the extra funds needed to ensure that the innovator breaks even. However, in the absence of government intervention, the fact that $S > \rho$ means that no such R&D efforts will occur.[3]

This brings us back to the role of patent length in influencing innovative activity. Recall that $S = K/M$ and relax our assumption that the discount factor is zero. Then we can rewrite $S$ as $S(T) = K/M(T)$, where $M(T)$ is the discounted value of monopoly profits with a patent of length $T$. Clearly, $S$ gets smaller as $T$ increases. (Formally, $dS(T)/dT < 0$.) This brings attention to the obvious parameter that the Patent Office can influence in $S(T)$ is patent length as measured by $T$. Suppose first that $T$ is "short," giving $S(T) = S_1$ in Figure 21.2. Then there is no innovation to develop the new product. If, however, patent length is increased, giving $S(T) = S_2$ in Figure 21.2, one firm will be induced to undertake the innovation. Because this is profitable for the firm *and* is likely to generate consumer surplus even during the period of patent protection, this can be expected to be socially desirable.

So far so good. Note, however, that if patent length is increased still further, giving $S(T) = S_3$ in Figure 21.2, both firms will be induced to undertake the innovation and we have socially wasteful innovation. In other words, we have yet another reason for being cautious in the design of patent policy. Very lengthy patents are attractive to firms and so may induce socially wasteful patent races.

**21.2**

**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, thirty 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 both labs are successful, then each firm will earn a profit (before paying for the research) of \$138.89. Consumer surplus in this case will be \$277.78
c. Now show that the expected profit to BMI (or ECN) if it is the only firm to establish an R&D division is $\$250 - K$ while the expected profit to each firm if they both establish R&D divisions is $\$138.89 - K$. 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 \$416.67. Hence, show that the second lab is socially desirable only if $K < \$141.67$.

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 lead firms to pursue more risky innovations. The intuition behind this argument can be illustrated fairly simply.

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

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 also assume that each discovery generates the same profit of $1 million per period during the time that it is utilized and is protected from imitation by a patent.

Because 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 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.[4] The reason is once again that 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, because 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.

## 21.4 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 firms may hold a large number of patents all related to the same process or product, many of which are never acted upon. (Return to Table 21.1 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"?

One important motivation for holding 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 forty-eight patented new products, that 60 percent were imitated within four years of 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

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

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.

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 its hold on its monopoly position. The question 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 its 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 its 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 its 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 it earned previously but still better than nothing. From this it should be clear why the monopolist places a greater value on discovering the alternative process than does the entrant. Under Bertrand competition, the entrant never earns 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 limit of \$700 on the incumbent's profit. So, it is worth $\$900(= \$1600 - \$700)$ to the incumbent to acquire the process first and thereby prevent the imposition of this profit ceiling.

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, $\pi^m(c_I = 20) = \$1600$, 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 its current profit, taken to be zero.[5] Therefore, the entrant's net gain from developing the technology is \$400. Hence, in the Cournot case also, 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 it wins the race, it gets to keep its current monopoly position. If the entrant

[5] 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.

**Table 21.2** Patent use by inventor's employer

| | *Internal Use* | *Licensing* | *Cross-Licensing* | *Licensing & 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, P., Mariani, M., et al (2005) p. 20.

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 21.2 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.[6] The survey asked the inventors to rate the importance that they put on different motives for patenting.

In this table, "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 the table). 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.

Historic examples of an incumbent using a "sleeping patent" strategy to inhibit rival expansion are common. 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, stage 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.

## 21.5 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, because this implies a "price for information" of near zero, it leaves little incentive for anyone to produce new

[6] 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.

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 US 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.

### 21.5.1 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 because then it will be unable to compete as 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.[7]

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 earns a profit 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 each. Profit to the innovating firm increases to \$900 while profit to each of the other firms falls to \$100.[8]

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 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. The rivals' profits are just as high and the innovator's profit is higher with licensing rather than without it.

Licensing is, indeed, potentially quite profitable. Moreover, the licensing fee of \$10 that we have chosen is not the best that the innovating firm can do.[9] We show in the Appendix 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 because 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.

