23

Patents and Patent Policy

In 1769, an English inventor, Richard Arkwright, patented a spinning frame that would revolutionize the production of cotton cloth. Two years later, in 1771, Englishman James Hargreaves introduced another invention, the spinning jenny. With these inventions, Britain entered the Industrial Revolution. Equally important, the inventions allowed Arkwright and Hargreaves to establish a commanding position in the production of cloths and, more generally, textile products. This allowed the inventors to reap large profits and to sell at a high price in the American colonies even after these became independent states.

The British energetically protected their monopoly position. Westbound ships out of London were searched thoroughly to make sure that no passenger was a former Arkwright or Hargeaves employee or had a copy of the design plans for the Arkwright–Hargreaves machines that firms outside of Britain might copy. Such restrictions along with the high textile price for British textiles vexed many Americans. Consumers did not like paying the monopoly prices and firms were eager to get some version of the machines that would permit them to compete with the British producers. Some firms offered "bounties" for English apprentices who would be able to obtain the necessary information. Finally, in 1789, an enterprising young Englishman and former Arkwright partner, Samuel Slater, responded to just such a bounty offer. After completely memorizing the engineering details of the Arkwright-Hargreaves machines, he disguised himself as a common laborer and set sail for America. Shortly thereafter, Slater arrived in Pawtucket, Massachusetts and established the first of many New England textile mills consolidating the region's manufacturing base and finally breaking the British monopoly.

The issues raised by Slater's entrepreneurship (what some might call theft) lie at the heart of this chapter. The central question is how strongly a firm's innovation should be protected from imitative competition. On the one hand, information about an innovation is a public good so that once the information is produced, efficiency requires that access to this information, i.e., new production techniques and new products, should be unrestricted to prevent the rise of monopoly. On the other hand, if the government does not protect innovators against imitation, there may be little incentive to do the hard work that led to the invention in the first place.

The patent system was designed to create incentives for innovative activity. Patents and copyrights confer ownership to new inventions, new designs, and new creative works. In turn, those property rights permit innovators to restrict the use of their ideas just as the British restricted the flow of information on their textile technology. The patent holder can act as a monopolist regarding its discovery and earn a monopoly profit as a result. Yet while that profit may create an incentive to undertake R&D efforts, the monopoly that generates the profit reduces the total surplus below what it could be given that the invention has occurred.

Getting this balance right is not easy. We can imagine just how much less productive the economy would be if the science behind electric lighting, the aerodynamics of airplanes, and semiconductors had never been developed. However, production would also suffer were those same technologies not widely available to all firms. At some point, policy must shift from a stance of protecting innovators from imitation to one of permitting the use of the innovation on as wide a basis as possible. The sixty-four-million-dollar question is, exactly where does that point arise? When has protection of the innovator extended sufficiently far that we ought to start thinking about protection of consumers?

The issue as to how far patent rights should extend has two dimensions. First, what is the length of time for which any patent rights ought to extend? Secondly, to what range of products should the patent apply? Should the developer of a new AIDS treatment based on a special combination of protease inhibitors be protected against a rival's later development of an alternative AIDS treatment based on a different combination of protease inhibitors? What about a new AIDS treatment that is not based on protease inhibitors? Or what if a protease inhibitor treatment originally created as a treatment for AIDS is now applied as a treatment for multiple sclerosis? These issues—typically referred to as patent length and patent breadth—are the central questions in patent policy.

23.1 OPTIMAL PATENT LENGTH

Current patent law establishes a patent duration that varies from country to country. In the United Kingdom and the United States patent law grants protection for twenty years from the date of filing the application. In both countries it is up to the patent holder to ensure that the patent is renewed during its life and to ensure that the patent is not infringed.

Economic theory can provide some insights as to whether what duration makes sense. The key is to find a balance between the innovator's ability to earn a return on its R&D investment and the benefits that will accrue to consumers once the patent expires and competition emerges. The basic model, which is due to Nordhaus (1969), is presented below.

Imagine a competitive industry in which each firm is pursuing a nondrastic innovation. Innovative efforts incur costs. Each firm's unit operating cost is currently c. If a firm invests in R&D at some intensity x, it expects to reduce its unit operating costs from c to $c - x$. The cost of undertaking R&D at intensity x is $r(x)$. We assume that such costs rise as the level of research intensity increases and that they do so at an increasing rate. Formally, this means that $dr(x)/dx > 0$ and $d^2r(x)/dx^2 > 0$. Thus, R&D is expensive to do and exhibits decreasing returns in that a doubling of research intensity will give less than double the reduction in operating costs.

Our assumption of a competitive market implies that price equals marginal cost, which means that the initial market price is c and that the output level is Q_0^C. This is shown in Figure 23.1. A successful innovator will be able either to produce at the lower unit cost of $c - x$ and drive out all its rivals by setting a price just one penny less than the current price, or to license its discovery to its competitors for a fee of $c - x$ per unit produced. Either way, the current market price and volume of output remain unchanged. The innovator, however, will earn a profit equal to area A in Figure 23.1. Assuming that the life of the innovator's patent is T years, this profit will last for T years as well.

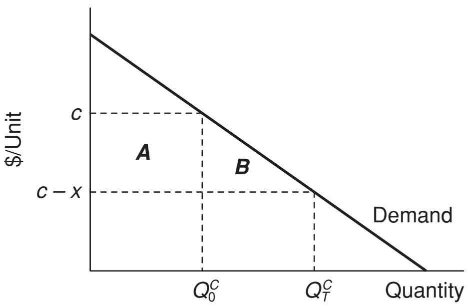

Figure 23.1 Innovation gains during period of patent protection (T years) and after patent protection
The innovator receives profit of area A for the T years that the patent is in effect. When the patent expires, competition lowers the price to $c - x$. Consumers gain the former profits A and also area B as consumer surplus. However, only the area B is a net increase in welfare.

When the patent expires, all firms will have access to the technology for free. Competition will reduce the price to $c - x$, and output will expand to Q_T^C. The profit that the innovator used to earn becomes consumer surplus. This is simply a transfer from a producer to consumers and so does not reflect a net gain. However, the expansion of industry output to the higher level Q_T^C does bring such a net benefit by virtue of the additional consumer surplus this generates. This additional surplus is shown in Figure 23.1 as area B.

The longer the duration of the patent (the higher is T), the longer is the time over which the innovator earns the profit A and the greater is the innovator's incentive to do costly R&D. Denote the per period profit flow to the innovator (area A in Figure 23.1) as $\pi^m(x; T)$ and the discount factor as R. The present value of the innovator's profit from R&D is[1]

$$V_i(x; T) = \sum_{t=0}^{T-1} R^t \pi^m(x; T) = \frac{1 - R^T}{1 - R} \pi^m(x; T) \tag{23.1}$$

Therefore, the R&D has a net value to the innovator of

$$V_i(x; T) - r(x) \tag{23.2}$$

For a given value of T chosen by the patent office, the innovator will select a level of R&D activity, $x^*(T)$, that maximizes this expression. This choice will just balance the marginal gain of additional discounted profit against the marginal cost of doing more R&D work.

Of course, a rational patent office recognizes that its choice of patent life T affects the firm's choice of R&D effort. We suppose that the patent office can work out this relationship precisely. In other words, the patent office can determine the innovator's profit-maximizing research intensity, $x^*(T)$, as a function of T. To choose T optimally, the patent office will wish to pick the patent duration that maximizes the net social gain to both consumers and producers given how firms choose their research intensities. Let us denote by $cs(x; T)$ the per-period increase in social surplus that the innovation generates once it becomes freely

[1] This result uses the following equation in calculating discounted value. Assume that a sum A is to be received each period for T periods, and recall from section 2.2 Chapter 2 that $R = (1 + r)^{-1}$ where r is the interest rate. Then the discounted value of these cash flows is $S = A + RA + R^2A + R^3A + \ldots + R^{T-1}A = A(1 + R + R^2 + \ldots + R^{T-1}) = A(1 - R^T)/(1 - R)$.

available, which as we have seen is the area $A + B$ in Figure 23.1. The present value of this increase in surplus is then

$$CS(x; T) = \sum_{t=T}^{\infty} R^t cs(x; T) = \frac{R^T}{1-R} cs(x; T) \tag{23.3}$$

The total net social surplus from the innovation is

$$NS(x^*(T); T) = V_i(x^*(T); T) + CS(x^*(T); T) - r(x^*(T)) \tag{23.4}$$

and the objective of the patent office is to choose the patent duration T^* that maximizes this net surplus. This is a complicated expression but we can develop an intuitive argument to support a very important proposition, namely, that *the optimal patent duration is finite*.

To see why, note that as the patent office initially increases patent duration it induces greater R&D effort and, at first, a greater discounted net surplus to producers and consumers. If patent duration is zero, the returns to an innovator are also zero since the results of the innovation will be imitated immediately. Accordingly, there will be no R&D and no change in the social surplus. If we now increase the patent length to a value $T > 0$, we will induce some innovation and, thereby, some increase in the total surplus. Beyond some point, however, continued increases in T will reduce net social surplus even though they lead to more R&D and therefore greater reductions in production cost. Two forces work to limit the optimal value of T. The first is our assumption of diminishing returns to R&D activity. Because it becomes progressively more expensive to lower production costs, it will take progressively greater increases in T to achieve a given additional cost saving. The second force limiting optimal patent duration is the fact of discounting. The consumer benefits shown as area B in Figure 23.1 will not be realized until after the patent expires. If the patent office chooses a very long duration time T the present value of those benefits will be very small indeed.

This is particularly important since it has sometimes been argued that innovation should be granted patent protection forever. Such a long patent duration puts far too heavy a value on the monopoly profits that patent protection generates and too little consideration on the additional consumer surplus that will emerge only after the patent protection has expired.[2]

Practice Problem 23.1

Let the inverse demand function for a particular product be $P = 100 - Q$, and let it be provided by a group of competitive firms, each with an identical marginal (and average) cost of \$70 per unit.

a. Show that the current market output and price are, respectively, $Q = 30$ and $P = \$70$.
b. Imagine that one firm can conduct R&D at a pace x, at a cost of $r(x) = 15x^2$. Let the interest rate, r, be 10 percent so that the discount factor, R, is 0.9091. Show that a patent length of 25 years will induce the firm to pursue R&D at a level of approximately $x = 10$. Note that if $x = 10$, the firm's research activity will reduce the unit cost from \$70 to \$60.
c. Would the firm's R&D effort increase or decrease if patent duration was reduced to 20 years?
d. Would total social welfare increase or decrease if patent duration was reduced to 20 years?

---

[2] Author Mark Helprin has similarly argued for an infinite copyright for creative works ("A Great Idea Lives Forever, Shouldn't Its Copyright?," *New York Times*, May 20, 2007. Note, the argument for an infinite patent life is moot if there is continual innovation that effectively limits the economic life of any one patent.

