equation (A9) to solve for

$$\beta = \beta\left(B_1, T, X_0^2, s\right) \tag{A10}$$

(we normalize s to unity).

Marshallian welfare (the sum of producer's and consumer's surpluses) for price discrimination is

$$\begin{aligned} W_d = &\int_0^\infty B_1(R) X_0^1 e^{-rt} \mathrm{d}t + \int_0^\infty \frac{1}{k} B_2(R) X_0^2 e^{-rt} \mathrm{d}t \\ &+ \int_T^\infty \frac{1}{2}(X_1^1 - X_0^1) B_1(R) e^{-rt} \mathrm{d}t \\ &+ \int_T^\infty \frac{1}{2k}(X_1^2 - X_0^2) B_2(R) e^{-rt} \mathrm{d}t - sR \end{aligned} \tag{A11}$$

where X_1^j is the post-patent demand in market j, when unit cost is driven to $\bar{c}$ by competition.

We let the linear downstream demands be given by $X^j = \gamma_j - \eta_j p_j$. After integration and some manipulation, social welfare can be written as

$$\begin{aligned} W_d = &\frac{(\eta_1 + \eta_2)(1-\phi)}{2r} B_1^2(R) + \frac{(1 + X_0^2) + \eta_2 \frac{1-k}{k}(1-\phi)}{r} B_1(R) \\ &+ \frac{\frac{1-k}{k} X_0^2}{r} + \left(\frac{1-k}{k}\right)^2 \frac{\eta_2(1-\phi)}{2r} - sR. \end{aligned} \tag{A12}$$

We maximize this welfare function to find the optimal patent life, making various assumptions about the elasticity of the IPF (α) and the demand slopes (η_1, η_2). We find β by assuming that the average equilibrium innovation size for a ten-year patent is a 40 per cent cost reduction in market 1 (B_1), and that market 1 pre-innovation cost is 1.5 times as great as market 2 pre-innovation cost (k).

Optimal patent life with uniform pricing

Under the assumptions above which assure that the innovation is 'run-of-the-mill' to both markets, the optimal uniform pricing strategy will be to either charge a license fee of $B_u = B_1$ and not serve market 2 (Case A), or charge $B_u = B_2/k$ and serve both markets (Case B). The value of a patent with life T in the two cases will be

$$\begin{aligned} V_A &= \int_0^T B_1 e^{-rt} \mathrm{d}t - sR \\ V_B &= \int_0^T \frac{1}{k} B_2 (1 + X_0^2) e^{-rt} \mathrm{d}t - sR. \end{aligned} \tag{A13}$$

Using the relationship between B_1 and B_2 derived above, V_B can be rewritten as

$$V_B = \int_0^T \left(1 - K\eta_1 + X_0^2\right)\left(B_{1B} + K\right) e^{-rt} \mathrm{d}t - sR_B \tag{A14}$$

where B_{1B} is the optimal value of the innovation to market 1 under Case B; R_b is the optimal level of innovative effort under Case B; and $K = (1-k)/k$. These values can be found by solving the necessary condition for a maximum of V_B; they are:

$$B_{1B} = \beta \left(\frac{(1 - K\eta_1 + X_0^2)\alpha\beta\phi}{rs} \right)^{\frac{\alpha}{1-\alpha}} \tag{A15}$$

$$R_B = \left(\frac{(1 - K\eta_1 + X_0^2)\alpha\beta\phi}{rs} \right)^{\frac{1}{1-\alpha}}. \tag{A16}$$

Case A can be likewise solved for optimal innovative effort and equilibrium invention level. Since Case A involves only one end-use market, it is identical to the case studied by Nordhaus. For each set of parameters considered, and each candidate optimal patent life, T^*, we calculate V_A and V_B, and determine which pricing strategy maximizes firm profits.

Marshallian social welfare for Case B is straightforwardly calculated as before. After integration and manipulation, the welfare function can be written as

$$\begin{aligned} W_B \; = \; & \frac{\theta(1 - K\eta_1 + X_0^2) + K(1 - K\eta_1)}{r} + \frac{(\eta_1 + \eta_2)\theta^2(1-\phi)}{2r} \\ & - s\left(\frac{(1 - K\eta_1 + X_0^2)\alpha\beta\phi}{rs} \right)^{\frac{1}{1-\alpha}} \end{aligned} \tag{A17}$$

where $\theta = K + B_{1B}$. As shown in Nordhaus (1972), welfare in Case A can be calculated as

$$W_A = q\left(\phi^{\alpha v} + \frac{qr\phi^{2\alpha v}\eta_1(1-\phi)}{2} - \alpha\phi^v\right) \tag{A18}$$

where $q = \beta^v \alpha^{\alpha v}/r^v$.

For each candidate optimal patent life under uniform pricing, welfare is calculated after determining whether the firm uses pricing strategy A or B.

After calculating optimal patent lives under uniform and discriminatory pricing, the attained social welfare levels are divided by the first-best level of attainable welfare (which would follow from the government buying the optimal level of research effort) to obtain indices of attained welfare under the two policies. The first-best welfare level is

$$\begin{aligned}\hat{W} &= \frac{(\eta_1+\eta_2)}{2r}\beta\left(\frac{\alpha\beta(1+X_0^2)}{rs}\right)^{2\alpha v} \\ &+ \frac{(1+X_0^2)+\eta_2 K}{r}\beta\left(\frac{\alpha\beta(1+X_0^2)}{rs}\right)^{\alpha v} \\ &+ \frac{KX_0^2}{r}+\frac{K^2\eta_2}{2r} - s\left(\frac{\alpha\beta(1+X_0^2)}{rs}\right)^{v} \qquad \text{(A19)}\end{aligned}$$

(Since we care only about welfare index ratios, it doesn't matter which particular assumption about the first-best welfare attainable we make, as long as the measure is invariant with respect to T.)