7 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.

8 These numbers come from simple application of the equations for the Cournot-Nash equilibrium that we have developed in previous chapters.

9 For details see Katz and Shapiro (1985).

### 21.5.2 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 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 its 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 itself, there is no other firm to which it can license!) It should be clear that in such cases the innovating firm will license the monopolist. Because the patent holder is not active in the market, the only way it can obtain any revenue from its 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 it charge a fixed fee independent of output? Or should it 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 16, 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 its 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.

#### *Patent Licensing, Social Welfare, and Public Policy*

The foregoing cases indicate that most of the time an innovator has an interest in licensing its 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 occur unless it raises profit for both the innovating firm and in total. 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 because 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 reduces wasteful R&D that either duplicates existing research efforts 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 in order to enter has to incur an R&D expenditure of \$3 million to develop its own product alternative. In the absence of any licensing, the entrant will pursue this investment because it yields a net gain of \$2 million. Yet if this is the case, then the monopolist firm knows 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 \$3 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 because, 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 wants 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 that 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 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 US 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.

### 21.5.3 Patent Thickets and Sequential Innovation

Indeed, the increasing complexity of technical advances has resulted in what some call a "patent thicket." [See Shapiro (2000).] 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 be built upon 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 described in Chapters 8 and 16, among others.

To understand the patent thicket issue better consider a product that is produced competitively but requires the input of three different patented inputs. That is, one unit of each input is required to make the final product. While each input is freely available on the market at cost $c$, the patent owner establishes a patent fee per $f_i$ unit of the associated input, so that the final price for each input to the producers: $p_1 = f_1 + c$; $p_2 = f_2 + c$; and $p_3 = f_3 + c$. Demand for the final product is linear and given by $P = \mathrm{A} - \mathrm{B}Q$. Because the product is produced competitively, the final product price $P = p_1 + p_2 + p_3$. We may thus write the demand equation as: $p_1 + p_2 + p_3 = \mathrm{A} - \mathrm{B}Q$

We imagine that each patent holder takes the price of its rivals as given. Because every unit of final product demand translates into exactly one unit of patented input demand for each input we may write the inverse demand facing each patent holder as:

$$\begin{aligned}
&\text{Patent Holder 1}: p_1 = \mathrm{A} - p_2 - p_3 - \mathrm{B}Q\\
&\text{Patent Holder 2}: p_2 = \mathrm{A} - p_1 - p_3 - \mathrm{B}Q\\
&\text{Patent Holder 3}: p_3 = \mathrm{A} - p_1 - p_2 - \mathrm{B}Q
\end{aligned} \tag{21.9}$$

It follows that the profit from patent license fees for each patent holder is:

$$\begin{aligned}
\pi_1 &= f_1 q_1 = f_1 Q = (p_1 - c)Q\\
\pi_2 &= f_2 q_2 = f_2 Q = (p_2 - c)Q\\
\pi_3 &= f_3 q_3 = f_3 Q = (p_3 - c)Q
\end{aligned} \tag{21.10}$$

The usual profit-maximizing condition that marginal revenue equal marginal cost must apply. So, we can write

$$\begin{aligned}\mathrm{MR}_1 &= \mathrm{A}-p_2-p_3-2\mathrm{B}Q = c\\ \mathrm{MR}_2 &= \mathrm{A}-p_1-p_3-2\mathrm{B}Q = c\\ \mathrm{MR}_3 &= \mathrm{A}-p_1-p_3-2\mathrm{B}Q = c\end{aligned} \tag{21.11}$$