23.2 OPTIMAL PATENT BREADTH

The question of the optimal patent breadth is trickier than that of patent length, mainly because there is no universally accepted measure of breadth comparable to time as a measure of duration. Conceptually, the idea is to set a minimum amount by which a new innovation must differ from an existing process (or good) in order for the new one either to avoid infringement on an existing patent or to be itself patentable. The larger this required minimal degree of difference, the more difficult it is for other firms to "invent around" the patent and to cut into the inventor's profit. We could in principle work out the optimal patent breadth just as we worked out the optimal patent length. But the lack of a clear method for measuring breadth makes implementing this plan very difficult.

This lack of precision is reflected in the language of the patent office. Each application for a patent is required to specify all the "related" existing patents and to indicate not only how the patent being applied for is a discovery distinct from those already patented, but also to show that the discovery is "novel, nonobvious, useful." Such language leaves the patent office a lot of discretion regarding how it will rule in any particular case.

What makes the question of the optimal patent breadth even more difficult is that it cannot be divorced from the question of optimal duration. Patent policy must set both dimensions of patent protection. Typically, this amounts to choosing between a system in which patents should have a short duration but a broad coverage, the "short and fat" approach, or a long duration combined with a very narrow coverage, the "long and thin" solution. As always, these choices involve balancing the need to maintain the incentive to innovate against the need to distribute the benefits of innovation as widely as possible.

Unfortunately, introducing a second dimension of breadth into patent policy is more difficult than one might suspect, largely because breadth is less obviously defined. Consider for instance the analysis of Gilbert and Shapiro (1990). They define patent breadth in term of the extent to which the patent-holder can charge a price above marginal cost. Broader patents decrease consumer substitute options and permit a higher price-cost margin. This margin of course is the source of the patent-holder's profit. Suppose that we know the desired innovative effort level x, and hence the cost $r(x)$ necessary to achieve that effort. The trick then is to do this with a patent design that produces the necessary (discounted) profit at the lowest possible social cost. That is, we may frame the objective as choosing patent breadth and length so as to minimize the deadweight loss per unit of innovator profit subject to that profit level being sufficient to undertake the desired inventive activity.

Given the social objective and their definition of patent breadth, Gilbert and Shapiro (1990) then demonstrate that the optimal patent is to have very narrow but infinitely long patents. Why? The underlying intuition is as follows. If we think of time as a sequence of equally long intervals, then each interval may be thought of as a separate market. A standard condition for welfare maximization is that it should not be possible to raise welfare by shifting production from one market to another, i.e., the net marginal value of an extra unit should be the same in each market or, in our case, in each period. A finite patent though will not typically satisfy this condition. During the patent life, the price will be high due to the patent-holder's monopoly power. Once the patent has expired, however, the price will fall to marginal cost. The only way to avoid this discontinuity is to keep the price above marginal cost in all markets, i.e., for all time periods into the infinite future, while limiting the accompanying distortion that this brings by restricting the patent's breadth so that price is just a enough above cost permanently that the necessary profit level is achieved. In other words, optimal patent policy induces small but equal price distortion for many periods rather that a few large distortions in some periods and none in others.

However, the Gilbert and Shapiro (1990) approach is not the only way to model patent breadth. Klemperer (1990) for example offers an alternative that relates patent breadth more directly to product differentiation. If we think of a Hotelling line segment of finite length, Klemperer's view is that a useful definition of patent breadth is the fraction of the line segment that is covered by the patent. Again assuming that the goal is to minimize the ratio of social cost to innovator profit subject to covering desired innovation costs, Klemperer (1990) then shows that there will be cases in which optimal patent design is just the opposite of that implied by the Gilbert and Shapiro analysis. That is, it will now often be the case that the best patent design is one of broad patents that are short-lived.

To understand Klemperer's (1990) argument, consider a simple example. Suppose that the new good costs nothing to produce (all the costs are sunk design costs) and that there are ten potential customers for the product. Assume further that each customer values the good at \$10. If there were no other substitutes available, the monopolist firm would then simply set a price of \$10, sell one unit to each of the ten consumers and claim all of the \$100 surplus from the market. There would be no consumer surplus but also no welfare loss. Total output would also have been ten had the monopolist priced at marginal cost (= 0).

Now suppose that the transport cost of buying an alternative legal substitute is different for each of the ten consumers. Specifically, let one consumer incur a transport cost of \$1 per unit of distance the substitute is from the product; a second incur a transport cost \$2 per unit; and so on. In this setting, patent breadth w is interpreted as how far consumers have to travel to obtain a legal alternative brand. A very wide breadth or high value for w effectively puts one back in the setting of no alternatives. Hence, if w is very broad, the market outcome will again be a price again of \$10, and there will be no deadweight loss. Now, however, consider what happens if we limit the patent width to $w = 1$ (or just a bit less). In this case, at any price $p \geq \$1$, the patent holder will lose some customers. At a price of \$1, he will lose one client. At a price of \$2, he will lose a second and so on. His best bet (assuming whole-dollar prices) is then to set a price of \$5, in which case he will sell five units and earn profit of \$25. Now there is a deadweight loss. Real resources are being used to produce the less desired substitutes into which consumers are shifting. Consumers are now incurring transport costs, as well. Accordingly, the narrower patent results in the greater deadweight loss. Its length should then be set at the minimum necessary to achieve the desired innovative expenditure.

As noted, Klemperer's (1990) analysis also yields conditions under which a narrower but longer-lived patent is preferred. This occurs for example when, unlike the case above, it is the transport cost that is the same for all and the valuation that varies across consumers. However, the crucial result is that when consumer variation primarily reflects differences in transport cost or strength of preference for the brand of the new good and not in the basic valuation of that good, broad patents of relatively short derivation are preferred.

Gallini (1992) provides a further reason why short-lived broad patents may be best. She makes the important point that imitators can often get around patent protection if they spend enough money to imitate the product without infringement. They will be particularly encouraged to do so when patents are long because otherwise, entry into the market will be greatly delayed. When patents are short, imitation is less attractive because firms now find it cheaper simply to wait for the patent to expire than to engage in costly efforts to imitate legally now. In other words, Gallini (1992, 2002) makes the important point that costly imitation efforts need also to be accounted for in considering the welfare effects. If these imitation costs are sizeable, then broad but short-lived patents are preferable.

Denicolò (1996) synthesizes many of these features in a framework that also incorporates the extent of market competition. He finds that "Loosely speaking, the less efficient is the type of competition prevailing in the product market, the more likely it is that broad and short patents are socially optimal" (p. 264). By "efficient," Denicolò means roughly the extent to which competition drives firms close to the competitive ideal. Denicolò's statement implies that markets in which firms have a greater degree of monopoly power will do best with the "short and fat" approach, while markets characterized by a good bit of competition will do best with patents that are "long and thin."

As a policy recommendation, a major drawback of Denicolò's proposal is that it seems to suggest applying different standards to different innovators depending on the structure of the innovator's basic industry. In reality, the rule of law cannot be applied so selectively without risking serious inconsistency. Even apart from that consideration, there is a further difficulty in implementing any of the proposed standards. The problem again is that it is not always easy to make the concept of breadth operational. We do not have an easy way to translate real markets into a spatial representation and no obvious measure of distance. Indeed, as Scotchmer (2004) has noted, Klemperer's (1990) horizontal concept of breadth is itself too limiting. There is also a vertical component reflecting how much better (or how much worse) a rival's product has to be before it infringes on the patented good. Recognizing this second dimension of patent breadth makes its measurement all the more difficult from a practical perspective. While it is probably fair to say that we will not go too far wrong if we adopt a one-size-fits-all policy of granting patents with "reasonable" breadth but constrained length, precisely what this means in practice is a lot less clear.

23.3 PATENT RACES

Our discussion of market structure and innovative activity was largely motivated by Schumpeter's observation that innovation is a crucial and different kind of competition. The Schumpeter vision is one in which firms vie with each other by racing to develop new technologies or new goods and in which this sort of rivalry is potentially deadly for those who come up short. This is particularly true when innovations are eligible for patent protection. With patents, coming in first is all that matters whether one wins by several lengths or by just a nose. The first firm to discover a cure for male baldness or to engineer a successful system for producing "talking" pictures leaps far ahead of its rivals and stays there for some time by virtue of patent or copyright protection. Patent awards have a "winner-take-all" feature so that finishing second can be no better than finishing third or fourth or, for that matter, tenth.

Innovative competition can be regarded as a race in which one player's success is the other player's serious defeat. The loser of a patent race may see years of investment and hard work wiped out overnight when the rival announces its breakthrough. We now turn to some of the issues that arise when we consider the implications of a patent system for generating a race in which finishing first is all that matters. What are the consequences of such races? Do they lead to inefficient investment in R&D? Does the innovative activity generated by the race influence market structure?

Consider a patent race between two firms that can choose to invest in research with a view to developing a new product. The first to make the breakthrough wins the race and files a patent giving that firm exclusive rights to its invention. This is what gives the race its winner-take-all aspect. The loser walks away less than empty-handed, having expended resources on R&D with no return.

9781405176323_4_023.qxd 10/19/07 8:16 PM Page 611

Consider two firms, BMI and ECN, who are Cournot competitors and who are both considering doing the R&D necessary to create a new product. They each estimate that if the innovation is successful they can produce this new product at a marginal cost of $\underline{c}$ and that demand for the new good is $P = A - BQ$. They are also confident that the new product is a sufficiently radical departure that it will have a negligible impact on their existing business and so will not affect their existing profits—that is, there is no replacement effect.

The R&D effort by each firm requires a research division that will cost a fixed sum, K. This sum covers both the costs of research and of development if the research is successful and, once sunk, can never be recaptured. Given that such a division is established, the probability of a successful innovation is ρ. If only one firm is successful in its R&D efforts, we assume that the innovation is protected from imitation, perhaps by a patent or by some other means. If both are successful simultaneously, we assume that both firms can make the new product, in which case they will be involved in Cournot competition in selling it. To keep matters reasonably simple, we assume that both firms discount the future heavily—that is, the interest rate r is so large that the discount factor $R \approx 0$.

In order to identify the incentives each firm has to establish the research division, we need to identify their profits with and without a successful innovation. If neither firm attempts to develop the new product, neither firm will enter this new market. As a result, each will earn zero profit in this new market. Conversely, if both firms undertake R&D and are successful in making the innovation, we know that their profits, ignoring the cost of establishing the R&D division, will be the Cournot duopoly profits[3] at marginal costs $\underline{c}$:

$$\pi_b = \pi_e = \frac{(A - \underline{c})^2}{9B} \tag{23.5}$$

If one firm, say BMI, is successful in its R&D efforts but ECN is not, then BMI will be a monopolist in the new product market earning the monopoly profits. ECN will earn nothing from the new market. The profits of the two firms in this case, again ignoring the costs of establishing the R&D division, will become

$$\pi_b = \frac{(A - \underline{c})^2}{4B} \tag{23.6}$$

Of course, if ECN is successful but BMI is not, these profits will be reversed.

