22

Research and Development

The final results of the human genome project indicate that we humans are not as complicated as we thought we were. Rather than consisting of the approximately 100,000 genes that were initially predicted, it appears that we have only 30,000 genes, not much more than twice as many as the humble roundworm with its 19,098 genes.[1] This finding is important for many reasons. From our perspective, there is an important economic aspect to this result. It is well understood that genes are a crucial factor in predicting and curing many diseases. Therefore, identifying and understanding the workings of each gene could lead to the creation of a new family of custom-made drugs. The rough equation quoted by the pharmaceutical companies was "one gene, one patent, one drug."[2] If, as initially expected, there were 100,000 genes then there was potentially a vast number of revenue-generating patents. The finding that the actual number of genes is far less than 100,000 has suggested to many that genes hold many fewer of the keys to the treatment of disease. As a result, understanding genes and their functions may offer a much less lucrative source of new patentable treatments.

However, all is not lost. It is being suggested that much of human biology is determined at the protein level rather than at the DNA level, and we have well over 1,000,000 different proteins in our bodies. So now we have a whole new science, proteomics—studying how genes control proteins—as a method for creating tailored drugs. Proteomics is being pursued by an increasingly wide number of companies and institutions: Harvard University, for example, has created a new Institute of Proteomics.

The race to understand the proteomic causes of diseases and to develop new drugs targeted at those diseases will not come as a surprise to anyone familiar with the popular business literature of the past twenty years. That literature is characterized by the dominant theme that the most successful firms find new ways of doing things, or develop new products and new markets.[3] The now prevalent view is that firms become industry leaders by conducting research and development (R&D) leading to innovations in their production technologies or

[1] If you are interested, the complete human genome is available as a free download from http://gdbwww.gdb.org/.

[2] "Scientists, Companies Look to the Next Step After Genes," *New York Times*, February 13, 2001.

[3] This is virtually the mantra in the best-selling book by Peters and Waterman, *In Search of Excellence: Lessons from* ***America's*** *Best Run Companies* (1982). However, the argument is repeated frequently in other business books, including, as noted herein, Porter's (1990) encyclopedic volume.

the products they provide. Michael Porter's *The Competitive Advantage of Nations* (1990) serves to make the point. Porter writes that any theory of competitive success:

> must start from the premise that competition is dynamic and evolving . . . Competition is a constantly changing landscape in which new products, new ways of marketing, new production processes, and whole new market segments emerge . . . [Economic] theory must make improvement and innovation in methods and technology a central element. (p. 20)

Porter's quote could almost have been taken verbatim from Joseph Schumpeter's classic work written almost fifty years earlier. Schumpeter was both an economist and a historian. He brought a historical perspective to his study of competition and the rise and fall of corporate empires. The following dramatic passage appears in his book *Capitalism, Socialism, and Democracy* first published in 1942.

> [I]t is not . . . [price] . . . competition which counts but competition from the new commodity, the new technology, the new source of supply, the new type of organization . . . competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and outputs of existing firms but at their foundations and very lives. (p. 84)

Interest in the forces behind innovative activity is perhaps stronger today than it was when Schumpeter wrote.[4] An important issue, raised by Schumpeter, concerns the market environment most conducive to R&D activity. Schumpeter conjectured that R&D efforts are more likely to be undertaken by large firms than by small ones. He speculated secondly that monopolistic or oligopolistic firms would more aggressively pursue innovative activity than would firms with little or no market power. Accordingly, Schumpeter argued that the benefits of an economy comprised largely of competitive markets populated by small firms reflected the rather modest gains of allocating resources efficiently among a *given set of goods and services produced with given technologies*. In contrast, the benefits of markets dominated by large firms, each with sizable market power, stems from the much larger dynamic efficiency gains of developing new products and new technologies. As Schumpeter wrote, "a shocking suspicion dawns upon us that big business may have had more to do with creating (our) standard of life than with keeping it down" (p. 88).

The validity of Schumpeter's ideas—which have come to be jointly referred to as the Schumpeterian hypothesis—is the key issue addressed in this chapter. Do larger firms do more R&D? Does a concentrated market structure provide a better environment for the development of new innovations than a competitive structure? Table 22.1 lists the ten companies awarded the most patents by the U.S. Patents and Trademark Office (USPTO) in 2006 as well as their ranks in 2005 and 2004. Each of these is a large company. Most operate in oligopolistic markets with only a few large competitors. Moreover, there is considerable stability in the rank ordering, at least over these three years. It is tempting to conclude on the basis of such data that Schumpeter was right and that large firms in concentrated markets are more innovative. However, great care is needed before reaching that conclusion. Rather than implying that large firms do more R&D, these results could imply that firms that do more R&D become large.

[4] For example, see the story by C. J. Whalen, "Today's Hottest Economist Died 50 Years Ago," *Business Week*, December 11, 2000. Ever since Solow's (1956) classic work, macroeconomists studying growth have also focused intensively on technological progress and innovation as the primary source of improved living standards over time. See, e.g., the books by Barro and Sala-I-Martin (1995) and Romer (2006).

Table 22.1 Top ten U.S. patent-receving firms in 2006 and their rank in 2005 and 2004

Company	*# of patents in 2006*	*Rank in 2005*	*Rank in 2004*
International Business Machines	3,621	1	1
Samsung Electronics	2,451	5	3
Canon Kabushiki Kaisha	2,366	2	4
Matsushita Electric Industrial	2,229	4	2
Hewlett-Packard	2,099	3	6
Intel Corporation	1,959	7	5
Sony Corporation	1,771	12	7
Hitachi	1,732	8	8
Toshiba Corporation	1,672	9	9
Micron Technology	1,610	6	11

Source: USPTO

Table 22.2 Top ten patent-receiving industries in 2006 and cumulative patents to that year

Industry class	*Patents granted in 2006*	*Cumulative patents granted*
Semiconductor device manufacturing: process	4,467	50,224
Active solid-state devices (e.g., transistors, solid-state diodes)	4,287	42,436
Drug, bio-affecting and body treating compositions	2,784	73,234
Multiplex communications	2,754	25,157
Chemistry: molecular biology and microbiology	2,423	46,466
Telecommunications	2,271	20,624
Stock material or miscellaneous articles	2,263	54,326
Static information storage and retrieval	2,019	24,710
Optical: systems and elements	2,013	28,024
Computer graphics processing and selective visual display systems	2,004	22,507

Source: USPTO

The most active areas for research activity are likely to vary over time. Table 22.2 lists the top patent-receiving industries or research areas in 2006. It also shows the cumulative patents in that area up to that year. While there is some consistency across the two columns there is also considerable variation. Thus, while semiconductor and solid-state devices have led the patent parade in recent years, bio-science drugs and molecular chemistry have accounted for many more patents in total over time.

Introducing a new product can often undermine the marketability of the firm's existing products. Similarly, the development of a new production process requiring new equipment reduces the value of existing productive capacity. Because introducing new products or processes inevitably means the destruction of old ones, Schumpeter dubbed such competition by innovation "creative destruction." In addition, since some of the products and processes that are made obsolete may well be those of the innovating firm itself we can ask our central question in a somewhat different way. Why do firms undermine existing activities (including

Reality Checkpoint

Creative Destruction in the Pharmaceutical Industry: Will the Prozac Work if the Viagra Fails?

Perhaps no market offers better examples of Schumpeter's "creative destruction" than that for pharmaceuticals. Consider the market for antidepressants. For several years after its introduction in 1987 by Eli Lilly & Co., Prozac dominated this market. Originally envisioned as a treatment for high blood pressure and, when that failed, an anti-obesity drug, Lilly was pleasantly surprised when hospital tests on mildly depressed patients showed a marked and widespread positive effect. Lilly took its fluoxetine drug as it was then called, and asked Interbrand, one of the major branding companies in the world, to develop a new name and sales campaign. Prozac was born and soon it dominated the antidepressant market. By the early 1990s, Prozac accounted for nearly a quarter of Lilly's $10 billion revenue.

However, rivals were soon at work inventing around Lilly's patent. In 1992 Pfizer, Inc., introduced a rival Zoloft, which quickly jumped to a third of the market. This was shortly followed by SmikthKline Beecham's Paxil, which soon had 20 percent of the market. All three drugs increased the levels of the neurotransmitter, serotonin, in the brain. But all three had slightly different chemical bases and different side effects. Finally, in 2001, the Prozac patent expired. Within two weeks, prescriptions for generic fluoxetine exceeded those for brand-name Prozac. Within a year, Lilly had lost 90 percent of its Prozac prescriptions. Lilly countered with a new drug, duloxetine, with the brand name of Cymbalta that works on two neurotransmitters, serotonin and norepinephrine.

This brings us to the story of the Viagra, the original drug to combat male impotence patented by Pfizer in 1996. Like Prozac, the drug originally known as sildenafil citrate, was originally envisioned as a treatment for high blood pressure and angina. Even though it failed in that regard, its many male users reported a dramatic increase in sexual function. Pfizer received approval from the Food and Drug Administration (FDA) to market the drug as a treatment for erectile dysfunction and re-launched the drug under the name, Viagra, in 1998. Sales topped $1 billion within a year.

Once again, success brings competition. By 2003, two new drugs were approved by the FDA as treatments for male impotence. These were Levitra (made by Bayer and GlzaxoSmithKline) and Cialis (developed by a small startup firm, ICOS, and marketed by Lilly). Both work in much the same way as Viagra but, again, there are important differences and side effects. Levitra penetrates cell walls in as little as 16 minutes rather than the 30 minutes minimum required for Viagra. Cialis works about as fast as Viagra but has a serum half life of nearly 18 hours. Hence, it can be effective for up to 36 hours or 9 times longer than the typical duration of Viagra. The relative efficacy of these two products has made them fierce competitors to Viagra. Within the U.S., the two drugs combined quickly for as much as 40 percent of the market. In countries such as Australia and France, Cialis alone claimed 40 percent of the market within its first year. Of course, the real competition will start in 2011 when the Viagra patent expires. If the Prozac experience is any guide, competition from generics will very quickly become fierce and Viagra sales will decline even as the general market rises. Drug firms like Lilly and Pfizer could then take a double hit as the change in lifestyles that may then occur may lead to a decline in the demand for antidepressants among both men and women.