If we sum up the three equations in (21.11), we then have:

$$3\mathrm{A}-2p_1-2p_2-2p_3-6\mathrm{B}Q = 3c = \overline{c} \tag{21.12}$$

where $\overline{c}$ is the actual unit cost of producing the product in the absence of any patent fees. Given the inverse demand and the fact that competition implies $2(p_1 + p_2 + p_3) = 2P$, we may rewrite equation (21.12) as

$$3P-2P-3\mathrm{B}Q = P-3\mathrm{B}Q = \overline{c} \tag{21.13}$$

Hence, we have:

$$\frac{P-\overline{c}}{P} = 3\mathrm{B}\left(\frac{Q}{P}\right) = 3\frac{\Delta P}{\Delta Q}\left(\frac{Q}{P}\right) = \frac{3}{\varepsilon} \tag{21.14}$$

Equation (21.14) says that the industry price-cost margin as measured by the Lerner Index introduced in Chapter 4 will be three times the inverse of the market elasticity of demand $\varepsilon$. Yet we know from that chapter as well as from our work in Chapter 16 that an integrated firm that owned all three patents would set a price such that the Lerner Index would be:

$$\frac{P-\overline{c}}{P} = \frac{1}{\varepsilon} \tag{21.15}$$

In other words, with three independent patent owners each holding a critical patent, the final markup is three times that which would maximize the industry's joint profit. Of course, this much higher price hurts consumers as well. Moreover, the result in equation (21.14) readily generalizes. If $n$ firms each hold an essential patent, then the final price-cost margin rises to $n/\varepsilon$ or $n$ times the optimal amount. Clearly, this patent thicket creates a major obstacle to production.

There are a number of responses to the patent thicket. One is to permit the patent holders to form a patent pool that permits others to acquire the use of all the patents for one overall fee—in effect, producing the integrated solution. Another is to permit cross-licensing agreements by which firms agree to license their patents to each other. However, these all 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. Beyond these cooperative arrangements, though, another alternative is simply to recognize that the intricate web of interconnected patents may be the result of policy that grants patents too liberally. Hence, it may be that a weak patent policy in which patents are only granted for the clearest and most dramatic of innovations is best. Such a weak patent policy becomes even more attractive when patents are sequential so that each patent builds on an earlier one as argued by Bessen and Maskin (2009).[10]

[10] See Lerner and Tirole (2004).

To get a sense of the Bessen and Maskin (2009) model, we consider the case of a three-period duopoly under two regimes—one with clear and strongly enforced patents and one with no patents in which case competitive imitation is quick to emerge. We will imagine that a firm that does R&D in any period incurs a cost of $c$ and, with probability $p$, creates an innovation that raises the market value of a product by $V$. This is true in each period. Thus, if R&D is successful in period 1, it creates a total additional value of $V$. A second innovation in period 2 would then add a further $V$ in value *if combined with the product now enhanced by the first innovation*. Thus, the sequential nature of the innovative process is embodied in the fact that no innovation can take place in period $t + 1$ (or later period) *unless* there is an innovation in period $t$, e.g., digital music players and "Bluetooth" digital music speakers cannot be created unless someone first creates a process for digitalizing music.

To keep matters concrete, we will work with a numerical example in which $V$ is \$200; the probability of successful R&D by any one firm is $p = 0.5$; and that the cost of R&D is $c = \$25$. Thus, the net expected social value of one firm doing R&D is $pV - c = \$75$.

To drive the point home as clearly as possible, we assume that under a patent regime the first patent winner has to be paid for rights to use the enhanced product and all bargaining power resides with this original innovator so that it can claim as payment the value of any subsequent improvement. For example, suppose that firm 1 (alone) successfully innovates in period 1. If firm 2 then makes a breakthrough in period 2 worth $V$ dollars, it would only be able to apply this innovation to the product by also employing firm 1's patent for which it would have to pay a fee. Because firm 2 has already sunk the R&D cost of $c$, there is nothing to prevent firm 1 from setting a fee equal to the full value $V$ of the new innovation. Yet that would mean that firm 2 would foresee that its risky research effort would ultimately not be profitable and therefore not engage in research. The implication is that in our model, a patent regime results in just one firm doing R&D once a breakthrough has been made.[11]

Now consider the first of our three periods and assume there is a breakthrough. If there is a patent regime, then as just noted, the innovating firm will be the only firm that then does R&D in either period 2 or period 3. What if there are no patents?