We can now calculate the expected profit for each firm depending on whether or not it establishes a research division. If neither firm sets up such a division, neither will innovate and each will earn zero profit in this new market. Now consider the expected profit if only one firm, say BMI, establishes an R&D division. For BMI expected profit is made up of two components:

1. profit if the R&D division is unsuccessful, which is zero and occurs with probability $(1 - \rho)$;
2. profit if the R&D division is successful, which is the monopoly profit $(A - \underline{c})^2/4B$ and occurs with probability ρ.

[3] See section 9.4 in Chapter 9 for a derivation of Cournot profit.

As a result, the expected profit of BMI if it is the only firm to establish an R&D division is

$$\pi_b = \rho\frac{(A - \underline{c})^2}{4B} - K \tag{23.7}$$

Of course, the expected profit of ECN, given that only BMI has established an R&D division, is zero. By symmetry, we reverse these payoffs to get expected profits if ECN is instead the only firm to establish a research division.

If both firms establish R&D divisions, the expected profit to either firm is given by

1. profit if the firm's R&D division is successful and the rival's is not, which is $(A - \underline{c})^2/4B$ and occurs with probability $\rho(1 - \rho)$;
2. profit if both R&D divisions are successful, which is $(A - \underline{c})^2/9B$ and occurs with probability ρ^2.

Of course if neither firm is successful in R&D they earn nothing from the new market. This means that the expected profit of each firm, given that they both operate R&D divisions, is

$$\pi_b = \pi_e = \rho(1 - \rho)\frac{(A - \underline{c})^2}{4B} + \rho^2\frac{(A - \underline{c})^2}{9B} - K = \frac{(A - \underline{c})^2}{36B}\rho(9 - 5\rho) - K \tag{23.8}$$

Before we put these payoffs into a payoff matrix, we can do a bit of simplifying. The profit equations share a common expression, the monopoly profit, which we denote as $M = (A - \underline{c})^2/4B$. We can use this to define a parameter $S = K/M$, which is the share of the monopoly profits that are needed to establish the R&D division. With the substitution of S and M, the expected profits are summarized in the payoff matrix of Table 23.1. This matrix allows us to identify the possible Nash equilibria for this R&D game. As we shall see, these will be dependent upon the relative magnitudes of the two parameters, S and ρ.

There are three possibilities that have to be considered:

1. *Neither firm wishes to establish an R&D division*. For this to be a Nash equilibrium, the payoff to BMI, for example, from not having an R&D division, given that ECN also has no R&D division, must be greater than the expected profit from investing in R&D, again given that ECN does not. In other words, BMI expects to make more profit from the

Table 23.1 Payoff matrix for the Duopoly patent race

		BMI	
		No R&D Division	*R&D Division*
ECN	*No R&D Division*	0, 0	0, $M(\rho - S)$
	R&D Division	$M(\rho - S)$, 0	$M\left(\frac{\rho(9-5\rho)}{9} - S\right), M\left(\frac{\rho(9-5\rho)}{9} - S\right)$

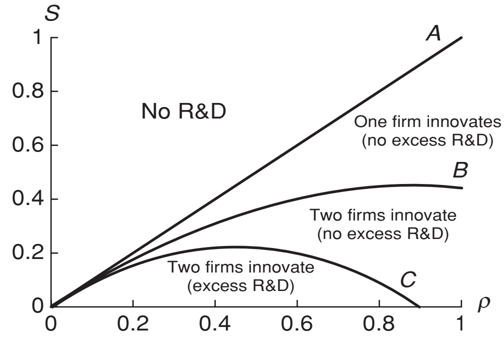

Figure 23.2 A patent race with a duopoly

strategy combination (No R&D, No R&D) than from the combination (No R&D, R&D). This requires that $M(\rho - S) < 0$, which implies that $S > \rho$, the probability of success is less than the fraction of monopoly profit required to fund the R&D. This expression is illustrated by the line 0A in Figure 23.2. All parameter combinations above 0A give the Nash equilibrium (No R&D, No R&D).

2. *Only one firm wishes to establish an R&D division*. Assume that the firm that establishes the R&D division is BMI. Then for the strategy (No R&D, R&D) to be a Nash equilibrium, two conditions must be satisfied:

(a) BMI expects its expenditure on R&D to be profitable, given that ECN is not investing in R&D—that is, BMI expects to make more profit from the strategy combination (R&D, No R&D) than from the strategy combination (No R&D, No R&D). This is just the opposite of the expression derived in part 1. It requires that $S < \rho$.

(b) ECN does not expect its expenditure on R&D to be profitable, given that BMI is investing in R&D—that is, ECN prefers the strategy combination (No R&D, R&D) to (R&D, R&D). For this to be the case, the following must be true:

$$M\left(\frac{\rho(9-5\rho)}{9} - S\right) \text{ which requires that } S > \frac{\rho(9-5\rho)}{9}$$

This relationship is illustrated by the curve 0B in Figure 23.2. All parameter combinations that lie between 0A and 0B are such that only one of the firms will establish an R&D division.

3. *Both Firms Wish to Establish an R&D Division*. For this to be a Nash equilibrium, the payoff to, for example, BMI from having an R&D division, given that ECN also has an R&D division, must be greater than the expected profit from not investing in R&D, again given that ECN does. In other words, BMI expects to make more profit from the strategy combination (R&D, R&D) than from the strategy combination (R&D, No R&D). For this to be the case we must have that

$$M\left(\frac{\rho(9-5\rho)}{9} - S\right) > 0 \text{ which requires that } S < \frac{\rho(9-5\rho)}{9}$$

Of course, exactly the same condition guarantees that ECN prefers the strategy combination (R&D, R&D) to the strategy combination (No R&D, R&D). Thus, all parameter combinations below 0B are such that both firms will establish an R&D division.

One question about patent races is whether the potential profit from successful innovation can lead the two firms to overinvest in R&D. Neither of the firms will establish an R&D division unless this division is expected to be profitable. For the strategies (R&D, No R&D), (No R&D, R&D), and (R&D, R&D) to be equilibria, they must each give positive expected profits to the two firms. This tells us that no equilibrium in which only one firm invests in R&D is characterized by "excessive" R&D in the sense that the firms would be better off without the R&D. The question that is left is whether there is "too much" R&D when both firms establish R&D divisions. Are there situations in which the strategy combination (R&D, R&D) is a Nash equilibrium but generates less aggregate profit than the strategy combinations (R&D, No R&D) or (No R&D, R&D)? For this to be the case it must be that

$$2M\left(\frac{\rho(9-5\rho)}{9}-S\right)<M(\rho-S) \text{ which requires that } S>\frac{\rho(9-10\rho)}{9}$$

This is illustrated by the curve 0C in Figure 23.2. All parameter combinations between 0C and 0B lead to excessive R&D as the two firms race to be first to discover and introduce the new product.

Our simple model delineates three distinct possibilities. First, neither firm will invest in R&D unless it is expected to be profitable. Hence R&D must have a reasonably low cost relative to the monopoly profits that it might generate (low S), or a reasonably high probability of success. Second, for any given probability of success, a larger number of firms will establish R&D facilities when there is a lower cost of R&D relative to the profit the innovation is expected to generate. Thus, for any given probability of success ρ, the equilibrium number of firms with R&D divisions increases from zero to one and finally to two as S is reduced. Third, there is an intermediate range of values for the cost of R&D in which there is excessive R&D in that both firms establish R&D divisions although this reduces their aggregate profits. In this range, the lure of profit from innovation involves the firms in a competitive R&D race that they would be better to avoid.

So far, we have only considered the gain that research brings in terms of the expected profit of the two firms. From a public policy perspective, however, increased profit is not the only potential benefit of innovation. We should also consider the gain in consumer surplus that development of this new product will generate. While we have just shown that the level of R&D activity can be excessive from the viewpoint of the firms' combined profits, we have not demonstrated that this is the case when viewed with the objective of maximizing the total gain of profit plus consumer surplus. The R&D which seems excessive to the firms may still be worthwhile to society overall if the additional consumer surplus more than offsets the reduction in aggregate profit. However, R&D can be excessive even when evaluated with this broader criterion, (see Practice Problem 23.2). The patent race can lead both firms to establish research divisions even when the total cost of such divisions is not justified by the sum of expected producer and consumer surplus.

Even more interesting is that we can easily show that the possibility of too little R&D —as judged from a social welfare criterion—is quite real. Consider the case in which neither firm does any R&D. As we know, this happens when $S > \rho$. Suppose that although this inequality holds, S is so close to ρ that one firm could almost expect to break even if it pursued the innovation (and its rival did not). If the sale of the product generates any significant consumer surplus at all, then it is socially desirable that the research takes place. The value of the expected consumer surplus more than provides the extra funds needed to ensure that the innovator breaks even. Yet, in the absence of some sort of government intervention, the fact that $S > \rho$ means that no such R&D efforts will occur.[4]

23.2 Practice Problem

Consider the BMI-ECN example of a patent race. Assume that demand for the new good is $P = 100 - 2Q$, and that each firm believes that it will be able to produce this good at a unit cost of $c = \$50$. Assume further that the discount factor R is so small that each firm cares only about the one-period profit it will make. (Alternatively, assume that one period is of a very long duration, say, 30 years or more.) The probability that such a lab will be successful and actually produce a discovery is $\rho = 0.8$.

a. Show that if one firm is successful in introducing the product, it will have a monopoly price of \$75, sell 12.5 units, and earn monopoly profits (before paying for the research) of $M = \$312.50$. Show also that consumer surplus is \$156.25.
b. Show that if each firm sets up a lab and if each lab is successful, the Cournot equilibrium output for each firm will be 8.33 units, the price will be \$66.67, and each firm will earn a profit (before paying for the research) of \$138.89. Confirm that consumer surplus is now \$277.78.
c. Now show that the expected profit (before paying for the research) to BMI (or ECN) if it is the only firm to establish an R&D division is \$250 while the expected profit to each firm if they both establish R&D divisions (again before paying for the research) is \$138.89. Use these results to construct the payoff matrix for this case, now including the cost, K, of establishing an R&D division.
d. Show that if K, the cost of setting up the research lab, is such that $K > \$250$, neither firm will set up a lab, while if $K < \$138.89$, both firms will set up a lab.
e. Show that expected social surplus ignoring research costs if one firm establishes a research lab is \$375, and if two research labs are established, is \$505.56. Hence, show that the second lab is socially desirable only if $K < \$130.56$.

We have focused on the risk that patent races may yield either too much or too little R&D investment. Another issue to consider is the possibility that patent races will lead firms to pursue more risky innovations. The intuition behind this argument can be illustrated fairly simply. Suppose that firms can choose to invest either in a relatively safe R&D route that has an expected discovery time uniformly distributed between one and three years or a more risky route that has an expected time of discovery uniformly distributed between zero and four years. Both discoveries are equally costly, and both are expected to become redundant or worthless in five years' time. We will also assume that each discovery will generate the same profit of \$1 million per period during the time that it is utilized and is protected from imitation by a patent.