Sources: R. Langreth, "High Anxiety: Rivals Threaten Prozac's Reign," *Wall Street Journal*, May 9, 1996, p. A4; A. Pollack, "Lilly Pays Bid Fee up Front to Share in Rival of Viagra," *New York Times*, October 2, 1998, p. C1; S. Carey, "Lilly Reports 22% Decline in Net Income as Generics Hurt Sales of Prozac," *Wall Street Journal*, April 16, 2002, p. C2; and "Viagra Rival Cialis Wins up to 40 Percent Market Share," *Reuters News Wire*, March 23, 2004.

their own ones) in this way? More generally, what are the incentives to engage in innovative activity and how do these vary with firm size and market structure?

Both the professional and the popular business literature have had much to say on the Schumpeterian hypothesis in recent years. In this chapter, we approach this topic using the tools of economic analysis and strategic interaction that we have built throughout the book. However, before we begin a more formal analysis, we need to establish some definitions or classifications to which we can easily refer.

22.1 A TAXONOMY OF INNOVATIONS

Research and development consists of three related activities. The first is *basic research*. This includes studies that will not necessarily lead to specific applications but, instead, aim to improve our fundamental knowledge in a manner that may subsequently be helpful in a range of activities. The derivation and validation of the theory of laser technology is a good example. A second category is *applied research*. Such research generally involves substantial engineering input and is aimed at a more practical and specific usage than basic research. The creation of the first laser drill for dentistry would be an example of applied research. Finally, there is the *development* component of R&D. Here, the goal is to move from the creation of a prototype to a product that can be used by consumers and that is capable (to some extent at least) of mass production. To continue our analogy, the transformation of the first laser drill into a small, handheld product that is affordable and usable by a large number of dentists would be an example of the development stage. For the most part we shall be concerned with applied research rather than development, but we shall touch upon some of the important issues that characterize the decision to move from research to development.

In considering the output of R&D, it is common to distinguish between two kinds. *Process innovations* are discoveries of new, typically cheaper methods for producing existing goods. *Product innovations* are the creation of new goods. For the most part, we shall concentrate on process innovations, but we shall also present examples showing how the analysis can be extended to product innovations.

Finally, with respect to process innovations, there is a further distinction that can be made. This is the division of innovations into *drastic* or major innovations, and *nondrastic* or minor innovations. Roughly speaking, drastic innovations are ones that reduce a firm's unit cost to such an extent that even if it charges the profit-maximizing monopoly price associated with that low cost, it will still undercut all competitors. Hence, a drastic innovation creates a monopolist unconstrained by any fear of entry or price competition—at least for some time. By contrast, a firm making a nondrastic innovation may gain some cost advantage over its rivals but not one so large that the firm can price like a monopolist without fear of competition.

The formal distinction between drastic and nondrastic innovations is illustrated in Figure 22.1. Assume that demand for a particular product is given by $P = 120 - Q$ and that before the innovation all firms can produce the product at a constant marginal cost of \$80. Assume also that the existing firms are Bertrand competitors so that the price is \$80 and total output is 40 units.

Now suppose that one firm gains access to a process innovation that reduces its marginal costs to \$20 as in Figure 22.1(a) and that, perhaps because of a patent, this firm is the only one able to use the new low-cost technology. If this innovator were alone in the market, it would set the monopoly price corresponding to its new, lower marginal costs of \$20. Given our demand function we know that marginal revenue is $MR = 120 - 2Q$. Equating this with

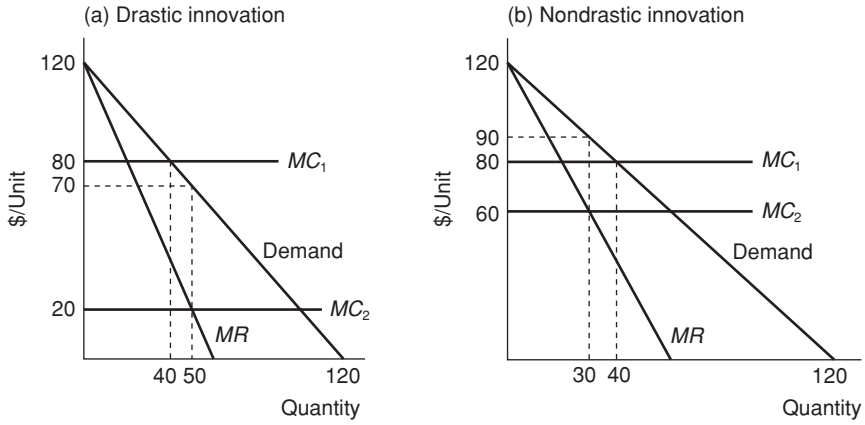

Figure 22.1 Drastic and nondrastic process innovations

marginal cost of \$20 gives an output of 50 units and a monopoly price of \$70. Setting this monopoly price forces all the other firms out of the market. The innovation is a drastic one because the reduction in cost is so great that the innovating firm can charge the full monopoly price associated with the new low cost and still be able to undercut the marginal costs of all other firms.

Suppose by contrast that the innovation reduces marginal costs to \$60 as in Figure 22.1(b). By exactly the same argument as above, the innovating firm acting as a monopolist wants to produce an output of 30 units and set a price of \$90. The problem is that this will not work. The remaining firms can profitably undercut this price. So the best that the innovating firm can do now is to set a price of \$80 (more accurately, \$79.99) and an output of 40 units. This still eliminates the other firms but only by the innovator lowering the price that it charges. Hence, this is a nondrastic innovation.

22.1 Practice Problem

Assume that demand in a competitive market is given by the linear function: $P = 100 - 2Q$ and that current marginal cost of production is constant at \$60. Now assume that there is a process innovation that reduces marginal cost to \$28. Show that this is a nondrastic innovation. How much would the innovation have to lower marginal costs for it to be drastic?

22.2 MARKET STRUCTURE AND THE INCENTIVE TO INNOVATE

We now turn to some of the basic questions economists have asked regarding how the incentives to spend on R&D are affected by market structure.[5] We assume the demand for a particular good is linear. Specifically, the inverse demand curve is again assumed to be given

[5] This analysis owes much to Nobel Prize winner Kenneth Arrow's path-breaking work (1962).

by the equation $P = 120 - Q$. We also assume that each producer of the good has a marginal cost of \$80. Accordingly, if the market is competitive and there are many such producers, the current price is also \$80.

22.2.1 Competition and the Value of Innovation

Suppose that a research firm, not involved in the actual manufacture of this good, discovers a new production process by undertaking research at some cost K. Using the notation from above, we consider the case of a nondrastic process innovation that reduces the marginal production cost to \$60. We further assume that the innovation is protected by a patent of unlimited duration that cannot be "invented around" by other potential or actual firms. What benefits does the innovation bring, and does the market mechanism work to convey such incentives to the research firm?

Let us first consider how society as a whole values the innovation. Imagine a social planner whose goal is to maximize total social surplus (producer surplus plus consumer surplus) and, moreover, who has the power to command that prices be set at whatever level the planner requests. Such a benevolent dictator would reason as follows: With or without the innovation, optimality requires that price be set to marginal cost. The per-period value that the social planner places upon the innovation is the increase in consumer surplus when price equals (constant) marginal cost as then there is no producer surplus. Prior to the innovation, consumer surplus at a price of \$80 is \$800. After the innovation, when firms set the price equal to the new lower marginal cost of \$60, consumer surplus increases to \$1,800.[6] The increase in consumer surplus is \$1,000, the shaded area in Figure 22.2(a). This additional surplus will be realized not just in one period but also in all present and future periods following the innovation. Hence, using the discounting techniques discussed in section 2.2 Chapter 2, the total present value of the additional surplus created by the innovation is $V^p = 1{,}000/(1 - R)$ where $R = (1 + r)^{-1}$ and r is the interest rate. The more this value exceeds the cost K, that is, the more it exceeds the present value of the expenses associated with discovering the process, the more desirable is the innovation.

Of course, we don't have a dictator and if we did it is doubtful that a dictator would succeed in maximizing social welfare. What we have are markets. The issue is how the structure of the market affects the realization of the value of this innovation. What is the incentive of a research firm to pursue the innovation if when it is successful it can auction the rights to the innovation to a competitive industry comprised of many firms. Prior to the innovation all firms sell at a price equal to the marginal cost of \$80 and earn zero profit. Total output each period prior to the innovation is just 40 units.

Now consider the behavior of a firm that has the rights to the innovation? Quite evidently, its best strategy is to undercut its erstwhile competitors just slightly, driving them out of the market and giving it an effective monopoly. The firm that wins the rights will set a price that is one cent less than the old competitive price, \$80. At this price, the industry's total output remains identical to what it was prior to the adoption of the innovation. Consequently, the firm will earn per-period profit of $\$(80 - 60) \times 40 = \$800$. This is illustrated by the shaded rectangle in Figure 22.2(b). The present value that a competitive firm places on the innovation is the maximum amount it willingly bids for the rights to it and

[6] Given our demand function $P = 120 - Q$ and assuming that $P = MC$, consumer surplus is the area of a triangle with height $120 - MC$ and base $120 - MC$. That is, $CS = (120 - MC)^2/2$.