In the case of no patents, we assume that any innovation by one firm is followed immediately by imitation by its rival leading to competition and that each firm receives the payoff $SV$, where $S < 0.5$ to reflect the fact that competition results in some of the surplus going not to the firms but to consumers. For our purposes, we will assume that $S = 0.4$.

We start with the third and final period. At that time, each firm will have to decide whether to incur a cost $c = \$25$ to engage in research with a potential social payoff of $V = \$200$ and for which no prior inventor can claim any fee. However, if a breakthrough is made by say firm 1, then the rival firm 2 imitates and competes with the result that the revenue flow to both firms is $SV = \$80$. Because $pSV - c = \$15$, we know that at least one firm will find it profitable to pursue the innovation. Is there a possibility that both firms will do so?

Recalling that if it fails, a second innovator can still copy rival success and letting either firm's R&D success be independent and equal to 0.5, the answer is no. If a second firm spends \$25 on R&D, it will be successful half the time and earn \$80. It will fail the other half but in half of these cases it can copy its successful rival. So the expected net gain of a second firm's R&D expense is:

$$[p + p^2]SV - c = 0.75 \times 0.4 \times \$200 - \$25 = \$35 \tag{21.16}$$

[11] More generally, firm 1 will only permit firm 2 to use its patent if the joint profit of both firms with both patents in use exceeds firm 1's monopoly profit and there is no reason to assume that this is the case.

However, with no patents, instead of spending on R&D like its rival, a firm can choose not to do so and simply imitate any successful innovation that the rival develops. In this case, the imitating firm has expected profit

$$pSV = 0.5 \times 0.4 \times \$200 = \$40 \tag{21.17}$$

As the payoff is greater to imitating rather than being a second innovator, the period 3 outcome in the no-patent regime must be one in which one firm does R&D and the other does not. There are of course two such equilibria—one in which firm 1 does the R&D and one in which firm 2 does the R&D. We will imagine that this choice is made by random assignment, i.e., a coin toss.

Now consider period 2. In that period, we know that at least one firm will innovate because again, $pSV - c = \$15$. In considering whether a second firm will innovate or simply imitate though, the calculation now is a little changed. This is because a breakthrough in the second period helps *both* firms in the third period because that is what makes possible an innovation at that later time. Suppose for example that firm 1 pursues R&D in period 2, but firm 2 does not. Then firm 2's expected profit over both periods 2 and 3 is:

$$\begin{aligned} E(\pi_2) &= 0.5 \times 0.4 \times \$200 + 0.5 \times 0.5[(0.5 \times 0.4 \times \$200 - \$25) + (0.5 \times \$0.4 \times \$200)] \\ &= \$53.75 \end{aligned} \tag{21.18}$$

Firm 2's expected profit if it imitates in period 2 has two parts. The first is the probability that firm 1 makes a breakthrough in period 2, which happens with probability 0.5 and in which case firm 2 imitates and earns \$80. The second term is firm 2's expected third-period gain. Getting to that period with any opportunity for further innovation is simply the probability $p = 0.5$ that firm 1 is successful in period 2. In that third period however, we know that only one firm pursues R&D. Given that the chances of success are $p = 0.5$, that firm has expected profit of $0.5 \times 0.4 \times \$200 - \$25$, and its rival has expected profit $0.5 \times 0.4 \times \$200$. Of course, firm 2 has a fifty percent chance of being either one of these firms in period 3.

If on the other hand, firm 2 decides to follow firm 1 and pursue R&D in period 2, its expected profit over the remaining two periods is:

$$\begin{aligned} E(\pi_2) &= 0.75 \times 0.4 \times \$200 - \$25 + 0.75 \times 0.5[(0.5 \times 0.4 \times \$200 - \$25) + (0.5 \times \$0.4 \times \$200)] \\ &= \$55.625 \end{aligned} \tag{21.19}$$

Here again, firm 2's expected profit has two components. The first is the now higher probability that there is a breakthrough in period 2 times the revenue that brings and less the cost of doing research *plus* the also higher probability that there is opportunity for a further breakthrough in period 3 times the expected net gain for firm 2 in that period which depends on whether it is the one firm then doing any R&D.