Since the expected date of discovery is the same, namely two years for both routes, then assuming neither firm had any competition, a risk-neutral firm considering them would be

[4] See Reinganum (1989) for a masterful survey of patent races and the timing of innovation, including the consequences for social welfare.

indifferent between the two options, and a risk-averse firm would go for the less risky route. However, when firms are involved in a patent race, competition between the firms may lead them to choose the more variable or risky route in which success can come anytime between zero and four years.[5] The reason is that again, when innovation is protected from imitation, all that matters is winning the race. The second-place firm loses the same amount no matter how close it is behind the winner. In our example, if my rival chooses the less risky R&D route, I have an incentive to choose the more risky route, since this offers the possibility of success and a quick victory right away. Similarly, if my rival adopts the risky strategy, I can see that unless I do the same there is a real possibility that I will be left behind in the race. Of course, my rival can work out all this too. The result is that both of us choose the more risky route.

23.4 MONOPOLY POWER AND "SLEEPING PATENTS"

Another way in which the patent system and innovative competition can interact to affect market structure is through "sleeping patents." Many students at first find it puzzling that a firm will hold a large number of patents all related to the same process or product, and some patents are never acted upon. (Return to Table 23.1 for some evidence on this point.) What possible reason can a firm have to earn patent rights to products and processes that it never uses, that is, what could be the rationale for a firm to create and hold what is called a "sleeping patent"?

The motivation behind a sleeping patent is to create a buffer of protection for the monopoly profits generated by the truly valuable patent. Legal history and economic analysis have both documented that the protection granted by a single patent is often very limited. Edwin Mansfield and his associates (1981) found in a study of 48 patented new products, that 60 percent were imitated within 4 years after their introduction. Firms often can and do "invent around" patent protection, as we discussed earlier in the case of pharmaceuticals. Frequently, there are several technical solutions to a particular problem such as is the case for the production of the whitening agent, titanium oxide. Each such alternative production technique is a threat to the firm holding a patent on a particular process or product. Hence, by patenting as many of these alternatives as it can, a firm increases the protection it has in using whatever process it actually decides upon.

Suppose, for example, that market demand is given by $P = 100 - Q$ and that the incumbent firm has a proprietary technology with a constant marginal cost of $c_I = \$20$. The firm has a patent that protects its technology. Let us also suppose that this technology is so efficient that entry is not possible, and thus the incumbent is free to set the monopoly price and earn a monopoly profit each period of $\pi^m(c_I = 20)$. To be precise, the monopolist will sell 40 units at a price of \$60 and earn a profit of \$1,600.

Assume now that there is also an alternative technology that the monopolist has discovered, which permits production at the higher constant marginal cost of $c_E = \$30$. Clearly, the monopolist has no incentive to switch to this technology. However, if \$30 is a low enough unit cost so that another firm could acquire this technology and enter the industry, then the incumbent's current monopoly would be eroded. The entrant would either be the high-cost member of a Cournot duopoly or, if Bertrand competition prevailed, the entrant's cost based on using this alternative technology would at least establish a clear upper bound of \$30 on the incumbent's price—one that we know (by construction of the example) is below the incumbent's current monopoly price.

[5] This type of case is discussed in Klette and de Meza (1986).

It is easy to see that the incumbent has an incentive to patent the higher-cost technology as well as the lower-cost one, even though it will never use this alternative, higher-cost technology. By acquiring this patent and letting it lie dormant or sleep, the incumbent strengthens his hold on his monopoly position. The question that we need to ask is whether the incumbent's incentive to acquire the higher-cost technology is so strong that it actually exceeds the incentive of the entrant to acquire the technology and enter.

The surprising answer is yes. Acquiring a patent to the high-cost technology is worth more to the incumbent monopolist than to his potential rival. This is obvious in the case of Bertrand competition. In that case, the rival's entry with a high unit cost of \$30 would provoke a price war in which the incumbent would have to lower his price from its current monopoly level to the marginal cost of the entrant, namely, \$30. Of course, when this happens, the entrant earns nothing. The incumbent, however, because of his lower cost, will still earn \$30 − \$20 = \$10 per unit. At this price, the incumbent will now sell \$70 units and earn a profit of \$700. This is less than what he earned previously but still better than nothing. From this it should be clear why the monopolist will place a greater value on discovering the alternative process than would the entrant. Under Bertrand competition, the entrant will never earn any money with this innovation. Hence, for the entrant, discovering the process is worthless. Yet, even though the entrant cannot make money with this higher-cost process, it can put pressure on the incumbent. Specifically, discovery of the process by the entrant imposes a ceiling of \$30 on the incumbent's price. Hence, it is worth something to the monopolist to acquire the process first and thereby preclude the imposition of this price cap altogether.

The same basic result holds in a Cournot model. The gain to the monopolist from acquiring the second, sleeping patent on the high-cost process is the profit earned as a monopoly firm using the low-cost technology, $\pi^m(c_I = 20) = \$1,600$, less the profit earned as a duopoly firm with the low-cost technology facing a rival with the high-cost technology, or $\pi_I^d(c_I, c_E) = \pi_I^d(20, 30) = \$900$. So, the total net gain to the monopolist is $\pi^m(c_I = 20) - \pi_I^d(20, 30) = \$700$. In contrast, the gain to the potential entrant is the profit earned as the high-cost firm in a duopoly, $\pi_E^d(c_I, c_E) = \pi_E^d(20, 30) = \$400$ less his current profit, taken to be zero.[6] Therefore, the entrant's net gain from developing the technology is \$400. Hence, in the Cournot case, the gain to the monopolist incumbent exceeds that of the potential entrant.

The reason that the incumbent monopolist is more willing than a potential rival to develop the high-cost process and patent it is now familiar. It is because the monopolist has a lot more at stake. If he wins the race, he gets to keep his current monopoly position. If the entrant wins the race, the best the entrant can hope for is to be the high-cost member of a duopoly. The incumbent acquires the patent on the high-cost process to make sure that nobody else will use it. Viewed in this light, acquiring "sleeping patents" amounts to broadening the patent's width.

The data in Table 23.2 provide some empirical support for this proposition. These data are drawn from a PatVal–EU survey of 9,017 patents issued by the European Patent Office between 1993 and 1997 to individuals located in France, Germany, Italy, the Netherlands, Spain and the United Kingdom.[7] The survey asked the inventors to rate the importance that they put on different motives for patenting.

[6] We leave it to the reader to show that the Cournot equilibrium has the incumbent producing 30 units and the entrant producing 20 units implying the profit amounts we have used here.

[7] For a detailed description and analysis of these data see Giuri and Mariani (2005) "Everything you Always Wanted to Know about Inventors (But Never Asked): Evidence from the PatVal-EU Survey," available at http://www.lem.sssup.it/WPLem/files/2005-20.pdf.

Table 23.2 Patent use by inventor's employer (%)

	Internal use	*Licensing*	*Cross-licensing*	*Licensing and use*	*Blocking competitors*	*Sleeping patents*
Large companies	50.0	3.0	3.0	3.2	21.7	19.1
Medium-sized companies	65.6	5.4	1.2	3.6	13.9	10.3
Small Companies	55.8	15.0	3.9	6.9	9.6	8.8

Source: Giuri et al. (2005, p. 20)

In Table 23.2 "blocking competitors" refers to sleeping patents that are used specifically for the strategic reasons we have been discussing in this section (what the researchers term sleeping patents are patents that, according to the respondents, were not used for any of the other six purposes identified in Table 23.2). As can be seen, large companies used fewer of their patents and used a higher proportion of patents for blocking purposes than did medium sized or small companies, consistent with our "protecting monopoly power" analysis.

Reality Checkpoint
The Light That Failed

Carlile Stevens, an inventor, and Bill Alling, his business partner, endured a legal odyssey of longer duration than the fabled ten years of wandering suffered by Ulysses. The two men met in 1969 when they both worked for Singer Corporation. Mr. Stevens was a physicist then employed by Singer on a project to make traffic lights brighter. In the course of his work, he hit upon an idea for a solid-state electronic ballast to be used in fluorescent lamps. At that time, all ballasts were magnetic ones that wore out quickly, causing the fluorescent lights first to "hum" incessantly and, eventually, to fail. Mr. Stevens got together with Mr. Alling, who was then in Singer's marketing department. The two went out on their own and persuaded 175 investors to put up $3.6 million in seed money. By the late 1970s, they patented their product, which was shown not only to outlast the existing magnetic ballasts but also to offer a 50 to 70 percent improvement in energy efficiency. In 1982, Universal Manufacturing Corporation, which owned, Magnetek, one of the two major manufacturers of magnetic ballasts, approached Stevens and Alling about acquiring the rights to their new technology. The two entrepreneurs agreed in return for a share of the royalties Magnetek earned from licensing the process. By 1984, however, Stevens and Alling had received no royalties and became convinced that Magnetek had no intention of putting their new ballast on the market. Meanwhile, Motorola, a firm that had originally approached Stevens and Alling before Magnetek, was able to invent around the patent and introduce its own solid-state ballast.

Stevens and Alling filed a suit against Universal and Magnetek. They argued that Magnetek acquired the rights to the invention with no intention to use it but simply to protect their own magnetic ballast product. In 1996, two juries ruled in favor of Stevens and Alling and awarded them $96 million in damages. The awards were upheld by an appeals court in 1997. The saga of this sleeping patent finally came to rest.

Source: T. Riordan, "Patents: Two Investors Hope They Will Finally Win Compensation for a Device That Was Squelched," *New York Times*, July 21, 1997, p. D2; and A. Salpukas, "Award to Lighting Inventors Upheld On Appeal," *New York Times*, September 1, 1997, p. D32.

The Reality Checkpoint on the patent for solid-state ballast to be used in fluorescent lighting is one example of the use of sleeping patents. Other examples—all instances of an incumbent attempting to inhibit rival expansion—also exist. Alcoa achieved its dominant market position largely on the strength of Charles Martin Hall's electrolytic process for the reduction of aluminum bauxite ore. Fifteen years after it was formed, the company bought up the competing Bradley patents on an alternative reduction process—one that Alcoa never used. Similarly, Du Pont's patent of the synthetic fiber nylon was accompanied by the company's filing of literally hundreds of other patents all based on variants of the same molecule. Perhaps the best example of the use of sleeping patents comes from Hollywood. Film companies regularly buy the film rights to books, staged plays, and submitted screenplays knowing that many of these script ideas will never be turned into a final product. In part, each film company simply wants to make sure that a rival producer does not get the chance to make a film based on this material.