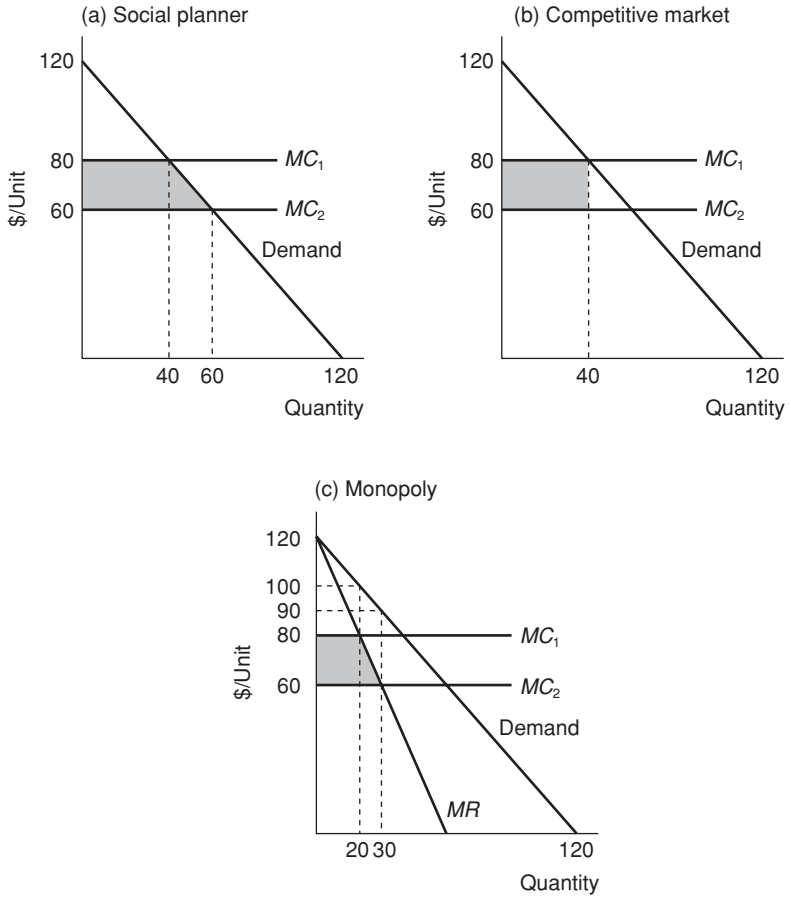

Figure 22.2 Market structure and the incentive to innovate

this is $V^c = 800/(1 - R)$. This is less than the social value of the innovation. The reason is simple: the competitive firm only considers the profit it can earn as a result of the innovation. It ignores the additional benefit from increased consumer surplus that the innovation brings.

Now consider the potential value when it is a monopolist who has the rights to the innovation and who faces no threat of entry. For such a firm, the gain from introducing the innovation is the additional profit it makes as a result of being able to produce at a lower marginal cost. Since the monopolist maximizes profit by setting marginal revenue equal to marginal cost, we can measure this gain by comparing the monopolist's per-period profit at its current marginal cost with its per-period profit at the lower marginal cost that the innovation permits. This is illustrated in Figure 22.2(c).

Given our demand function we know that marginal revenue is $MR = 120 - 2Q$. So, prior to the innovation, the monopolist produces an output of 20 units, sets a price of \$100 and earns profit per period of \$400. After the innovation output is increased to 30 units, price is reduced to \$90 and per-period profits are \$900. As a result, the per-period value placed by the monopolist on the innovation is \$500 – the difference between profits with and without

the innovation. This is illustrated by the shaded area in Figure 22.2(c).[7] In turn, the total present value the monopolist places on the innovation is $V^m = 500/(1 - R)$.

From the foregoing analysis, it is obvious that $V^p > V^c > V^m$. Both the competitive firm and the monopolist undervalue the innovation relative to the social planner interested in maximizing total welfare. However, a competitive firm values the innovation more than the monopolist.

The reason that the value placed upon the innovation by the monopolist is smaller than the value of the innovation to a competitive firm and to society is again explained. A competitive firm is just breaking even prior to adopting the innovation and so values the innovation at the full additional profits it will generate. By contrast, the monopolist is already earning a monopoly profit with its existing technology. Introducing the new process displaces and therefore undermines that investment, and as with the competitive firm, the monopolist ignores the increase in consumer surplus. This is often referred to as the *replacement effect* but the term is misleading. After all, society also values the innovation by comparing it to the technology that it is replacing. The important reason the monopolist undervalues the innovation is because the monopolist restricts output to less than the socially optimal level. To see why suppose, by contrast, that the monopolist could employ first-degree price discrimination. Then the monopolist's valuation of the innovation would exactly equal society's valuation.

While the comparison just drawn is between a monopolist and a firm in a perfectly competitive market, the results would be the same if we instead compare a monopolist with a firm in an oligopoly market characterized by Bertrand competition. (Why?) Moreover, the same qualitative result will be obtained in a comparison of a monopoly firm with firms engaged in Cournot competition. The basic reason remains. While the Cournot firm does enjoy some positive, pre-innovation profits, these are much smaller than those of a monopolist. Therefore, the Cournot competitor has much less to lose than does the monopolist from pursuing the innovation. While the case just described considered a nondrastic process innovation, the same ordering, $V^p > V^c > V^m$, holds for a drastic one. In other words, the social gain from a drastic innovation exceeds the gain to a firm engaged in Bertrand (or Cournot) competition, which in turn exceeds the gain to a monopolist. Finally, while our analysis assumes a specific linear demand, the same results are obtained for any demand function even if it is nonlinear.[8]

22.2 Practice Problem

Assume that demand for a homogeneous good is $P = 100 - Q$, where P is measured in dollars, and that a process innovation reduces marginal costs of production from \$75 to \$60 per unit. Assume that the discount factor is $R = 0.9$.

a. Confirm that this is a nondrastic innovation and that marginal costs would have to be reduced to less than \$50 per unit for the innovation to be drastic.
b. Calculate the maximum amount that a monopolist is willing to pay for the innovation.

[7] This is derived from the property that one way to represent the monopolist's profit is the area between the monopolist's *MR* and *MC* curves.

[8] Gilbert (2006) shows that our results generalize to any demand function.

Now assume that the market is served by Cournot duopolists who have identical marginal costs of \$75 prior to the innovation.

c. Confirm that the pre-innovation price is \$83.33 and that at this price each firm has profits per period of \$69.44.
d. Confirm that if one of these firms is granted use of the innovation, the price will fall to \$78.33.
e. Show that this firm is willing to pay more for the innovation than the monopolist.

22.2.2 Preserving Monopoly Profit and the Efficiency Effect

The analysis in the previous section assumed that there was only one innovator, namely, a lab outside the industry. If that laboratory company does not innovate, no one does. This view does not truly capture the spirit of Schumpeter's contention. Instead, Schumpeter's point is precisely that firms compete by means of innovation. This means that firms have their own labs and that each firm is a potential innovator. As a result, even if one firm does not innovate, another might. This can reverse the previous results.[9]

Suppose that demand is given by $P = 120 - Q$ and that the current technology allows production at a marginal cost of \$60. An incumbent monopolist and a potential entrant play the following three-stage game. In stage 1 the incumbent decides whether or not to undertake R&D, which we assume reduces marginal cost to \$30. In stage 2 a potential entrant decides whether or not to enter. If the incumbent has not undertaken R&D, the entrant then chooses whether or not to undertake R&D. Without R&D the entrant's marginal cost is \$60 and with it marginal cost is \$30. No matter who innovates, the innovation is protected by a patent of unlimited duration that cannot be "invented around" by other potential or actual firms. If entry occurs then in stage 3 the entrant and the incumbent act as Cournot competitors. The extensive form of this game is illustrated in Figure 22.3.

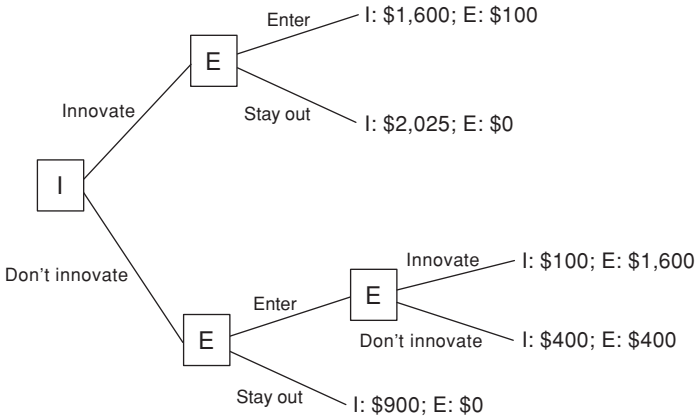

Figure 22.3 Extensive form for the Innovation and Entry game

[9] The underlying analysis can be found in Gilbert and Newbery (1982). Reinganum (1983) shows, however, that this conclusion might not hold when the timing of the successful breakthrough is uncertain. The incumbent monopolist might delay innovation in order to enjoy its current profits. A potential entrant has no such incentive to delay its innovative activity.

As usual we solve this game "backwards." Suppose that the incumbent has undertaken R&D. Then the entrant will enter with a cost of \$60 and the incumbent earns per period profit of \$1,600 and the entrant earns per-period profit of \$100. (This assumes, of course, that there are no sunk costs of entry. We return to this point below.) Now suppose that the incumbent does not innovate. The entrant will certainly enter. Innovation by the entrant gives the entrant per-period profit of \$1,600 while no innovation leads to per-period profit of \$400.