As can be seen, firm 2 does better by investing in research in period 2 and thereby adding its innovative efforts to those of firm 1 rather than simply imitating its rival. Thus, conditional on a breakthrough in period 1, the no-patent regime leads to both firms doing research in period 2 and only one doing it in period 3 while the patent regime leads to only one firm doing research in both periods 2 and 3, i.e., the no-patent regime yields more R&D following the initial period than the patent regime does. Moreover, this result easily generalizes to periods 2 and 3 in a four-period model; periods 2, 3, and 4 in a five-period model; and so on. The intuition is straightforward. By adding to the innovation effort of its rival, firm 2 not only increases the chance of an innovation in that period but also to the chance of an innovation later which helps firm 2 even if it plans only to imitate at those later times. This makes doing R&D more profitable.

We have not addressed the first period. It is easy to guess though that a strong patent system will generate as much if not more R&D effort as a weak one in that first interval. However, the sequential nature of innovation suggests that once an innovation occurs, the obstacles that a strong patent system and associated patent thicket creates may well lead to *less* innovative activity in every subsequent period. This conclusion follows directly from the sequential nature of innovation. Hence, to the extent that innovation has that sequential feature, there is again concern that recent patent policy has erred too strongly in the protection of intellectual property.

**Practice Problem 21.3**

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?

## 21.6 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 United States. The first and perhaps most crucial step was a legal reorganization that gave the Court of Appeals for the Federal Circuit (CAFC), in 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 US 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. Because shutting down a high volume production line is very expensive—even if only

## 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 be ideas for using the Internet profitably. These business method innovations differ from manufacturing technological advances but may be even more valuable ever since the granting of patent protection to business methods with the 1998 U.S. Court of Appeals ruling that legitimized 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 on-line giant Amazon, for example. Its 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 processed. Amazon actually received a patent for this 1-Click feature and touts it to all potential customers. How effective that patent is was revealed in 1999 when traditional "brick and mortar" bookseller, Barnes & Noble, added internet selling and included a one-click feature in its web Express Lane checkout. Amazon successfully sued for patent infringement and Barnes & Noble had to move to finalizing an order in two clicks.

The 1-Click case is not unique. A customer at Burger King, for instance, might order a Whopper, 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 so-called upsale, Burger King receives more revenue. Yet, a good chunk of the extra funds will go to Walker Digital as a licensing fee because Walker owns a patent on this process.

In the wake of the *State Street* decision, filings for business method patents nearly tripled. Many economists, including Gallini (2002) and Hall (2003), suspect business method patents may slow down technical progress. It would be ironic if at the same time that technical breakthroughs made information cheap, expanded business method patents made information more expensive to exploit.

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

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 United States 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 19th 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 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 US jury found the Canadian firm guilty of sixteen counts of patent infringement. On appeal, seven of these were dismissed in 2004 but that still left nine 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 US 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 of 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 both for its unanimity and for its clear concern that patent system abuse could undermine innovation. 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

## Reality Checkpoint

## It Was Patently Obvious and Therefore, Not Patent Worthy

In April, 2007, the US Supreme Court issued an important ruling that substantially raised the bar for 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. KSR won a 1999 contract with General Motors to provide such an adjustable pedal with an electronic sensor on GM vehciles. Teleflex, which had a patent for a specific type of electronic pedal sensor claimed infringement and demanded royalties. KSR refused to pay arguing that the Teleflex patent was invalid because it combined elements of exisiting sensors in an obvious manner. KSR won in District Court but lost on Tlelefex's appeal to the CAFC, the main patent appellate court.

Although the Supreme Court had a history of letting CAFC judgments stand unreviewed, in this case the Court took issue with the CAFC and especially, its use of the so-called "teaching, suggestion, or motivation" test (TSM test). This test only finds an invention obvious if "some motivation or suggestion to combine the prior art teachings" can be found in the previous work. Thus, the TSM test would imply that even when a new device yields very predictable results that any professional could logically foresee, the device would nonetheless be patentable unless that expecation had actually been tested in practice. The Supreme Court found that this test awarded protection too easily. It 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 was 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.