Nordhaus calculated welfare indices and optimal patent lives by fixing the equilibrium invention size across all variations of patent life and demand elasticities. Since the purpose of the patent grant is to provide incentives to firms to invest in innovations, this calculation is incorrect. Our approach chooses one particular combination of equilibrium invention size and patent life to fix the parameters of the invention production function (IPF), and then let firms optimally choose levels of inventive effort and resulting invention size as a function of patent life and derived demand elasticities.

For our calculations, we assumed that with a ten-year patent life, the average equilibrium invention would reduce costs by 40 per cent in market 1;6 we also assumed that pre-innovation unit costs in market 1 were 1.5 times as high as in market 2. We then calculated optimal patent lives for all combinations of market 1 final demand elasticities between 0.25 and 2.0, and market 2 final demand elasticities between 0.25 and 5.0. The results are reported in Appendix Table 1.

References

Baxter, W. (1966). Legal restrictions and exploitations of patent monopoly: An economic analysis. *Yale Law Journal*, **76**, 267–.

Boiteux, M. (1971). On the management of public monopolies subject to budgetary constraints. *Journal of Economic Theory*, **3**, 219–40.

Bound, J., Griliches, Z., Hall, B., and Jaffe, A. (1984). Who does R&D and who patents? In *R&D, patents, and productivity*, (ed. Z. Griliches). University of Chicago Press.

Bowman, W. S. (1973). *Patent and antitrust law*. University of Chicago Press.

International Trade Commission (1985). 'Findings of Fact', In the Matter of Certain Aramid Fiber, ITC No. 337–TA–194.

International Trade Commission (1985). 'Respondent's Prehearing Statement', In the Matter of Certain Aramid Fiber, ITC No. 337–TA–194.

Kaplow, L. (1984). The patent-antitrust intersection: a reappraisal. *Harvard Law Review*, **97**, 1813–92.

Lieberman, M. B. (1984). The learning curve and pricing in the chemical processing industries. *Rand Journal of Economics*, **15**, 213–28.

Neale, A. D., and Goyder, D.G. (1982). *The Antitrust Laws of the U.S.A.* Cambridge University Press.

Nordhaus, W. D. (1972). *Invention, growth and welfare: a theoretical treatment of technological change.* MIT Press.

Ramsey, F. (1927). A contribution to the theory of taxation. *Economic Journal*, **37**, 47–61.

Robinson, J. (1933). *Economics of imperfect competition.* Macmillan, London.

Scherer, F. M. (1980). *Industrial market structure and economic performance.* Houghton Mifflin, Boston.

Schmalensee, R. (1981). Output and welfare implications of monopolistic third-degree price discrimination. *American Economic Review*, **71**, 242–47.

Spence, A. M. (1981). The learning curve and competition. *Bell Journal of Economics*, **12**, 49–70.

Varian, H. R. (1985). Price discrimination and social welfare. *American Economic Review*, **75**, 870–75.

Williamson, O. (1979). Assessing vertical market restrictions: antitrust ramifications of the transaction cost approach. *University of Pennsylvania Law Review*, **127**, 49–85.

Appendix Table 1

Optimal patent lives and welfare indices

η_2	η_1 = 0.25		0.50		0.75		1.00		1.25		1.50		2.00	
0.25	24	0.87	19	0.83	16	0.80	14	0.77	13	0.75	12	0.74	11	0.71
	40	0.72	28	0.66	22	0.61	19	0.58	17	0.55	16	0.52	14	0.48
0.50	20	0.84	17	0.81	15	0.78	13	0.76	12	0.74	12	0.73	11	0.71
	40	0.67	28	0.62	22	0.58	19	0.55	17	0.52	16	0.50	14	0.46
0.75	18	0.82	16	0.79	14	0.77	13	0.75	12	0.73	11	0.72	11	0.70
	40	0.63	28	0.59	22	0.55	19	0.52	17	0.50	16	0.48	14	0.45
1.00	17	0.80	15	0.77	13	0.75	12	0.74	12	0.72	11	0.71	10	0.69
	40	0.60	28	0.56	22	0.53	19	0.50	17	0.48	16	0.46	14	0.43
1.25	15	0.78	14	0.76	13	0.74	12	0.73	11	0.72	11	0.71	10	0.69
	40	0.56	28	0.53	22	0.50	19	0.48	17	0.46	16	0.44	14	0.42
1.50	15	0.77	13	0.75	12	0.73	12	0.72	11	0.71	11	0.70	10	0.68
	40	0.54	28	0.50	22	0.48	19	0.46	17	0.44	16	0.43	14	0.40
2.00	13	0.74	12	0.73	12	0.72	11	0.71	11	0.70	10	0.69	10	0.67
	40	0.49	28	0.46	22	0.44	19	0.43	17	0.41	16	0.40	14	0.38

Key:

T^*_{PD}	I_{PD}
T^*_U	I_U

(T^*_{PD}, T^*_U) are optimal patent lives under discriminatory and uniform pricing respectively.
(I_{PD}, I_U) are the welfare indices (fraction of first-best attained) under the two pricing schemes.