We can now calculate how much the innovation is worth to the incumbent and to the entrant. For the entrant, innovation increases per-period profit from \$400 to \$1,600. Accordingly the present value of the innovation to the entrant is $V^e = \$1{,}200/(1 - R)$. What about the incumbent? No innovation by the incumbent will lead to innovative entry provided only when the cost of the innovation is less than $\$1{,}200/(1 - R)$. In that case, the incumbent then earns per-period profit of \$100. By contrast, if the incumbent innovates and pre-empts innovation by the entrant the incumbent earns per-period profit of \$1,600. As a result, the value of the innovation to the incumbent is $V^i = \$1{,}500/(1 - R)$. Clearly this exceeds the value placed on the innovation by the entrant. Hence, the monopolist has the stronger incentive to innovate.

Our analysis illustrates the potential for innovation to deter entry, protecting the incumbent's monopoly position and profit. Suppose that sunk entry costs S are such that $\$100/(1 - R) < S < \$400/(1 - R)$. That is, imagine that sunk costs are greater than the profit that the entrant expects to make if the incumbent innovates but less than the profit that the entrant expects to make if neither entrant nor incumbent innovates. The value of the innovation to the entrant is unchanged at $\$1{,}200/(1 - R)$. This is not the case for the incumbent. Now innovation deters entry, allowing the incumbent to maintain her monopoly position with per-period profit of \$2,025. Failure to innovate, by contrast, leads to innovative entry and per-period profit to the incumbent of \$100. The value of the innovation to the incumbent is now even greater at $V^m = \$1{,}925/(1 - R)$.

The foregoing result is not peculiar to the numbers we have assumed. It is in fact quite general. Suppose first that innovation by the incumbent does not deter entry. Denote the per-period duopoly profit of the incumbent as $\pi_i^d(c_i, c_e)$ and of the entrant as $\pi_e^d(c_i, c_e)$, where c_i is marginal cost of the incumbent and c_e is marginal cost of the entrant. Innovation reduces marginal cost from c_h (high) to c_l (low). The incumbent knows that innovation gives per-period profit $\pi_i^d(c_l, c_h)$ while failure to innovate leads to innovative entry and profit $\pi_i^d(c_h, c_l)$. For the entrant, innovation is possible only if the incumbent has not innovated. Innovation then gives per-period profit of $\pi_e^d(c_h, c_l)$ while the failure to innovate gives per-period profit of $\pi_e^d(c_h, c_h)$. The per-period value of the innovation to the incumbent is $\pi_i^d(c_l, c_h) - \pi_i^d(c_h, c_l)$ while the per-period value of the innovation to the entrant is $\pi_e^d(c_h, c_l) - \pi_e^d(c_h, c_h)$.

Symmetry between the two firms tells us that $\pi_i^d(c_l, c_h) = \pi_e^d(c_h, c_l)$ and $\pi_e^d(c_h, c_h) = \pi_i^d(c_h, c_h)$. As a result, for the incumbent to place a higher value on the innovation than the entrant requires that $\pi_i^d(c_h, c_l) < \pi_i^d(c_h, c_h)$. This condition is always satisfied. The profit of the incumbent firm when it faces a low-cost rival is less than its profit when it faces a high-cost rival, no matter what the incumbent's marginal costs are.

Now suppose that innovation by the incumbent deters entry. The per-period value of the innovation to the entrant is unchanged. By contrast, the per-period value of the innovation to the monopolist is now $\pi^m(c_l) - \pi_i^d(c_h, c_l)$. This is clearly greater than the value of the innovation with entry since $\pi^m(c_l) > \pi_i^d(c_l, c_h)$. A low-cost incumbent always prefers monopoly to sharing the market, even when the sharing is with a high-cost rival.

To summarize, no matter whether innovation by an incumbent monopolist maintains that monopoly or not, the incumbent firm values the innovation more highly than a potential entrant. Replacing oneself is better than being replaced by a newcomer. This effect is called the *efficiency effect*.

22.3 A MORE COMPLETE MODEL OF COMPETITION VIA INNOVATION

What drives the efficiency effect is the fact that the cost of not innovating becomes higher once we recognize that it is precisely in this case that a rival may innovate. Such an increase in the opportunity cost of non-innovation makes the incumbent monopolist much more willing to pay for the innovation. Clearly, the strategic interaction from potential entry through innovation seems closer to the view Schumpeter (1942) presents.

We can get even closer to the Schumpeterian spirit by making the decision to spend on R&D an explicit part of a firm's strategy. The simplest model in this spirit is one due to Dasgupta and Stiglitz (1980). Their model is attractive both for its key insights and because it builds on the Cournot model developed in section 9.5 in Chapter 9. We present the essentials of their analysis here.

Dasgupta and Stiglitz assume an industry comprised of n identical Cournot firms each of which has to determine the level of output, q_i it will produce and the amount, x_i, that it will spend on R&D. While R&D is costly, the benefit of R&D spending is that it lowers the firm's unit cost of production, c. Specifically, each firm's unit cost is a decreasing function of the amount it spends on R&D, $c_i = c(x_i)$ and $dc(x_i)/dx_i < 0$. Total net profit for any firm, π_i, is:

$$\pi_i = P(Q)q_i - c(x_i)q_i - x_i \tag{22.1}$$

Suppose that each firm spends a specific amount, x^*, on research. Each firm then has a unit cost of $c(x^*)$. Accordingly, if we know the value of x^*, we know each firm's unit cost, and we can work out the equilibrium output for the individual firm and the industry in total using the analysis from section 9.5.[10] In particular, we know that the outcome in this symmetric, n-firm Cournot model is an equilibrium price-cost margin, or Lerner Index, given by

$$\frac{(P - c(x^*))}{P} = \frac{s_i}{\eta} \tag{22.2}$$

Here, P is the industry price, s_i is the ith firm's share of industry output, η is the elasticity of market demand, and x^* is the amount that each firm spends on R&D in equilibrium. We have dispensed with the subscript on the term x^* because for identical firms the amount chosen is the same in equilibrium for each firm. We can simplify further by recognizing that since all firms are identical, s_i is just $1/n$. So, equation (22.2) can be written as

$$P\left(1 - \frac{1}{n\eta}\right) = c(x^*) \tag{22.3}$$

[10] If we set the derivative of equation (22.1) with respect to q_i to zero, taking the production of all firms other than the ith, Q_{-i} as given, and then solve for q_i we obtain each firm's best response function.

Equation (22.3) does not by itself tell us the amount of R&D expenditure, x^*, that each firm will find optimal in the ultimate equilibrium. To determine that value we must add a second equilibrium condition indicating when a firm will know that it has spent the right amount on research activities. This is obtained by differentiating the profit equation (22.1) with respect to the R&D expenditures, x_i, to give the condition

$$\frac{\partial \pi_i}{\partial x_i} = -\frac{dc(x_i)}{dx_i} q_i - 1 = 0 \tag{22.4}$$

which can be simplified to the condition that in equilibrium we must have

$$-\frac{dc(x_i)}{dx_i} q_i = 1 \tag{22.5}$$

What does this mean? Remember that an increase in R&D expenditures reduces marginal cost $c(x_i)$, so that $dc(x_i)/dx_i$, the amount by which marginal cost changes as a result of an additional dollar of R&D expenditures, is negative. The left-hand side of equation (22.5) is, therefore, positive and is equal to the full marginal benefit of an extra dollar of R&D spending. The marginal cost of an extra dollar spent on R&D is simply \$1. At the equilibrium level of R&D expenditures, x^*, the marginal benefit of an extra dollar spent on R&D just equals its marginal cost.

What are the implications of the equilibrium conditions of equations (22.3) and (22.5)? The most obvious conclusion is that an increase in the number of firms in the industry will decrease the amount that each firm is willing to spend on R&D. An increase in the number of firms in the industry decreases the amount that each firm will choose to produce. This is, actually, a direct implication of equation (22.3). But equation (22.5) makes clear that the marginal benefit of extra R&D spending is directly proportional to the volume of a firm's output. Hence, the reduction in a firm's output that results from increasing the number of firms also reduces the marginal benefit that R&D spending yields to an individual firm. It follows that the equilibrium level of such spending per firm, x^*, will fall as the number of firms rises.

This does not necessarily imply, however, that the total industry spending on R&D, which is nx^*, will also fall. It is perfectly possible that each firm spends less on R&D but total R&D spending increases. Dasgupta and Stiglitz show that aggregate spending on R&D may actually either increase or decrease as the number of firms in the industry increases. The key point is that for aggregate R&D spending to increase, the elasticity of market demand must be fairly large. When demand is relatively elastic, the expansion of industry output resulting from a greater number of firms will not decrease the price too much and, as a result, will not decrease the marginal revenue of equation (22.3) very much either. Since this difference between price and cost is what finances a firm's R&D expenditure, such expenditure can be expected to rise in total with the number of industry firms so long as η is relatively large. If, however, the elasticity of market demand declines as output expands (as is the case with linear demand curves), then increasing the number of firms will, beyond some point, lead to a reduction in total R&D efforts. Even for a relatively small number of firms in the market adding one more firm induces a decline in total R&D spending. Therefore, the Dasgupta and Stiglitz model may be taken as partial support for the Schumpeterian hypothesis that concentration fosters innovation.

The foregoing does not explain what determines the number of firms in an industry. Dasgupta and Stiglitz invoke a third equilibrium condition that, in the long run, free entry will lead to an increase in the number of firms until each firm makes zero profit. In other words, industry structure is determined endogenously by the firms' output and R&D expenditure decisions. The zero profit condition, when applied to equation (22.1) tells us that

$$P(Q^*)q^* - c(x^*)q^* - x^* = 0 \tag{22.6}$$

Aggregating this over the equilibrium number of firms in the industry, n^*, gives

$$P(Q^*)Q^* - c(x^*)Q^* - n^*x^* = 0 \tag{22.7}$$

which implies that $(P(Q^*) - c(x^*))Q^* = n^*x^*$. Now since each of the n firms is of the same size, each has a market share equal to $1/n$. By equation (22.2) we know that $P - c(x^*) = P/n^*\eta$. Using this substitution, the equilibrium R&D outcome derived by Dasgupta and Stiglitz is

$$\frac{n^*x^*}{P(Q^*)Q^*} = \textit{industry R\&D spending as a share of industry sales} = \frac{1}{n^*\eta} \tag{22.8}$$

Comparing across industries, equation (22.8) suggests that the share of an industry's total sales revenue that will be devoted to R&D is likely to be smaller in less concentrated industries. In other words, those industries with a naturally more competitive structure will undertake less R&D effort, all else equal. This may then be seen as offering fairly strong support for Schumpeter's basic claim that imperfect competition is good for technical progress and more imperfect competition is even better.