Sources: *KSR International Co. vs. Teleflex Inc.*, 550 U.S. 398 (2007); *Microsoft Corporation v. AT&T Corp.* 550 U.S. 437 (2007); and L. Greenhouse, "High Court Puts Limits on Patents," *New York Times* 1 May 2007.

leeway to dismiss patent infringement suits. Further evidence that the Court had moved to a less protective patent policy came in June, 2013 when it rules that firms could not legally patent human genes.

## 21.7 EMPIRICAL APPLICATION: PATENT LAW AND PATENT 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 $20^{th}$ 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 life cycles 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 a fixed time interval as:

$$f(\lambda, p) = \frac{e^{-\lambda}\lambda^{p}}{p!} \tag{21.20}$$

The Poisson distribution has a very nice feature in that it is fully characterized by the parameter $\lambda$, 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 $\lambda_{it}$ that is an exponential function of $X_{it}$ as follows:

$$E(p_{it}|X_{it}) = \lambda_{it} = \exp(X_{it}\beta + \gamma_t) \tag{21.21}$$

where $\gamma_t$ 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:

$$\ln \lambda_{it} = X_{it}\beta + \gamma_t \tag{21.22}$$

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 $\lambda$ is taken to be conditional on a set of firm characteristics $X_{it}$ and time dummies $\gamma_{it}$. 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 (MLE). Because $\lambda$ 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 21.3, below.

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.

**Table 21.3** 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) |

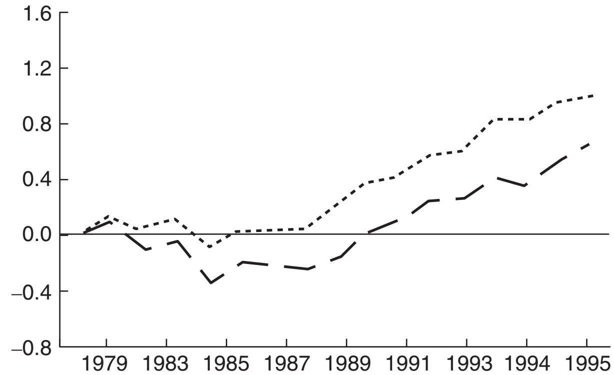


**Figure 21.3** Pattern of regression time coefficients in semiconductor patent behavior

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 shown in Figure 21.3. 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.

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 firms 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 allow an innovator to earn a greater return on its R&D investments than it 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. It also reflects 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 at that time.

More recently, there has been concern that US patent protection has been overly strong, especially because there is little evidence that it has led to faster innovation. In particular, patent decisions may have created a thicket of necessary approvals and fees that strongly deters new product development. This is particularly true if patents have a sequential feature such that one patent builds on those before it. Accordingly, recent court decisions have pared back these protections. However, as our discussion of licensing and recent legal developments makes clear, there is no easy way to eliminate the tension between allocative and innovative activity that a patent system is meant to balance.

## 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 $\rho$.
   a. Assume that the new process is expected to have marginal costs of \$70. Derive a relationship between $K$ and $\rho$ 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 $\rho$ 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 1(a)?

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 4, 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?

## References


Bessen, J., and E. Maskin, 2009.

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.

Gallini, N. 2002. "The Economics of Patents: Lessons from Recent U.S. Patent Reform," *Journal of Economic Perspectives* 16 (Spring): 131–154.

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

Giuri, P., 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): 113–30.

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 No. 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: 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: MIT Press.

Shapiro, C. 2000. "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting," in *Innovation Policy and the Economy*, A. Jaffe, J. Lerner, and S. Stern, eds. National Bureau of Economic Research, 1190–1250.