23.5 PATENT LICENSING

Efficiency requires that the existing stock of information should be available to all buyers at the marginal cost involved in sharing such knowledge. However, since this would imply a "price for information" of near zero, it would leave little incentive for anyone to produce new information as embodied in new goods or new technologies. Patent protection is an effort to cut a middle path between these two pressures. The firm receiving the patent is protected (to some extent) from sharing its discovery with others for free. In fact, it does not have to share it at all.

One interesting possibility is that an innovating firm might be willing to share its technical advance with other firms for a price. When this happens, it results in a licensing agreement between the patent owner and the patent user. Not sharing the patent at all can be interpreted as charging a very high (perhaps infinite) licensing fee. Actual licensing reflects a movement away from such a high fee and toward a price for information that is closer to—if still some way off—the efficient charge of near zero. In this sense, the licensing of a patent is unambiguously a good thing. The question is, does an innovating firm have a profit incentive to license its discovery?

The most obvious case in which a firm would prefer to license an innovation is if the licensee operates in a totally different market from the licensor. For example, a U.S. firm that has a patent on a particular product or process innovation may prefer to license a foreign firm to use this patent (for a fee, of course) rather than setting up a foreign subsidiary or exporting. About the only reasons for not licensing are, first, that the licensor may not be able to secure a satisfactory payment for the license except after extensive bargaining. If such negotiations will be prolonged, either or both parties may decide that it is simply not worthwhile. Second, the licensor may fear that, ultimately, the foreign licensee will produce in some market where it competes directly with the licensor. Finally, there is the fear that the licensee may—by acquiring rights to use the new process or product—improve its ability to develop the next generation of this technology by itself and thereby enhance its future ability to compete.

While these fears are undoubtedly real, there are considerable offsetting benefits to licensing agreements. Licensing gains revenue for the innovator today. Because the cost of sharing the information is low, any such revenue translates into profit.

What about cases though where the licensor and licensee are not separated by large geographic distance but instead are competitors in the same market? Will an innovating firm license its patented discovery for use by some or all of its rivals? The answer depends on market structure and the strength of competition in the market.

23.5.1 The Incentive for an Oligopolist to License a Nondrastic Innovation

Consider the toughest type of competition—price competition between firms making identical products. A firm that obtains a patent on a new technology that permits it to sell at a lower cost has little incentive to license the process to a competitor. Suppose, for example, that both firms are currently selling at a price equal to their (constant) marginal cost of $15 and that one firm has discovered a way to reduce this cost to $12. Without licensing its rival, the innovating firm can supply the entire market at $14.99 and drive its competitor from the market while earning a $2.99 profit on every unit sold. If it tries to sell a license to its rival, the only sensible royalty rate is $2.99 per unit. The rival firm will pay no higher royalty since then it will be unable to compete because its cost will be $12 plus the royalty, which is no better than its current cost of $15. At any lower royalty, the rival will force the innovating firm to lower its price below the current $14.99. But at a royalty of exactly $2.99, both firms will sell at $14.99 and split the market. The licensing firm loses $2.99 on those units it would have sold if it had not licensed but stayed a monopolist. It then gains the $2.99 back as a royalty payment on each of those same units now sold by its rival. In short, licensing gains the innovator nothing. Hence, the incentive to license is very small when the competition is Bertrand.[8]

By contrast, consider a market in which firms are Cournot competitors. In this case, a patent holder has a strong incentive to license, as a simple example shows. Assume that demand for the product in question is $P = 120 - Q$ and that there are three firms in the market, each with constant marginal costs of $60. Then we know from our earlier analysis that the Cournot equilibrium output of each firm is 15 units, total output is 45, the equilibrium price is $75, and each firm is making profits of $225.

Suppose now that one firm makes a nondrastic process innovation lowering its cost to $40 per unit, while the other two firms continue to produce at the higher value of $60 per unit. If the innovating firm does not license the innovation, then the Cournot–Nash equilibrium price falls to $70. The innovating firm increases its output to 30 units, while the other, high-cost firms reduce their outputs to 10 units. Profit to the innovating firm increases to $900 while profit to each of the other firms falls to $100.[9]

Now assume that the innovating firm agrees to license the innovation to its rivals at a fee of $10 per unit that each rival produces. This means that the innovator's costs are $40 per unit and the other firms' costs are $50 per unit. At the post-licensing equilibrium the

[8] For the patent holder that is selling in a *differentiated products market*, the analysis is a bit more complicated. Here, each additional license has three effects. First, it adds licensing revenue. Second, however, it makes the market more competitive and hurts the patent holder in its product market. Third, and as a result of the second effect, each additional license sold drives down the market value of licenses in general. In other words, the demand curve for licenses will be downward sloping because the more that are sold, the more competitive is the market and therefore the less any licensee can afford to pay for a license. Because the patent holder is the monopoly supplier of such licenses, its marginal revenue curve for selling will lie below the demand curve for licenses.

[9] These numbers come from simple application of the equations for the Cournot–Nash equilibrium that we have developed in previous chapters.

Derivation Checkpoint
Optimal License Price

Suppose that demand is $P = A - BQ$ and that the innovation gives marginal costs of c. Further suppose that the innovator charges a royalty of r per unit to its rivals. Then the innovator's profit is:

$$\pi = \frac{(A - Nc + (N-1)(c+r))^2}{B(N+1)^2} + r(N+1)\frac{(A - Nc + (N-1)(c+r))}{B(N+1)}$$

The first term is profit from the innovator's sales and the second is revenue from the royalty agreement. Differentiating π with respect to r and simplifying confirms that the innovator's profit in increasing in the royalty price r. So the innovator should set as high a royalty price as is possible consistent with the non-innovator's being willing to pay the royalty price.

innovating firm's output is 25 units while the other firms produce 15 units each so that price is \$65. The profit of the innovating firm is now \$25 per unit on its own sales plus \$10 per unit on the sales of its two rivals, giving a total profit of \$925. For each non-innovating firm, profit is \$15 per unit, giving each firm profit of \$225.

Licensing is, indeed, potentially quite profitable. Moreover, the licensing fee of \$10 that we have chosen is not even the best that the innovating firm can do.[10] We show (in the inset) that the innovator should actually push the license price as close as possible to the difference in costs that the innovation generates—in our example, as close as possible to \$20. Suppose for example that the innovator charges a royalty rate of \$20 per unit (more accurately, \$19.99). This restores the equilibrium with the innovation but without licensing. The innovator produces 30 units and each non-innovating firm 10 units, giving a product price of \$70. Profit of each non-innovating firm is, once again, \$100 since their costs are \$60 per unit. By contrast, profit of the licensing firm is \$30 per unit on its own output and \$20 per unit on the output of its rivals, giving the licensor a total profit of \$1,300. The message then is clear. For a Cournot firm with a nondrastic innovation, licensing its discovery is very attractive.

23.5.2 Licensing, Drastic Innovations, and Monopoly Power

What if the innovation had been drastic? Or what if the industry had been a monopoly instead of an oligopoly? Consider each question in turn. If one firm in a Cournot oligopoly patents a drastic innovation, it will not want to license its discovery. Take the simple case of a duopoly. Without licensing, the innovating firm becomes a monopoly. The innovation offers such a dramatic reduction in cost that even when it sets the monopoly price associated with that cost, it still underprices its old duopolist rival, while earning considerable monopoly profit. Here, nothing can be gained by licensing. If the rival is permitted to compete, the market returns to being a duopoly except at lower cost. The most the rival would ever pay for the

[10] For details see Katz and Shapiro (1985).

license is therefore its share of the duopoly profit. Combining this with the innovator's share would yield the innovator a total profit with licensing equal to the profit earned by two duopoly firms. Yet we know that—because the firms cannot collude—this is generally a smaller amount than the innovating firm could earn as a pure monopolist without licensing. Accordingly, a Cournot firm that makes a drastic innovation will not share its discovery with rivals even for a fee. Of course, this is also true for firms engaged in Bertrand competition. In all such cases, the oligopolist that makes a truly dramatic breakthrough may be expected to emerge as a monopolist driving his former competitors from the field.

Turning next to the case of monopoly in the first place, we now have to permit the innovation to take place at an outside firm or laboratory if we are to consider any licensing. (If the monopolist makes the innovation himself, there is no other firm to which he can license!) It should be clear that in such cases—whether the innovation is drastic or nondrastic—the innovating firm will license the monopolist. Since the patent holder is not active in the market himself, the only way he can obtain any revenue from his discovery is to sell or license it to the monopolist.

The interesting point in this case is the precise form that such a licensing contract should take. Should the licensor charge a royalty of *X* per unit? Or should he charge a fixed fee independent of output? Or should he use some combination of both? You should recognize that charging a per unit royalty—while it has the advantage that it relates revenue directly to usage—runs into the familiar problem of double marginalization (see Chapters 17 and 18). It raises the licensed firm's marginal cost so that—after that firm adds its markup—the price to the final consumer is doubly distorted and sales volume is restricted. In this light, it should not be surprising that the innovating lab will do best by using a two-part tariff. The principal part of this scheme will be a fixed fee (per month or per year). The second part will be a small royalty per unit reflecting any per unit cost the patent holder incurs in licensing his technology. For a transfer of pure information, this per unit charge will be zero. But if the patent holder needs to offer services or technical advice that increases with the frequency with which the technology is used, this fee will be positive. The licensing contract is much like a franchising contract. In principle, the inventor can appropriate all the increased profit that the invention brings if the contract is written correctly, that is, with a fixed fee exactly equal to that additional profit. In practice, however, the patent holder's bargaining position will usually not be strong enough to achieve this outcome. When the manufacturer has a monopoly in the product market, the inventor needs the manufacturer just as much as the manufacturer needs the inventor.

23.5.3 Patent Licensing, Social Welfare, and Public Policy

The foregoing cases indicate that most of the time an innovator has an interest in licensing his discovery. This is a reassuring result because our intuition is that licensing is typically a desirable outcome. Katz and Shapiro (1985) have provided a formal argument that licensing nearly always increases social welfare. Specifically, they show that licensing is socially desirable if total output increases as a result of the licensing activity. To see why, note that licensing will not take place unless it is both profitable and increases aggregate profit. The license agreement will not be signed unless the licensees see some benefit from it and will not be offered unless the licensor also sees some benefit from it. If, in addition to this mutual gain in profit, the license agreement increases total output, then the price will be lower and consumer surplus will be increased too. In other words, if the license agreement increases total output, both consumers and producers gain from the agreement, and so the agreement is socially desirable. Yet even if this fails to happen—even if the industry output is unchanged—licensing is still likely to be socially beneficial since the licensing revenue at least increases producer surplus. Somebody then, either a producer or a consumer or both, is made better off by licensing.

Moreover, licensing may have other beneficial effects. First, if a firm knows that it is going to gain profits from licensing its research findings as well as (or instead of) exploiting the research itself, this should increase the incentive to undertake research. Further, the possibility that a firm can obtain a license to use a particular innovation will reduce wasteful R&D that either duplicates existing research effort or is intended merely to invent around an existing patent.