22.4 EVIDENCE ON THE SCHUMPETERIAN HYPOTHESIS

The debate over the Schumpeterian hypothesis cannot be resolved by an appeal to economic theory alone. We must also consider empirical evidence. To date, a number of statistical studies relating R&D effort to firm size and industry structure have been conducted. While these studies are far from uniform in their results, one general finding does emerge. R&D intensity does appear to increase with increases in industrial concentration but only up to a rather modest value after which R&D efforts appear to level off or even decline as a fraction of firm revenue.

Some of the earliest studies exploring the link between industry structure and R&D were those of Scherer (1965, 1967). His basic finding was that while firm size and concentration are each positively associated with the intensity of R&D spending, these correlations diminish beyond a relatively low threshold. That is, once firms reach a relatively small size and/or markets reach a relatively low level of concentration, any positive effects of firm size or market concentration on innovative activity tend to vanish. Subsequent studies, including those of Levin and Reiss (1984), Levin, Cohen, and Mowery (1985a and 1985b), Levin, Klevorick, Nelson, and Winter (1987), Lunn (1986), Scott (1990), Geroski (1990), Blundell, Griffith, and Van Reem (1995) have tended to confirm Scherer's (1965) basic finding.[11]

[11] See Cohen and Levin (1989) for an early summary.

Reality Checkpoint

Some Little Inventors That Could

While the jury is still out on the Schumpeterian hypothesis that larger firms and/or concentrated markets spur technological progress, there is certainly a good bit of anecdotal evidence regarding the prowess of individual inventors and small firms to come up with the big breakthrough. The personal computer, for instance, was mainly introduced by a then-small firm called Apple. The phonograph and wireless telegraphy were developed by individuals, Edison in the first case and Marconi in the second. George Westinghouse was a young man of 22, working alone, when he patented his model for a compressed air breaking system that was soon adopted by every train on both the Southern Pacific and Central Pacific railroads—and virtually all other U.S. trains within a few years.

The story does not end in the late nineteenth or early twentieth century. Xerox was a small firm called Haloid when it developed the Xerographic copying method. Intel, which now controls more than two-thirds of the microprocessor market, started out as a small firm packaging transistors on a sliver of silicon. More recently, small firms and entrepreneurs continue to be an active source of innovation. For example, two of the most heavily trafficked sites on the web, e-Bay and Amazon, were both the creation of small independent entrepreneurs. Genentech was just a tiny venture capitalist experiment when it launched the field of recombinant DNA. Larry Page and Sergey Brin were Stanford graduate students when they began collaboration on a search engine called BackRub that relied on a new kind of server environment using many low-end PCs instead of big, expensive machines. Within a few years BackRub became Google and Page and Brin were billionaires.

While it may be surprising that so many well-known inventions came from relatively small firms and entrepreneurs, what is perhaps even more surprising is how often larger firms turned their backs on those very innovations that later proved so successful. IBM for instance, totally ignored the PC market at its inception. Apple came out with the first personal digital assistant (PDA) but it was a little firm called Palm that solved the problem of making a connection between the PDA and the PC desktop. Edison himself, well after his firm had been established as the premier technical enterprise of its time, initially regarded his own invention of the motion picture camera as little more than a pleasant toy and his fights against the superior alternating current technology to provide the electricity for his light bulbs are famous.

The anecdotal record tempts one to suspect that the confidence that comes from being a large firm with market power easily crosses over to an arrogance that blinds the firm to major innovations. However, there is some rational explanation. First, as noted in the text, large firms based on existing goods and processes have much to fear from the replacement effect that innovations bring. Second, there may also be a natural division of effort in which small entrepreneurs introduce new technologies while large, established firms play the role of "fast second" that quickly capitalizes on the small-firm breakthroughs. The reason for this is that a large firm with many products may fear that any failure with a new good will in fact taint its entire product line. A new small firm focused on just a few goods does not have this risk.

Source: Markides, C. and P. Geroski, *Fast Second: How Smart Companies Bypass Radical Innovations to Enter and Dominate New Markets*, Jossey-Bass, San Francisco, 2005.

In examining the influence of firm size and market structure on innovative activity, a number of important issues must be addressed. The first of these is that in comparing R&D efforts across markets, we should control for the "science-based" character of each industry: recall the very different patent activity by industry category noted in Table 22.2. Markets in which the member firms produce goods such as chemical products or computer hardware have such a strong technical base and general advances in scientific understanding can rapidly translate into either product or process innovations. Other markets, however, such as those for haircuts or hairstyling, will have more difficulty in making use of scientific breakthroughs and have less direct contact with universities and research laboratories. It turns out that measures of such technological opportunities tend to be highly correlated with the degree of industry concentration. In other words, while the simple correlation between concentration and innovation may be positive, this correlation reflects the positive effects on innovation that come with increases in an industry's opportunity for technical advances. The more recent studies cited above demonstrate that controlling for this factor is very important.

A second factor is the distinction between R&D expenditures and true innovations as perhaps measured by patents. While innovative effort can be measured by the ratio of R&D spending to sales this approach really measures the inputs to the innovative process. Presumably though, what we are really interested in are the outputs of that process—the true number of innovations as perhaps measured by the number of patents a firm acquires. Even though different firms do the same amount of R&D spending, the Schumpeterian hypothesis might be validated if size or concentration leads that spending to be more productive. The studies cited above do in fact look at the patent output of firms. Here again, however, little evidence is found in support of the Schumpeterian claims. Cohen and Klepper (1996) conclude that the general finding is that large firms do proportionately more R&D than smaller firms but get fewer innovations from these efforts. A notable exception in this regard, however, is Gayle (2002) who finds that firms in concentrated industries do generate many more patents when patents are not simply counted but, instead, are measured on a citation-weighted basis.[12]

Finally, a third issue is the endogeneity of market structure. Some firms, for example Alcoa or Microsoft, came to dominate their industry on the basis of a dramatic innovation. In the case of Alcoa, it was its unique process for refining aluminum. In the case of Microsoft, it was its unique Windows operating systems for personal computers. In these and other cases, the key technology that led to the firm's dominant position was associated with a number of patents. If this experience is pervasive, a naïve researcher may find that large, dominant firms are also firms with many patents and wrongly conclude that the Schumpeterian hypothesis is validated. In these cases it is the innovative activity that leads to market power and not the other way around. If the firms that come to dominate their markets start out as small operations and then grow on the basis of entrepreneurial skill and technical breakthroughs, the implication would be quite to the contrary of Schumpeter's model.[13]

[12] When a patent application is filed, the applicant must cite all the prior patents related to the new process or product. It is plausible that the most important patents are those that are cited most frequently. Hence, in evaluating a firm's true innovative output, one may want to control for how often the firm's patents are cited.

[13] Generally, market structure and innovative activity evolve together. For example, if experience raises R&D productivity then older firms will tend to do more innovation because it has a higher return for them, so that early entrants will tend to dominate an industry over time. See Klepper (2002) for an analysis along these lines.

22.5 R&D COOPERATION BETWEEN FIRMS

Our final topic in this chapter addresses the issue of cooperation on R&D efforts among firms. Two features of the innovative process make such efforts attractive from the viewpoint of economic efficiency. First, modern technology is very complicated and often draws on different areas of expertise and experience. Because it is doubtful that the scientists and engineers in one firm possess all this knowhow, it is desirable that firms share their experiences, experimental results, and design solutions with each other so as to realize fully the benefits from scientific study. Second, there is a potential for wasteful R&D spending as firms duplicate each other's efforts in a noncooperative R&D race.

We do have explicit evidence on this score. One of the most dynamic and creative groups of firms in the U.S. economy in recent years has been the American steel minimills. These firms rely on small-scale plants using electric arc furnaces to recycle scrap steel. They are widely regarded as world leaders and have outperformed even the Japanese steel firms once thought to be invincible. Through a series of interviews, von Hippel (1988) found that these firms regularly and routinely exchanged technical information with each other. In fact, sometimes workers of competing firms were trained (at no charge) by a rival company in the use of specific equipment. Such exchanges of information and expertise were made with the knowledge and approval of management even though they had the effect of strengthening a competitor.

To analyze the implications of research spillovers, we again make use of the Cournot duopoly model, similar to the Dasgupta and Stiglitz (1980) model except that we now explicitly permit one firm's research to benefit others.[14] We address three issues. First, how do technical spillovers affect the incentives firms have to undertake R&D? Second, what is the impact of such spillovers on the effects of R&D? Finally, what are the benefits to be gained from allowing firms to cooperate in their research, for example, by forming research joint ventures (RJVs)? Are these benefits worth the risk that cooperation in R&D might facilitate collusion between the same firms in the final product markets?

To begin suppose that the demand for a homogeneous good is linear and given by $P = A - BQ$. Two firms, each of which has constant marginal costs of c per unit, manufacture the good. These costs can be reduced by research and development activity, but it is possible that the knowledge developed by one firm can spill over to its rival. This can happen, for example, because the firms fund common sources of basic research such as universities or research laboratories; or because the research direction that one firm is taking becomes known to its rival; or because some of the preliminary results of research effort leak out; or, of course, because of industrial espionage.