## Appendix

Here, we derive the optimal license fee for the simple case in the text. Demand is $P = A - BQ$, the initial marginal cost is $c$' and the innovation lowers marginal costs to $c$. For any royalty per unit $r \leq c' - c$. the innovator's profit is:

$$\pi = \frac{(A - Nc + (N-1)(c+r))^2}{B(N+1)^2} + r(N+1)\frac{(A - Nc + (N-1)(c+r))}{B(N+1)} \qquad (21.A1)$$

Since $(A - Nc + (N-1)(c+r)) = A - c + (N-1)r$, it is evident that the expression in (21.A1) is increasing in $r$. Hence, profit maximization implies raising $r$ as high as possible until the maximum $c' - c$ is reached.

# Part Seven

# Networks, Auctions, and Strategic Policy Commitment

In this final part of the text, we explore topics that do not fit easily within our earlier classifications. The first of these is network externalities, which are the topic of Chapter 22. For many goods, such as telephones, the value of the product to any one consumer rises as additional consumers buy it. Such network effects greatly alter both the nature of industry competition and the characteristics of the market outcome. Often, network externalities and the complementarities that underlie them give rise to multiple equilibria with no guarantee that the actual equilibrium chosen will be the best of these. Further, because network externalities act much like scale economies except that they work on the demand side, they create strong incentives for firms to operate on a large scale with the result that the market will inevitably be dominated by those few firms that survive. In turn, because not just some profit but a firm's very survival may be at stake, competition in industries with important network effects can be incredibly fierce. We explore these issues in some detail. We also include a discussion of Gandal's (1994) empirical study that tries to identify network characteristics in the market for computer spreadsheets.

In Chapter 23, we switch gears somewhat and turn to the topic of auctions. It has only been in recent years that auctions have re-emerged as a common market arrangement, especially for government privatization efforts and private exchanges on the Internet. Our analysis begins with a review of Vickery's (1961) classic piece leading to the Revenue Equivalence Theorem, which says that, under certain rather broad conditions, the final auction price is independent of the auction design. We then examine various ways in which this outcome might break down when bidders have affiliated values or signals. Perhaps most importantly, we show that the insights of auction theory may be usefully applied to industrial organization topics such as Bertrand pricing. We subsequently offer the Porter and Zona (1999) study of Ohio school milk auctions as an example of careful application of auction models to uncovering collusive behavior.

Finally, Chapter 24 applies the insights of industrial organization to international trade. We explore how knowledge of the Cournot and Bertrand models can help to develop strategic trade policies that depart from the economist's traditional free trade recommendations. This analysis allows us to make clear the critical role of commitment in any strategic setting. We conclude with an investigation of the use of strategic trade policy by the Canadian Wheat Board along the lines of a Cournot model as suggested by Hamilton and Stiegert (2002).

# 22

# Network Issues

On Thursday, November 15, 2001, Microsoft CEO Bill Gates introduced the video game console, *Xbox*. Many questioned the decision to launch this new entry in a market that was then dominated by Sony and Nintendo, and that also included Sega, another firm with a sizable installed base. Today, however, the *Xbox* brand is still very much alive. The *Xbox 360* (introduced in 2005) is the best-selling video game console in the world and claims a 47 percent share of the North American market—well above rivals Sony (*PS3*) and Nintendo (*Wii*). Sega has exited the market altogether. All of the current ten best-selling games from *Mass Effect* to *Kingdoms of Amalur* to *Just Dance* are available in *Xbox* format. In 2011, *Xbox* consoles sales totaled $2.1 billion alone, and Microsoft earned an additional $4.6 billion from sales of games and accessories. Current estimates imply that 67 million *Xboxes* have been sold since that November 2001 launch.

The *Xbox*, of course, is just one of many computer-based success stories associated with Microsoft. Unlike most of these, however, the *Xbox* is a hardware product. Rather than Microsoft, the major providers of the software that runs on the *Xbox* platform are game development firms such as Electronic Arts and Activision Blizzard. Indeed, both the game developers and Microsoft are tied in a sort of "chicken-and-egg" relationship. There is no way to market *Xboxes* successfully unless developers produce games for that console. Yet there is no reason for developers to produce games for the *Xbox* unless a lot of *Xboxes* are sold. Put differently, the more *Xboxes* are sold, the more game-makers want to develop *Xbox* compatible games, and the more *Xbox* compatible games there are the more consumers want to buy the *Xbox*. There is then an external effect to the consumers' decision. As more consumers buy the *Xbox*, more games are developed for that platform, raising the value of the *Xbox* for everyone.