Consider an entrant whose profit (in present value terms) under duopoly is $5 million but who would incur an R&D expenditure of $3 to develop its own product alternative. In the absence of any licensing, the entrant will pursue this investment since it yields a net gain of $2 million. Yet if this is the case, then the monopolist firm will know that whether it licenses or not, it will soon be a duopolist. If the monopolist firm licenses its technology to the entrant for $3 million, the entrant is just as well off and the monopolist now gets the licensing revenue. In addition, society avoids the unnecessary expenditure of $2 million that the entrant would otherwise have made. The moral of this section therefore seems quite clear. Public policy should actively encourage the licensing of innovations as much as possible.

There is, however, need for a cautionary note. Licensing might involve some risks. First, consider the risks associated with licensing based upon an output-related royalty. Imagine as well that the licensing agreement holds for the outstanding duration of the patent that is being licensed since, after that, the information becomes publicly available. If the royalty rate extracts almost all of the additional profits that the licensee might expect to make, there is the risk that the licensee will take the license in order to gain experience with the technology but then actually produce very little during the period of the license agreement, which means, of course, that very little is actually paid for the license. Alternatively, if output is difficult to monitor, the licensee has the incentive to lie about how much is actually being produced. What may be necessary is that the licensor tie the license agreement to some agreed minimum level of output on the part of the licensee but even this is not always easy to negotiate or enforce.

A further risk in licensing is that it can be difficult to write enforceable contracts that limit the ways in which licensees can use the license. Typically, the licensor will want to limit the markets into which the licensee can sell, for example, to avoid direct competition with the licensor or with other licensees. This may be possible within a particular jurisdiction such as the United States, although even here antitrust laws may prevent such market-limiting agreements. But it is almost impossible to write binding contracts that limit the international markets in which licensees can operate. In addition, access to a particular process or product technology may enhance the ability of a licensee to develop related technologies that are not covered by the patent being licensed. Once again, it is almost impossible to write enforceable contracts that protect the licensor from such imitation or at least give the licensor some return from the new technologies that licensees develop.

Licensing raises public policy issues that suggest caution in favoring and promoting every licensing agreement. One danger is that licensing contracts include restrictions on price or geographic territory that create monopolies with exclusive territories–monopolies would otherwise be illegal under the antitrust laws. Matters become particularly complicated when, as often happens, one patent leads to another, complementary development. One firm creates, say, a new antibiotic that has some occasional and serious side effects. Then another

9781405176323_4_023.qxd 10/19/07 8:16 PM Page 624

firm develops a means to undo the side effects of the first firm's drug. The two firms may strike a deal that licenses each to produce the other's product. Yet it is easy to see that this agreement may often include terms that exclude other firms. Such dangers are recognized by U.S. policy, which tends to limit severely the ability of reciprocal licensing agreements to include exclusive provisions. Still, the example serves to make clear that the tension between promoting licensing and realizing its associated benefits, on the one hand, and the potential risk of collusion that licensing may foster, on the other, is real.

Indeed, the increasing complexity of technical advances has resulted in what some might call a "patent thicket." As advance builds on advance, and technical progress increasingly draws from learning in different fields, the technology involved in bringing a new product to market may in fact build a host of patented techniques each of which is owned by a different entity. The innovator may then need to get the approval of each of the individual patent holders before proceeding. In turn, this can involve the coordination difficulties of complements that we first described in Chapter 8. Acting individually, each patent-holder may set too high a license fee with the result that all are worse off. Cross-licensing agreements by which firms agree to license their patents to each other and patent pools, by which a group of firms agrees to pool a set of patents and to license them as a package, have become increasingly popular ways to solve the coordination problems inherent in the "patent thicket." They run the risk of permitting cooperation beyond the technological sphere and giving the parties a chance to wield their technological power collectively against potential entrants. However, without such efforts, it may be impossible for any new entrants to cut their way through the thicket and thereby provide any competitive pressure.[11]

23.3 **Practice Problem**

Two firms compete in a Cournot-type duopoly. The industry demand is given by $P = 100 - 2Q$. Each firm has a constant average and marginal cost of \$60.

a. What is the current equilibrium price and quantity in the industry?
b. Suppose that one firm discovers a procedure that lowers its average and marginal cost to \$50.
   (i) If the innovator does not license its product but simply competes as the low-cost firm in a Cournot duopoly, what will be the innovator's profit?
   (ii) What will be the innovator's profit if it licenses the technology to its competitor at a royalty rate of \$10?
   (iii) Suppose instead that the innovator licenses the technology for a fixed fee. What is the highest fee that the non-innovator will be willing to pay. What will the innovator's profits be if it can charge the highest possible such fee?

23.6 RECENT PATENT POLICY DEVELOPMENTS

In the first half of the 1980s a number of events occurred that, together, greatly increased the legal protection of patent rights in the U.S. The first and perhaps most crucial step was a legal reorganization that gave the Court of Appeals for the Federal Circuit (CAFC), in

[11] See Lerner and Tirole (2004).

Washington, D.C., exclusive jurisdiction over patent appeals in an effort to unify the legal treatment of patent rights. This court is widely considered to have a very "pro-patent" view and, until recently, its decisions were left unquestioned by the U.S. Supreme Court. The CAFC emerged as the final and sole arbiter of patent disputes and its pro-patent views became widely reflected in lower court cases. Just how much stronger patent protection was to become became apparent in the 1986 patent infringement suit filed by Polaroid against Kodak regarding Kodak's production and sale of an instant-film and instant-picture camera.

Prior to that decision, losers in a patent infringement case had typically paid small penalties and been permitted to continue to produce so long as they paid appropriate royalties to the winner. However, when Polaroid won the suit Kodak was required to pay very large penalties and, most importantly, forced to stop producing its instant camera. Since shutting down a high volume production line is very expensive—even if only for a few weeks—the fact that the courts were now willing to impose such an outcome put all firms on notice that patent infringement cases were serious business. Moreover, the Kodak/Polaroid case was quickly followed by very aggressive behavior on the part of one firm, Texas Instruments (TI) in filing infringement suits (mostly against foreign firms) and raising royalty fees that also served to put high technology firms on notice. In the technology sector where reverse engineering has always been important, TI was so aggressive that its royalty fees and court awards began to outstrip its production activities as a source of revenue.

In short, a new legal environment of much stronger protection for patent-holder rights emerged in the U.S. in the 1980s. It may not be surprising then to discover that there was an explosion of patent activity over the next several years. Between 1983 and 2000, the annual number of patent applications doubled while the annual number of patents actually granted rose by an even greater 170 percent.

There has been increasing concern that the strengthened protection of patent rights has become too aggressive. To begin with, recent empirical evidence casts considerable doubt that stronger patent enforcement yields better innovation results. Drawing on a range of sources, Lerner (2000) identified 177 distinct patent policy changes in 60 countries over 150 years such as those that lengthened or broadened patents, those that reduced the patent filing fee, those that required compulsory licensing, and so forth. He then examined the effect of these changes on the rate of patenting. He found that increased patent protection sharply increased patenting by foreign firms but decreased patenting by domestic innovators. The overall effect was positive. However, the inference is that foreign companies used patents to protect themselves against domestic competitors. Hence, while patents may have enhanced international trade, their effect on innovation was negligible.

Moser (2005) constructed internationally comparable data using the catalogues of two nineteenth-century world fairs: the Crystal Palace Exhibition in London, 1851, and the Centennial Exhibition in Philadelphia, 1876. These included innovations that were not patented, as well as those that were, and innovations from countries both with and without patent laws. He found no evidence that patent laws increased levels of innovative activity. Instead, they simply affected the direction of innovation. Relative to countries with strong patent protection, inventors in countries without such protection simply concentrated their efforts in industries where secrecy was easily maintained, leaving the overall rate of innovative efforts unchanged. Similarly, Sakakibara and Branstetter (2001) found no evidence that a strengthening of Japanese patent laws in 1988 led to any increased R&D spending or innovative output.

Fears that patent protection had gone too far reached a dramatic high point in February 2007 when the three million customers of the BlackBerry wireless e-mail service were

Reality Checkpoint

Patent Policy in the Information Age: Getting One (click) Up On the Competition

The most valuable real estate lots bordering the information superhighway may simply be ideas about how to use this new tool to increase profit. Ideas about how to do business or so-called business methods are different from the technological innovations that we have discussed elsewhere in this chapter. Yet in the information age, they may be just as valuable.

A leading example in this regard is a patent issued to the online store, Amazon.com. Amazon customers shop the site and list the items that they wish to purchase. At the end of their visit, customers simply make one click of their mouse and their order is taken and then shipped. Amazon applied for and received a patent for this 1-Click feature and touts it to all potential customers.

In October, 1999 the traditional "brick and mortar" bookseller, Barnes & Noble, introduced an Express Lane at its recently opened website. The Express Lane checkout also permitted customers to finalize their shopping with one mouse click. Amazon instantly sued, claiming that the Express Lane model was a violation of its 1-Click patent. Amazon won and a federal appeals court subsequently issued an injunction preventing Barnes & Noble from using its Express Lane feature. Ultimately, it dropped the Express Lane and initiated an Express Checkout, which permitted finalizing an order in two clicks.

The 1-Click case is not unique. Business method patents have become common ever since a U.S. Court of Appeals rule in favor of Signature Financial Group's patent for an algorithm to manage mutual fund investments (*State Street Bank and Trust Co., Inc. v. Signature Financial Group*, 149 F.3d 1368, Fed. Cir (1998)). Consider the business method practice called upselling. A customer at Burger King, for instance, might order a Whopper sandwich, an order of french fries, and a small salad for a total of $7.14. When checking out, the cashier might say, "for just 86 cents more, you can also have a soft drink that regularly sells for $1.29." If the customer agrees to this upsale, Burger King obviously receives more revenue. Yet, Burger King will not get to keep all the extra funds. A chunk of it will go to Walker Digital as a licensing fee because Walker (owned by Jay Walker, the founder of Priceline.com), owns a patent on this process and Burger King must pay for it.

Since the *State Street* decision, filings for business method patents have nearly tripled to now reach between 1,000 and 2,000 per year. Such patents raise an interesting qualification to optimal patent policy. Where innovations require lots of development time and expense; where they can be clearly identified; and where they need protection against imitation, a patent award may well be necessary for technical progress to occur. However, when innovations are highly incremental and build on a host of other advances so that it is hard to identify any one actual breakthrough in any one application, a patent system can actually slow down technical progress. Many economists, including Gallini (2002) and Hall (2003), suspect method patents may fall in this second category. The irony is that just as the Internet and related developments are making information cheap, the rush to patent business methods and practices may be making the exploitation of that information ever more expensive.

Sources: S. Hansell "Barnes And Noble Injunction Lifted," *New York Times*, February 15, 1991, C1; and J. Angwin "Business Method Patents, Key to Priceline, Draw Growing Protest," *Wall Street Journal*, October 3, 2002, p. B1.