Specifically, if firm 1 undertakes R&D at intensity x_1 and firm 2 undertakes R&D at intensity x_2, the marginal production costs of the two firms become

$$\begin{aligned} c_1 &= c - x_1 - \beta x_2 \\ c_2 &= c - x_2 - \beta x_1 \end{aligned} \tag{22.9}$$

[14] The model is developed in d'Aspremont and Jacquemin (1988). A more general version of this type of investigation can be found in Kamien et al. (1992).

Here $0 \leq \beta \leq 1$ measures the degree to which the R&D activities of one firm spill over to the other firm.[15] If $\beta = 0$, there are no spillovers—firm 1's research effort x_1 yields benefits only to firm 1 itself. If $\beta = 1$, spillovers are perfect—every penny of cost reduction that x_1 brings to firm 1, it also brings to firm 2. For the intermediate case of $0 < \beta < 1$, spillovers are only partial—if firm 1's research lowers its own cost by one dollar per unit, it will lower firm 2's cost by some fraction of a dollar per unit.

Research is, of course, costly. Indeed, not only is it costly but we assume that R&D activity exhibits *dis*economies of scale, i.e., it becomes more costly the more research the firm does. Specifically, we assume that research costs are the same for both firms and given by the research cost function

$$r(x_i) = \frac{x_i^2}{2}, \quad i = 1, 2. \tag{22.10}$$

Thus, if the R&D intensity is $x_i = 10$, then the research budget $r(x_i) = 10^2/2 = \$50$. If the R&D intensity doubles to $x_i = 20$, the budgetary expense climbs to $20^2/2 = \$400$. A doubling of R&D effort therefore leads to a quadrupling of the R&D cost. This is an example of what we mean by a scale diseconomy.

22.5.1 Noncooperative R&D: Profit, Prices, and Social Welfare

Consider first what happens when firms do not cooperate on research. Suppose that we have a two-stage game and in the first stage, each firm chooses its research intensity x_i. In the second stage, each firm acts as a Cournot competitor in choosing its output. As usual, this game is solved backwards. From section 9.4, we know the Cournot equilibrium outputs for given values of c_1 and c_2 are

$$\begin{aligned} q_1^C &= \frac{(A - 2c_1 + c_2)}{3B} \\ q_2^C &= \frac{(A - 2c_2 + c_1)}{3B} \end{aligned} \tag{22.11}$$

and the firm profits after paying the research costs are

$$\begin{aligned} \pi_1^C &= \frac{(A - 2c_1 + c_2)^2}{9B} - \frac{x_1^2}{2} \\ \pi_2^C &= \frac{(A - 2c_2 + c_1)^2}{9B} - \frac{x_2^2}{2} \end{aligned} \tag{22.12}$$

[15] We confine our attention to the case in which the spillovers are positive. It is, however, possible that there might be negative spillovers. For example, firms might spread misinformation about their research or claim that they have made a breakthrough in order to discourage rivals from continuing with a particular line of research.

We know from equation (22.9) that $c_1 = c - x_1 - \beta x_2$ and $c_2 = c - x_2 - \beta x_1$. This allows us to express the final equilibrium outputs directly as a function of each firm's choice of R&D effort and the degree of spillover from one firm's findings to the other firm's costs. The resultant Cournot–Nash equilibrium outputs for each firm are

$$q_1^C = \frac{(A - c + x_1(2 - \beta) + x_2(2\beta - 1))}{3B}$$
$$q_2^C = \frac{(A - c + x_2(2 - \beta) + x_1(2\beta - 1))}{3B} \tag{22.13}$$

and their profits are

$$\pi_1^C = \frac{(A - c + x_1(2 - \beta) + x_2(2\beta - 1))^2}{9B} - \frac{x_1^2}{2}$$
$$\pi_2^C = \frac{(A - c + x_2(2 - \beta) + x_1(2\beta - 1))^2}{9B} - \frac{x_2^2}{2} \tag{22.14}$$

Equation (22.13) indicates that the output of each firm is an increasing function of its own R&D expenditures x_i. Such expenditures reduce a firm's costs and thereby make higher output more profitable. By contrast, the effect of the *rival's* R&D effort on a firm's production can go either way. Consider firm 1. On the one hand, the R&D activity of firm 2 spills over and lowers firm 1's costs, which has an expansionary effect on firm 1's own output. On the other hand, firm 2's R&D reduces firm 2's cost. This makes the firm 2 more competitive and permits it to expand its output leaving less market available to firm 1. The net result of these two countervailing forces is ambiguous. This ambiguity is reflected in the coefficient on x_2 in the q_1 equation and the coefficient of x_1 in the q_2 equation. In both cases, this coefficient, $2\beta - 1$, is positive only when the degree of spillover is large, that is, when $\beta > 0.5$. When spillovers are small, that is, when $\beta < 0.5$, a firm's output and profit are decreasing functions of the R&D expenditures of its rival. The same ambiguity appears in the profit equations (22.14).

We know that each firm will choose the level of research activity that maximizes its profit given the research effort of its rival. For every choice of effort that firm 2 makes, firm 1 will choose its own profit-maximizing response. The same is true for firm 2 in reverse. So we can in principle identify the best response or *research intensity reaction function* for each firm.

This is done mathematically in the inset to confirm an intuitive result that follows from our previous discussion. When research spillovers are low, the research intensity reaction functions for the two firms are downward sloping, indicating that the research expenditures of the two firms are *strategic substitutes*—more research by one firm reduces the amount done by the other. That is, research activity by one firm substitutes for research activity by the other. The intuition is that in this case the increased research effort by one firm, primarily reduces its costs and so gives it a competitive advantage with respect to the other rival firm. In turn, this results in a reduction in the profitability of the rival firm, which can be offset only by the rival reducing its expenditure on research.

By contrast, when spillovers are high, the research intensity reaction functions are upward sloping, meaning that the research expenditures of the two firms are *strategic complements*. When spillovers are this high, an increase in research intensity by one of the firms induces an increase in research intensity by the other. In this case the intuition is that if one firm

Derivation Checkpoint

Optimal Noncooperative R&D Effort in the Presence of Spillovers

Differentiation of the profit equation (22.14) with respect to research effort of firm i and setting the derivative equal to zero yields:

$$\frac{\partial \pi_i^C}{\partial x_i} = \frac{2(2-\beta)[A - c + x_i(2-\beta) + x_j(2\beta - 1)]}{9B} - x_i = 0$$

This result implies best response curves R_1 and R_2 for firm 1 and firm 2 of

$$R_1\colon\ x_1 = \frac{2(2-\beta)[A - c + x_2(2\beta - 1)]}{[9B - 2(2-\beta)^2]} \text{ and } R_2\colon x_2 = \frac{2(2-\beta)[A - c + x_1(2\beta - 1)]}{[9B - 2(2-\beta)^2]}$$

Clearly these are upward sloping if $\beta > 0.5$ and downward sloping if $\beta < 0.5$. The equilibrium must be symmetric since the two firms have identical costs in the absence of R&D and face the same demand function. Thus in equilibrium $x_1 = x_2$. Substituting this into R_1, for example, and solving for x_1 gives the Nash equilibrium research intensity:

$$x_1^C = x_2^C = \frac{2(A-c)(2-\beta)}{9B - 2(2-\beta)(1+\beta)}.$$

This is decreasing in β, implying that increased research spillovers decrease each firm's chosen research intensity. The solution for research effort x_i implies output levels and profits for the two firms of

$$q_1^C = q_2^C = \frac{3(A-c)}{9B - 2(2-\beta)(1+\beta)}$$

$$\pi_1^C = \pi_2^C = \frac{(A-c)^2[9B - 2(2-\beta)^2]}{[9B - 2(2-\beta)(1+\beta)]^2}.$$

opts for a high level of R&D effort, the benefits of that activity spill over to the other firm to such an extent that the other firm's profit increases, providing that firm with the funds and the incentive to increase its own R&D spending. Figure 22.4 illustrates typical research intensity reaction functions.

However, determining whether the reaction functions slope downward or upward—whether the two firms' R&D efforts are strategic substitutes or complements—does not tell us what the equilibrium level of R&D spending is. In particular, there can be no presumption that the presence of large R&D spillovers and hence the case of strategic complements will result in a higher equilibrium level of R&D spending than the case in which such spillovers are low. The Nash equilibrium occurs at the intersection of the two response functions, and the case in which this point is farthest from the origin is far from obvious.

In order to illustrate this last point, we shall focus for the remainder of our discussion on a numeric example. The insets in this and the next section give a more general mathematical

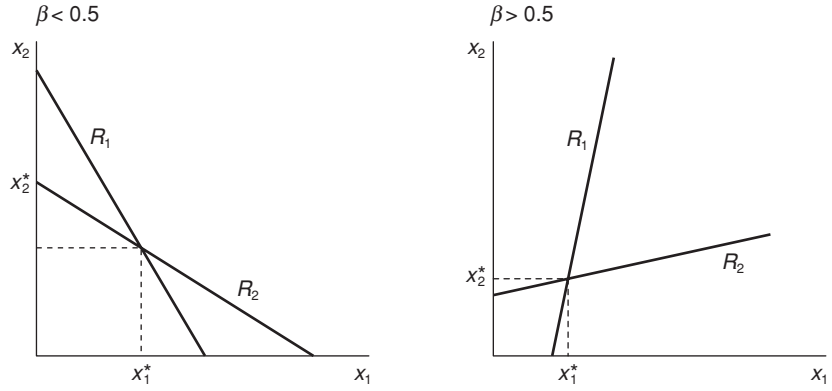

Figure 22.4 Best-response functions for research intensity in the noncooperative R&D game

analysis. Let demand for the good be $P = 100 - 2Q$, and each firm's marginal production cost is currently \$60. The firms can choose two levels of research intensity: $x_i = 10$ or $x_i = 7.5$. Further, we assume that the degree of research spillover (which is outside the control of the two firms) takes one of two values: a low value of $\beta = 1/4$ or a high value of $\beta = 3/4$.