When the value of a product to any one consumer increases as the number of other consumers using the product increases, we say that the market for that product exhibits network externalities or demand-side scale economies. Video game consoles exhibit such network effects but they are far from the only such goods and services. Telephone systems are perhaps the most obvious of the other examples. By itself, one person's ownership of a telephone is not very valuable. Yet as more and more phone users hook into the system, the value of a connected phone rises for all. The online auction firm eBay offers yet another illustration of network effects. Prior to the 1990s, direct trade in many items, especially collectibles, had been limited because of the extreme cost of matching a potential buyer with a potential seller. eBay founder, Pierre Omidyar, was among the first to recognize the enormous potential of the Internet—which makes it easy to disseminate a vast amount of information to a large number of buyers and sellers in a very short time—to solve this problem. As more buyers came to look for goods on eBay, more sellers found it worthwhile to sell there which, in turn, attracted more buyers and so on.

The telephone and eBay cases are examples of what are typically called *direct* network effects. In contrast, the video game console market illustrates an *indirect* network effect. Operating systems such as *Windows* and *Mac* also benefit from indirect network externalities. Analogous to the console, these systems become more popular as more software is developed to run on them while, in turn, more compatible software is developed as the systems become more popular. The difference between direct and indirect network effects should thus be clear. In the direct case, the effects pertain directly to the good or service, e.g., the telephone, itself. In indirect case, the network effect is mediated by a complementary good. In this chapter, we investigate these issues and the equilbria that are likely when important network effects are present.[1]

## 22.1 MONOPOLY PROVISION OF A NETWORK SERVICE

An early but insightful analysis of network issues is that provided by Rohlfs (1974). Rohlfs' approach is quite straightforward and draws attention to the main issues that arise in network settings. It simplifies the supply side by assuming a monopoly so that the analysis can focus on the central demand-side aspects that give rise to network effects. We present a modified version of Rohlfs model here.

Assume that the monopolist, say a telecommunications firm, charges an access fee but does not impose a per usage charge. That is, the consumer is charged a single price $p$ for "hooking up" to the network but each individual call is free, perhaps because the marginal cost of a call is zero.[2] We will also assume that there is a maximum size of the market, say one million, reflecting the maximum number $N$ of consumers who would ever willingly buy the product even if the access fee were zero. By fixing the total amount of potential customers at $N$, we can talk interchangeably about the fraction $f$ of the market and the actual number of customers $fN$ served. That is, if the maximum size of the market is one million, we can characterize a market outcome in which 100,000 purchase the service either in terms of the fraction $f = 0.10$ or the number of customers $fN = 100{,}000$ served.

Consumers all agree that the service is more valuable the greater the fraction $f$ of the market that signs up for it. However, even if everyone acquires the service ($f = 1$), consumers would still vary in their valuation or willingness to pay for the service. Specifically, we denote the valuation of the $i$th consumer when $f = 1$, as $v_i$. These valuations or $v_i s$ are assumed to be uniformly distributed between 0 and \$100. For example, the 1 percent of consumers who most value the service (roughly about 10,000 individuals in our case) would willingly pay close to \$100 for it if they knew that all other consumers would also acquire it. Unfortunately, when the consumer signs on, the consumer does not know the actual fraction $f$ of potential users that will sign on to the network. Instead, the consumer's willingness to pay depends on the fraction $f^e$ of users the consumer *expects* to sign on in that the *ith* consumer's willingness to pay for network membership is assumed

[1] For a formal but very readable introduction to network externalities, see Economides (1996).
[2] Note that this pricing policy is essentially that of a two-part tariff as described in Chapter 6.