Reality Checkpoint

It Was Patently Obvious and Therefore, Not Patent Worthy

On April 30, 2007, the U.S. Supreme Court issued an important ruling that substantially raised the bar for obtaining patents on new products that combine elements of pre-existing inventions. The case involved a patent infringement lawsuit filed by Teleflex, Inc. against KSR International over the development of an adjustable gas pedal for use on cars and trucks equipped with electronic engine controls. The position of the accelerator pedal in many cars is not adjustable. Instead, the driver adjusts the position of the seat until the pedal is a comfortable distance away. However, in the 1970s, a number of inventors began to develop adjustable pedals that could slide forward or backward without changing the effect of depressing the pedal a specific amount. In older cars, that effect is transmitted by means of a cable that typically opens up valves in the fuel injection unit of the engine. In more modern cars, however, the cable has been dispensed with and the mechanical connection has been replaced with a computer sensor that electronically transmits the acceleration or deceleration signal to the engine. KSR is a company with a history of making adjustable mechanical pedals for the major automobile companies. In 1999, it won a contract with General Motors to provide an adjustable pedal with an electronic sensor mounted at the pedal's fixed pivot point to communicate the necessary information. Teleflex, which had a patent for one type of electronic sensor for gasoline pedals, claimed patent infringement and demanded royalties. KSR refused to pay on the ground that Teleflex had combined existing elements including those in other patented sensors in an obvious manner so that its patent was therefore invalid. KSR won in Federal District Court in Detroit, but that decision was overturned in 2005 by the CAFC, the court with exclusive jurisdiction over patent appeals.

For many years, the Supreme Court has let the CAFC judgments stand unreviewed. However, the court now seems to have taken interest in the CAFC decisions. In particular, the court took issue with the way the CAFC employed an approach referred to as the "teaching, suggestion, or motivation" test (TSM test). That approach holds that a patent claim is only proved obvious if "some motivation or suggestion to combine the prior art teachings" can be found in the prior art, the nature of the problem, or the knowledge of a person having ordinary skill in the art. The Supreme Court said that the CAFC was applying the TSM standard too rigidly. In particular, it ruled that when the innovation simply yields predictable results based on existing technology then it is not entitled to patent protection regardless of whether that prediction has actually been tested in practice. The court found that the Teleflex patent on electric sensors was exactly this type of innovation and so, reversed the CAFC judgment and found in favor of KSR.

Because most inventions combine previously known elements, the decision in the *KSR v. Teleflex* case is widely recognized as a signal that two decades of aggressive patent enforcement were coming to an end. Indeed, it was accompanied by a second very similar decision in which the court found for Microsoft against a charge of patent infringement by AT&T. The result of these decision is almost surely that patents will be harder to obtain and harder to defend.

Sources: *KSR International Co. vs. Teleflex Inc.*, 550 U.S. ____(2007); *Microsoft Corporation v. AT&T Corp.* 550 U.S. _____(2007); and L. Greenhouse, "High Court Puts Limits on Patents," *New York Times*, May 1, 2007.

threatened with a shutdown due to a patent dispute. A small Virginia firm, NTP, had developed and patented the technology for a wireless e-mail device in 1990. However, NTP never produced a product nor did it make any effort to license the technology to others. In 1998, the Canadian firm, Research In Motion (RIM), unveiled its first wireless e-mail device. Sales took off sharply. Although RIM claimed that it had developed the technology on its own, NTP filed suit against BlackBerry in 2001. In 2002, a U.S. jury found the Canadian firm guilty of 16 counts of patent infringement. On appeal, seven of these were dismissed in 2004 but that still left 9 outstanding. In 2005, RIM offered a $450 million to NTP to settle the case, but that settlement was rejected by the trial judge. In January 2006, the Supreme Court refused to hear any further appeal and a hearing to order a shutdown of the BlackBerry service was scheduled for Friday, February 24, 2006. The hearing did not reach a final decision. However, after further negotiations, RIM and NTP reached a settlement in which RIM made a one-time payment of $615 million to the Virginian firm for unfettered use of the technology. The agreement came even as the U.S. Patent and Trademark Office (USPTO) was conducting a review of the legitimacy of NTP's patents. Many felt that the strong pro-patent laws had effectively allowed NTP to extort the payment from RIM and forced it to rush to a settlement before the USPTO completed its review.

In April 2007, the Supreme Court served notice that it too was concerned about excessive patent protection. (See Reality Checkpoint.) In *KSR International vs. Teleflex, Inc* the court ruled that new products that combine elements of pre-existing inventions and that result from nothing more than "ordinary innovation" with no more than predictable results," were not entitled to patent protection. The decision was notable for its clear statement that the patent system could be used to undermine innovation and its unanimity. As a result, most experts believe that the decision raised the bar substantially for future patent applications. It also opened the door to a re-examination of existing patents and gave judges much more leeway to dismiss patent infringement suits.

23.7 EMPIRICAL APPLICATION
Patent Law and Practice in the Semiconductor Industry

The semiconductor industry was not immune to the patent fever that spread through America in the last part of the twentieth century. As Hall and Ziedonis (2001) document, patent awards per million dollars of R&D spending in this industry doubled in the ten years following 1982. What makes this increase particularly striking is that semiconductor industry representatives have been surveyed repeatedly and consistently reported that patents are not a very effective way to appropriate the returns on R&D investments. Because the semiconductor industry is one of rapid technological change where product lifecycles are short, semiconductor firms have instead relied on lead time, secrecy, and product design tactics to reap the profits from their innovations. What then is the reason for the increased patent activity by semiconductor firms? How is it related to the changed legal environment?

Hall and Ziedonis (2001) examine the patent explosion in the semiconductor industry using data from 95 industry firms covering the years 1979 to 1995. These firms were awarded over 17,000 patents in this period. Hall and Ziedonis (2001) model these successful patents as the outcome of a patent production process that relates the *i*th firm's production of patents in year *t* or p_{it} to a set of variables X_{it} including the firm's R&D spending and its overall size. However, they recognize that p_{it} is what is called a count variable. That is, it counts the number of successes that take place during a time interval of given length. Thus, p_{it} can only take discrete integer values and often will be zero. If we consider p_{it} to have a random component, then we need to assume a probability distribution that recognizes these features. The natural choice for this purpose is the Poisson distribution, which gives the probability $f(\lambda, p)$ that there are p occurrences of a random variable in fixed time interval as:

$$f(\lambda, p) = \frac{e^{-\lambda}\lambda^{p}}{p!} \tag{23.9}$$

The Poisson distribution has a very nice feature in that it is fully characterized by the parameter λ, which is both its mean and its variance. Thus, Hall and Ziedonis (2001) model patent production as a Poisson process that has a conditional mean λ_{it} that is an exponential function of X_{it} as follows:

$$E(p_{it}|X_{it}) = \lambda_{it} = \exp(X_{it}\beta + \gamma_t) \tag{23.10}$$

where γ_t is a 1,0 dummy variable for each year reflecting factors in that year that are common to the patenting activity of all semiconductor firms. Of course, we can linearize this relationship by taking logs to yield:

$$\ln \lambda_{it} = X_{it}\beta + \gamma_t \tag{23.11}$$

Hall and Ziedonis (2001) measure p_{it} as the number of patents per employee. The variables in X_{it} include: (1) the log of firm R&D spending per employee; (2) a 1,0 dummy variable equal to one if the firm reported no R&D spending that year and zero otherwise; (3) the log of firm size measured as the number of employees in thousands; (4) the log of the plant and equipment value per employee as a measure of the capital intensity of the firm's production; (5) a 1,0 dummy variable equal to 1 if the firm entered the market after 1982 and zero otherwise; (6) a 1,0 dummy variable equal to 1 if the firm is a design firm that does no fabrication and zero if it is a manufacturing firm; (7) a 1,0 dummy variable equal to 1 if the firm is Texas Instruments and zero otherwise; and (8) the log of the firm's age.

The first three variables reflect the standard view of patents as the output of a process in which R&D is the input and in which there may be scale economies. The fourth variable allows Hall and Ziedonis (2001) to test the hypothesis that part of the increased patenting following the change in the patent enforcement environment reflects the decision by firms with large sunk costs who cannot afford to get "held up" in a patent dispute, to expand their patent portfolio rapidly to guard against such "hold-ups." The fifth and sixth variables allow them to test a second hypothesis, namely, that another reason for the rise in patent activity was that the new legal environment made it attractive for design firms who, unlike semiconductor manufacturers, rely heavily on patents to enter the market, and thereby they change the mix of firms in the semiconductor industry to one more likely to patent. The seventh variable captures the well-known super-aggressive patenting strategy adopted by TI, while the age variable allows for firm-specific learning.

Hall and Ziedonis (2001) observe the annual number of patents p_{it} by firm i in year t for 95 semiconductor firms from 1979 to 1995, and use these to estimate a Poisson process in which the mean λ is taken to be conditional on a set of firm characteristics X_{it} and time dummies γ_{it}. Because they are estimating a Poisson distribution, the assumptions of Ordinary Least Squares (OLS) do not hold. Instead, Hall and Ziedonis use maximum likelihood estimation

Table 23.3 Parameter estimates for expected patent output by semiconductor firms

Variable	*Estimated coefficient*	*Standard error*	*Estimated coefficient*	*Standard error*
Log R&D per employee	0.190	(0.084)	0.196	(0.117)
Dummy for no reported R&D	−1.690	(0.830)	−1.690	(0.840)
Log firm size	0.854	(0.032)	0.850	(0.034)
Log firm P&E per employee	0.601	(0.113)	0.603	(0.114)
Dummy for post-1982 entry	0.491	(0.169)	0.491	(0.199)
Dummy for design firm			−0.130	(0.185)
Dummy for Texas Instruments	0.799	(0.111)	0.798	(0.115)
Log of firm age			0.220	(0.146)

(MLE). Because λ is both the mean and the variance of the Poisson distribution, comparing the variance of the data with the mean is a natural test for the appropriateness of the underlying Poisson specification. The results of their two regressions that do best on this test are shown in Table 23.3.

The estimates in both regressions for the first four variables imply roughly constant returns to scale in patent production. As firm size (measured by the number of employees) grows semiconductor firms tend to increase their patent output proportionately. This is similar to the finding of other researchers, e.g., Hall, Griliches, and Hausmann (1986). It is also clear that TI has a markedly higher propensity to patent than do other semiconductor firms consistent with TI's well-known aggressive patenting policy during these years.

Most importantly, both of the key hypotheses are supported by these data. Firms with capital-intensive production as measured by the amount of plant and equipment per employee do significantly more patenting than others do. In addition, it appears that the firms that entered the industry following the 1982 centralization of patent law cases at the CAFC were much more likely to patent than the firms that were already in the industry. The coefficient on the post-1982 entry variable is highly significant in both regressions. While the coefficient on the design firm variable is not significant in the second regression, it is if the post-1982 entry variable is omitted indicating that the entry variable reflects mostly entry by design firms.