Consider first the low-spillover case and assume that firm 2 chooses the high research intensity of $x_2 = 10$. If firm 1 also chooses high research intensity, its output and profits will be, from equations (22.13) and (22.14)

$$q_1^C = \frac{(40 + 17.5 - 5)}{6} = 8.75;\ \pi_1^C = \frac{(40 + 17.5 - 5)^2}{18} - \frac{100}{2} = \$103.13$$

By contrast, if firm 1 chooses the low research intensity, its output and profits will be

$$q_1^C = \frac{(40 + 13.125 - 5)}{6} = 8.02;\ \pi_1^C = \frac{(40 + 13.125 - 5)^2}{18} - \frac{56.25}{2} = \$100.54$$

Now assume that firm 2 chooses the low research intensity of $x_2 = 7.5$. If firm 1 chooses the high research intensity, its output and profits will be

$$q_1^C = \frac{(40 + 17.5 - 3.75)}{6} = 8.96;\ \pi_1^C = \frac{(40 + 17.5 - 3.75)^2}{18} - \frac{100}{2} = \$110.50$$

By contrast, if firm 1 chooses the low research intensity, its output and profit will be

$$q_1^C = \frac{(40 + 13.125 - 3.75)}{6} = 8.23;\ \pi_1^C = \frac{(40 + 13.125 - 3.75)^2}{18} - \frac{56.25}{2} = \$107.31$$

The same calculations apply to firm 2. We then have the payoff matrix of Table 22.3(a). *The Nash equilibrium in this case of low spillovers is for both firms to adopt high research intensities.*

When the degree of R&D spillover is high, with $\beta = 0.75$, the same calculations lead to the payoff matrix of Table 22.3(b). *The Nash equilibrium in this case is for both firms to*

Table 22.3a Payoff matrix with low R&D spillovers, $\beta = 0.25$

		Firm 1	
		Low research intensity	*High research intensity*
Firm 2	*Low research intensity*	\$107.31; \$107.31	\$100.54; \$110.50
	High research intensity	\$110.50; \$100.54	\$103.13; \$103.13

Table 22.3b Payoff matrix with low R&D spillovers, $\beta = 0.75$

		Firm 1	
		Low research intensity	*High research intensity*
Firm 2	*Low research intensity*	\$128.67; \$128.67	\$136.13; \$125.78
	High research intensity	\$125.78; \$136.13	\$133.68; \$133.68

adopt low research intensities. An increase in the degree of R&D spillover causes the two firms to reduce their research intensities. Why? Consider first the case when R&D spillovers are weak. In this case the more firm 1 spends on R&D, the less firm 2 will spend because the primary effect of such spending is to strengthen the competitive position of firm 1. Yet somewhat paradoxically, this gives each firm an incentive to spend aggressively on R&D so as to avoid being the loser in this competition. If firm 1 spends a lot on R&D and firm 2 spends nothing, virtually all the benefits of firm 1's spending stay with firm 1. Firm 2 would find itself losing significant market share and profit to a much lower-cost competitor. When each firm tries to avoid falling behind in this manner, the net result can be a substantial amount of R&D effort, both at each firm and in total.

Just the opposite holds in the case of large spillovers. Yes, the more firm 1 spends on R&D, the more firm 2 is induced to spend by virtue of the strategic-complements setting, but this relation is a two-edged sword. Even if firm 1 spends only a little on R&D it knows that this will still induce firm 2 to do a fair bit of research activity. Moreover, the research activity at firm 2 will bring substantial benefits to firm 1 by virtue of the large spillovers. In this setting, the incentive for either firm to spend much on R&D can be quite small as each firm seeks to free ride on the other's efforts.

Where graphs fail to give clear results, algebra can often save the day. That this is true here is shown in the Derivation Checkpoint. We merely state the final result. The amount of research done by each firm *decreases* as β, the degree of R&D spillover, increases—the free-riding effect to which we have just referred.

22.5.2 Technology Cooperation

We now consider two alternative arrangements between the duopoly firms that can alter the outcome from that described above. The first possibility is that the two firms agree that while each will continue to do its own R&D, they will coordinate the extent of such research effort.

Derivation Checkpoint

Optimal R&D Effort with a Research Joint Venture (RJV)

The third and final case that we consider is where the firms form an RJV to coordinate their research activities and ensure that these are fully shared. This case is easily dealt with. It amounts to the firms ensuring that the degree of R&D spillover is perfect, that is that $\beta = 1$.

The optimal degree of research intensity, outputs and profits are, therefore, identified by substituting $\beta = 1$ in the equations derived for R&D cooperation. Thus:

$$x_1^{RJV} = x_2^{RJV} = \frac{4(A-c)}{9B-8} \quad q_1^{RJV} = q_2^{RJV} = \frac{3(A-c)}{9B-8} \quad \pi_1^{RJV} = \pi_2^{RJV} = \frac{9(A-c)^2}{(9B-8)^2}.$$

An RJV in which firms coordinate their research efforts to maximize their joint profits and share the results of their R&D activities dominates the other cases that we have considered in that it gives the highest per-firm profits and the lowest consumer prices.

Thus, the two firms now choose x_1 and x_2 to maximize their joint profit. They continue to recognize that they will compete as Cournot firms in the product market. The other alternative we consider is that the firms explicitly share their R&D activities by setting up a research joint venture (RJV). One way this scenario might work in practice would be for the two firms to jointly set up a laboratory for experimentation and analysis with all the discoveries made at that laboratory to be made fully available to both firms.

We introduce this RJV arrangement into the model by letting the two firms pick x_1 and x_2 cooperatively but by adding the further assumption that the degree of spillover is complete, that is, $\beta = 1$. Whatever is learned in the research lab—whether discovered by a firm 1 scientist or a firm 2 scientist—reduces the cost of both firms by the same amount.

We start with the simple coordination case. What we want to do is to pick the values of x_1 and x_2 that maximize the sum of the individual profit expressions shown in equation (22.14). The mathematical solution, as in the previous section, is shown in the Derivation Checkpoint.

We shall concentrate once again on our simplified example as given by the payoff matrices of Tables 22.3(a) and (b). When the firms coordinate their research efforts they choose the combination of R&D intensities that maximizes the sum of the profits in the cells of the relevant payoff matrix. *When R&D spillovers are low, coordination leads each of the firms to choose the low R&D intensity*. Cooperation then increases each firm's profits from \$103.13 to \$107.31. By contrast, *when the R&D spillover is high, coordination leads each firm to choose the high R&D intensity*, increasing their profits from \$128.67 to \$133.68.

What our example and the more detailed analysis of the inset show is the following: First, it is now the case that the higher is the level of R&D spillover—the larger is β—the more each firm spends on research. This is because the agreement between the two firms to set their R&D efforts jointly explicitly forces each firm to internalize the external benefits that such spending has upon its rival. In turn, this eliminates the free-riding problem that characterizes R&D competition when there are high spillovers. The ability to avoid this

problem also means that the two firms will each enjoy a profit at least as great as that which they would have earned in the absence of such cooperation.[16]

The second point to note is that the outcome under the simple coordination plan may not necessarily be good for the consumer. In particular, consumers are hurt by the technology cooperation when $\beta < 0.5$ and the extent of spillover is small. The reason is straightforward. When β is small, then without cooperation each firm tends to do a fair bit of research. It does so because a low value of β means that most of the benefit from its innovative efforts will accrue to it alone and because it knows that its rival is proceeding along the same line of attack. From the viewpoint of consumers, this is great since there has been considerable cost reduction and therefore a sharp decline in the price they pay. If, in this setting, we now introduce a cooperative R&D agreement, the two firms realize that their best bet is to reduce R&D intensity, which otherwise simply makes competition tougher in the product market. By decreasing R&D intensity, the firms increase their profits. Unfortunately, the lower rate of innovation also implies a higher price to consumers.

By contrast, when the degree of R&D spillover is high ($\beta > 0.5$), both firms and consumers benefit from a cooperative R&D agreement. This happens because now the primary effect of the R&D cooperation is to correct a market failure. In the absence of cooperation, a large degree of spillover leads each firm to free ride on the R&D efforts of its rivals and to take no account of the beneficial effects its own R&D expenditures have on the costs and profits of other firms. R&D cooperation internalizes these effects because it forces the cooperating firms to look at the impact their R&D expenditures have on aggregate profit rather than merely on their individual profit.

What about a research joint venture? As noted, an RJV can be best thought of as a case in which the firms take actions not only to coordinate their research expenditures but also to ensure that the spillover from one firm's research to the other's is complete, that is, so that $\beta = 1$. A little thought should convince you that an RJV will likely yield the maximum benefits to both firms and consumers. As we just saw, coordination of R&D levels benefits both producers and consumers whenever $\beta > 0.5$. Moreover, the profit outcomes in Tables 22.3 indicate that if the firms could find some way of increasing the technology spillover from $\beta = 0.25$ to $\beta = 0.75$ they would both benefit at *any* research intensity. We can make this discussion even more general. The inset on R&D cooperation shows that an increase in the spillover parameter β increases the research intensity and the profits of each firm *and* increases the output that each firm brings to the market. In other words, *both firms and consumers benefit* from an increase in β. The RJV takes this to its logical conclusion by ensuring that $\beta = 1$, its highest possible value.