In sum, Hall and Ziedonis (2001) interpret their results as indicating two major reasons for the jump in patenting efforts in the semiconductor industry after 1983. One is that the new pro-patent environment and the prospect of having production actually stopped by legal injunction was particularly threatening to capital-intensive firms with heavy sunk costs. The result was that they responded strategically by rapidly accumulating a portfolio of patents to protect their products and processes. The second was that the changed legal environment also induced entry of purely design firms that have an inherently greater propensity to patent their findings in any case.

There is also a third effect revealed by the Hall and Ziedonis (2001) findings. This is that the changed legal framework led all semiconductor firms to patent more. The evidence for this is in the time dummies (not shown). Normalizing so that the effect is zero in 1979, the pattern of these coefficients is seen in Figure 23.3. Here, the dotted line reflects the time dummy coefficients from the first of the two regressions above, while the dashed line reflect the same coefficients in the second regression estimates.

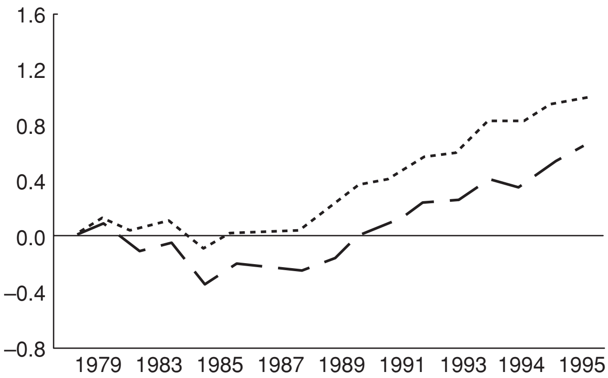

Figure 23.3 Pattern of regression time coefficients in semiconductor patent behavior

What both sets of estimates clearly show is that after 1986, even after controlling for the mix of semiconductor firm characteristics, there was a steady increase in the proclivity to patent with each successive year. The new pro-patent environment is the most obvious explanation for this rise that is common to all semiconductor firms.

Summary

By giving innovators a legally enforceable means of earning a return on their discoveries, patents and copyrights do provide incentives for innovative activity that might otherwise not be undertaken. Yet patents also confer monopoly power on the patent holder, with all the price distortions that such power entails. In addition, patent rules may enhance the ability of existing monopolies to maintain their current dominant position against would-be entrants. One mechanism by which this may occur is through the use of "sleeping patents" designed to buffer the invention against any and all attacks from rival innovations that might permit an entrant to "invent around" the original patent.

Licensing agreements by which firm permit the use of their patented knowledge for a fee can help ameliorate the patent tension. This is because such agreements both permit wider use of the innovation and also allows an innovator to earn a greater return on her R&D investments than she otherwise would receive. However, licensing contracts can be difficult to enforce except by imposing restrictions that can be harmful to competition.

Within the United States, the 1980s marked a sharp increase in the legal protection of patents against infringement. This was followed by an equally sharp increase in both patent applications and patent grants. Empirical evidence from the semiconductor industry suggests that this reflects in part the desire of firms with large sunk investments in products and processes to avoid disruption of their production by accumulating a large patent portfolio and the encouragement of entry by new firms that rely more heavily on patents to appropriate the gains of their innovations. That evidence also confirms that there was a general rise in patent proclivity across all semiconductor firms.

More recently, there has been concern that U.S. patent protection has been overly strong, especially since there is little evidence that it has led to faster innovation. Accordingly, recent court decisions have pared back these protections. Our discussion of licensing and of recent patent legal developments make clear that there is no way to eliminate the tension between allocative efficiency and innovative activity that a patent system raises either in theory or in practice.

Problems

1. Let the inverse demand for a particular product be given by $P = 250 - Q$. The product is offered by two Cournot firms each of which has a current marginal cost of \$100. Both firms can invest a sum K to establish a research facility to develop a new process with lower marginal costs. The probability of success is ρ.
   a. Assume that the new process is expected to have marginal costs of \$70. Derive a relationship between K and ρ under which
      (i) neither firm establishes the research facility;
      (ii) only one firm establishes a research facility;
      (iii) both firms establish a research facility.
   b. Can there be "too much" R&D? Illustrate your answers in a diagram with ρ on one axis and K on the other.
   c. Now assume that the marginal costs of the new process are expected to be \$40. How does this affect your answers to 8a?

2. In the text of this chapter we considered sleeping patents in the context of a process innovation. The same principles apply in the case of a product innovation. To see why, consider the following example: Assume that there are 100 aspiring Olympic swimmers whose tastes for low-water-resistance colored swimming suits are evenly distributed over the color spectrum from black to yellow. The "length" of this spectrum is normalized to be one unit. Each of these swimmers values the loss of utility from being offered swimming suits in other than their favorite color at \$10 per unit of "distance." Each swimmer will buy exactly one swimming suit per period provided that the full price for the suit—the price charged by the firm plus the value of utility loss if there is a color difference between the suits on offer and the swimmer's favorite color—is less than \$100 (these are very keen swimmers!). Production of low-water-resistance swimming suits is currently feasible only in black and is controlled by a monopolist who has a patent on the production of the black material. The marginal cost of making a swimming suit is \$25.
   a. What is the current profit-maximizing price per suit and what are the monopolist's per-period profits?

      Now assume that research can be conducted that will allow the swimming suits also to be manufactured in yellow at the same marginal cost of \$25.
   b. If the monopolist undertakes the research and introduces the new color what will be the resulting equilibrium prices of black and yellow swimming suits? What is the impact on the monopolist's per-period profit, ignoring research costs?
   c. If a new entrant undertakes the research and introduces the new color, what will be the resulting equilibrium prices of black and yellow swimming suits? What will the entrant's per-period profit be, again ignoring research costs?

3. Return to problem 2, above.
   a. Confirm that the incumbent monopolist will be willing to spend more on researching the new color than the potential entrant.
   b. Assume that the research costs can be split into some amount R, which is pure research cost, and another amount D, which is development cost—the cost of transforming a successful innovation into a viable product. Calculate limits on R and D such that the monopolist will be willing to undertake the research into manufacture of yellow swimming suits and patent it but then leave the patent sleeping.

4. Consider a Cournot duopoly in which inverse demand is given by $P = 120 - Q$. Marginal cost of each firm is currently \$60.
   a. What is the Cournot equilibrium quantity for each firm, product price, and profit of each firm?

      Now assume that one of the firms develops a new technology that reduces marginal cost to \$30.
   b. If it keeps control of this innovation itself, what will be the new Cournot equilibrium outputs, product price, and profits of the two firms?

c. If it licenses the innovation to its rival at some per unit fee r, calculate the innovator's profit as a function of r. What is the profit-maximizing value of r for the licensor?

5. Return to Problem 4, above. Assume now that the innovator licenses the innovation to its rival for a fixed fee of L. What is the maximum fee that it can charge? Will the innovator prefer to set a per-unit license fee or a fixed license fee? What kind of licensing arrangement would consumers prefer?

6. Consider the same Cournot duopoly as in question 9, but now assume that the research has been conducted by an outside research firm. Suppose that this firm agrees to license the technology at a per unit fee of r. What license fee will the research firm charge
   a. if it licenses to only one of the duopolists?
   b. if it licenses to both?
   c. How are your answers to a. and b. affected if the research firm chooses instead to charge a fixed fee of L for the license?

7. Consider the same Cournot duopoly as in question 9, but now assume that research takes the following form. It costs a firm $x^2/2$ to undertake research at intensity x. If this research is successful, it reduces the firm's marginal cost by x, but the research results spill over to its rival, reducing the rival's marginal cost by $\beta.x$. Research intensity can take one of two values, $x = 30$ or $x = 15$, and the degree of spillover can also take two values, $\beta = 1/3$ or $\beta = 2/3$.
   a. Determine whether the firms will prefer the low or the high research intensity for each value of the research spillover.
   b. Assume that the two firms agree to coordinate their research efforts to maximize their joint profits. This technology cooperation does not, however, extend to cooperation between the firms in the final output market. Which value of research intensity will the two firms prefer for each value of the research spillover? How are consumers affected by the technology cooperation?
   c. Now assume that the firms form a research joint venture to share the results of their research, so that we have $\beta = 1$. What research intensity will the joint venture choose? How are consumers affected by the research joint venture?

References

Denicolò, V. 1996. "Patent Races and Optimal Patent Breadth and Length." *Journal of Industrial Economics* 44 (March): 249–65.

Gallini, N. 1992. "Patent Policy and Costly Imitation." *Rand Journal of Economics* 23 (Spring): 52–63.

——. 2002. "The Economics of Patents: Lessons from Recent U.S. Patent Reform." *Journal of Economic Perspectives* 16 (Spring): 131–54.

Gilbert, R. and C. Shapiro. 1990. "Optimal Patent Length and Breadth." *Rand Journal of Economics* 21 (Spring): 106–12.

Giuri, P., Mariani, M., et al. 2005. "Everything You Always Wanted to Know about Inventors (But Never Asked): Evidence From the PatVal-EU Survey." Working Paper 2005/20, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, Italy.

Hall, B. H., Z. Griliches, and J. Hausmann. 1986. "Patents And R and D: Is There A Lag?" *International Economic Review* 27 (June): 265–83.

Hall, B. and R. H. Ziedonis. 2001. "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995." *Rand Journal of Economics* 32 (Spring): 101–28.

Katz, M. and C. Shapiro. 1985. "On the Licensing of Innovation." *Rand Journal of Economics* 16 (Winter): 504–20.

Klemperer, P. 1990. "How Broad Should the Scope of Patent Protection Be?" *Rand Journal of Economics* 21 (Spring): 233–30.

Klette, T., and D. de Meza. 1986. "Is the Market Biased against R&D?" *Rand Journal of Economics* 17 (Spring): 133–9.

Lerner, J. 2000. "150 Years of Patent Protection." Harvard Business School Working Paper 00-039.

Lerner, J. and J. Tirole. 2004. "Efficient Patent Pools." *American Economic Review* 94 (June): 691–711.

Moser, P. 2005. "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-century World Fairs." *American Economic Review* 95 (September): 1214–36.

Nordhaus, W. 1969. *Invention, Growth and Welfare*. Cambridge, MA: MIT Press.

Reinganum, J. 1989. "The Timing of Innovation: Research, Development, and Diffusion." In R. Schmalensee and R. Willig, eds., *The Handbook of Industrial Organization*. Amsterdam: North-Holland, 849–908.

Sakakibara, M. and L. Branstetter. 2001. "Do Stronger Patents Induce More Innovation? Evidence from the 1998 Japanese Patent Law Reforms." *Rand Journal of Economics* 32 (Spring): 77–100.

Scotchmer, S. 2004. *Innovation and Incentives*. Cambridge, MA: MIT Press.