In our example the benefits of an RJV are easily confirmed. Consider the case in which both firms choose the high degree of research intensity. Thus, with perfect R&D spillover, the profits to each firm are $(40 + 10 + 10)^2/18 - 50 = \$150$, while if each chooses the low research intensity, the profits to each firm will be $(40 + 7.5 + 7.5)^2/18 - 56.25/2 = \$139.93$. Clearly, the RJV will go for the high research intensity leading to the lowest costs. In turn, this will translate into the lowest consumer prices that these Cournot firms will offer.

The intuition behind the foregoing analysis is as follows. First, by maximizing the extent of spillovers, the RJV also maximizes the benefits of R&D. Every discovery is spread instantly to all firms in the industry. Second, despite this extensive spillover, the free-riding problem

[16] For $\beta = 0.5$ each firm earns exactly the same profit under uncoordinated R&D spending as each does with an R&D cartel. For all other values of β each firm's profit is higher with the R&D cartel.

Derivation Checkpoint

Optimal R&D Effort with R&D Cooperation

With cooperation, each firm's optimal research intensity is the R&D effort that maximizes the sum of the two firms' profits, given that output is determined noncooperatively in the second-stage output game. From equation (22.14) we know that aggregate profit is:

$$\pi_1^C + \pi_2^C = \frac{[A - c + x_1(2 - \beta) + x_2(2\beta - 1)]^2}{9B} - \frac{x_1^2}{2} + \frac{[A - c + x_2(2 - \beta) + x_1(2\beta - 1)]^2}{9B} - \frac{x_2^2}{2}$$

Differentiating this with respect to x_1 gives the first-order condition:

$$\frac{\partial(\pi_1^C + \pi_2^C)}{\partial x_1} = \frac{2(2 - \beta)[A - c + x_1(2 - \beta) + x_2(2\beta - 1)]}{9B} - x_1 + \frac{2(2 - \beta)[A - c + x_2(2 - \beta) + x_1(2\beta - 1)]}{9B} = 0$$

A similar condition applies for firm 2 but we do not need this. Rather, we can take advantage of the fact that the equilibrium for these two firms will be symmetric. Substituting $x_1^c = x_2^c = x^{RC}$ (where the superscript *RC* denotes "R&D cooperation") and simplifying gives:

$$\frac{2(1 + \beta)[A - c + x^{RC}(1 + \beta)] - 9Bx^{RC}}{9B} = 0$$

which gives the equilibrium R&D intensity as:

$$x_1^{RC} = x_2^{RC} = \frac{2(A - c)(1 + \beta)}{9B - 2(1 + \beta)^2}.$$

This is increasing in β.

The Nash equilibrium outputs and profits for the two firms when they cooperate in R&D can be identified by substituting into equations (22.13) and (22.14). After simplification this gives:

$$q_1^{RC} = q_2^{RC} = \frac{3(A - c)}{9B - 2(1 + \beta)^2}$$

$$\pi_1^{RC} = \pi_2^{RC} = \frac{9B(A - c)^2}{[9B - 2(1 + \beta)^2]^2}.$$

is now avoided. Because the two firms have agreed to coordinate their research efforts they fully internalize the otherwise external effects of research. Thus, firms will pursue extensive research, which, partly because of the extensive spillover effect of sharing, will lead to a sizable reduction in costs for every firm. This substantial cost reduction translates into an

equally impressive reduction in the price to consumers.[17] The policy implication of this is obvious and important. Research joint ventures should be encouraged because they benefit both consumers and producers so long as the antitrust authorities can ensure that such cooperation on research effort does not also extend to cooperation in production and prices, that is, to a price-fixing cartel.

The potentially large benefit from technology cooperation is undoubtedly the reason that research joint ventures—unlike price-fixing agreements—are not treated as *per se* violations by the antitrust authorities. Instead, they are evaluated on a rule of reason basis. Indeed, the U.S. Congress passed legislation in 1984 to require explicitly the application of a reasonability standard in the specific case of RJVs.

22.6 EMPIRICAL APPLICATION R&D Spillovers in Practice

R&D spillovers suggest a diffusion-like process. The greater is the spillover, the more rapid or the more complete is the diffusion of technological advances in one firm to the productivity of other firms. We might also suspect that a similar process is at work at a national and even international level. In particular, it seems likely that the R&D efforts of one country could spill over to enhance the productivity of its neighbors. Here again, the extent of such spillover is of interest. If technical advances in one country spread quickly to others, β will be high. In a world in which the international transfer of technical knowledge is weak, β will be low.

Wolfgang Keller (2002) explores the extent of international technical spillover by looking at data covering 12 broadly defined manufacturing industries from 14 countries over the years, 1970–95. To understand his basic approach, consider a simple Cobb–Douglas production function (see section 4.5) for industry i in country c.

$$Q_{ci} = K_{ci}^{1-\sigma} L_{ci}^{\sigma} \tag{22.15}$$

Here Q_{ci} is output (value added) and K_{ci} and L_{ci} are capital and labor inputs, respectively, in industry i in country c and σ is the share of costs accounted for by labor. Taking logarithms then yields:

$$\ln Q_{ci} = (1 - \sigma) \ln K_{ci} + \sigma \ln L_{ci} \tag{22.16}$$

For industry i, we define $\ln \bar{Q}_{ci}$ to be the average log of output across all countries. Similarly, for industry i, let $\ln \bar{K}_{ci}$ and $\ln \bar{L}_{ci}$, be the average amount of capital and labor inputs (again in logs), respectively, across all countries. If we define total factor productivity, TFP_{ci} in industry i and country c as the difference between the log of output and the weighted average level of inputs, i.e., $TFP_{ci} = \ln Q_{ci} - (1 - \sigma) \ln K_{ci} + \sigma \ln L_{ci}$, then the *relative* (to the mean) factor productivity F_{ci} of industry i in country c at a point in time is:

$$F_{ci} = (\ln Q_{ci} - \ln \bar{Q}_{ci}) - (1 - \sigma_{ci})(\ln K_{ci} - \ln \bar{K}_{ci}) - \sigma_{ci}(\ln L_{ci} - \ln \bar{L}_{ci}) \tag{22.17}$$

[17] While we have derived this result for a duopoly, Kamien et al. (1992) show that it extends to an n-firm oligopoly.

where we now let the cost share of labor σ_{ci} vary across countries and industries. Equation (22.17) is a measure of the extent to which output in industry i in country c remains above average even after correcting for any above average use of inputs. It is thus a measure of the productivity advantage (or disadvantage) in industry i in country c at any point in time. This is why it is called *relative* productivity. Of course, this will probably change over time due to R&D and other factors. For this reason, Keller (2002) measures relative productivity for each year from 1970 to 1995. This means that for each of the 12 industries in each of the 14 countries, Keller (2002) has a measure of relative productivity in each year, 1970 to 1995. Because we are now measuring this term over time as well as over industries and across countries, relative factor productivity now has an additional time subscript, i.e., F_{cit}. It is this series of relative productivity measures that Keller seeks to explain.

Because, by construction, variations in the relative productivity measure F_{cit} reflect variations in factors other than capital and labor inputs it is natural to identify these remaining differences as those due to differences in technology. In turn, these technical differences ought to reflect differences in R&D. Keller's (2002) approach in this respect is to distinguish a difference between R&D done domestically in industry i and that done abroad. The first research question is whether foreign R&D spills over to domestic productivity. The second is whether these spillovers are greater for countries that are closer to each other.

The 14 countries in Keller's (2002) sample are: Australia, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Sweden, the United Kingdom, and the United States. Five of these countries—France, Germany, Japan, the United Kingdom, and the United States—account for over 92 percent of all the R&D in the sample. Hence, Keller (2002) treats these G5 countries as the potential engines of technical change and examines how their R&D affects productivity in the remaining nine. Specifically, he estimates the parameters of the following equation:

$$F_{cit} = \alpha_{ci} + \alpha_t + \lambda \ln\left[S_{cit} + \gamma\left(\sum_{g \in G5} S_{git} e^{-\delta D_{cg}}\right)\right] + \varepsilon_{cit} \tag{22.18}$$

Here, F_{cit} is the relative productivity measure derived above for industry i in country c in year t, measured for each of the nine countries examined. The first term is a country and industry specific constant that permits for a time-independent productivity advantage or disadvantage for that sector in that country. The second is meant to pick up productivity increases over time that affect all firms in all countries in common. The key parameters are embedded in the next term. S_{cit} is a measure of the R&D done in industry i in country c up to time t. In contrast, S_{git} is a measure of R&D in that same industry but in one of the G5 countries and D_{cg} is the distance of that country from the domestic country in question. ($D = 1$ implies a distance of 235 kilometers.) Together, S_{cit} and the summation term for the G5 countries are meant to capture the R&D relevant to productivity in the domestic industry i at time t.

The effect of that combined industry-based R&D on productivity in that same industry in the domestic country is captured by the parameter λ. However, two adjustments are included to distinguish the impact of foreign from that of domestic R&D. To understand the first of these adjustments suppose that all G5 countries were right next to the domestic country in question ($D_{cg} = 0$). Then the contribution of their R&D on domestic productivity in industry i to total industry-relevant R&D, is adjusted by the parameter γ (taken to be same for all G5 countries). That is, if $\gamma < 1$, foreign R&D contributes less than the full effect of domestic R&D in adding to the knowledge relevant to a particular domestic industry's

productivity. In many ways, then γ comparable to the β of our industry analysis above. However, Keller (2002) also introduces a second source of distinction between domestic and foreign research lies in the distance term, D_{cg}. As this distance grows, the contribution of that G5 country's R&D to domestic productivity diminishes further if the parameter δ is positive. In short, the specification permits both for the possibility that simply by being foreign R&D may contribute less to domestic technology than home-grown research, and also for the more complicated fact that spillovers from foreign R&D grow smaller as the source of that R&D is farther away. Of course, the error term ε_{cit} picks up any remaining random factors that affect productivity.

The specification in equation (22.18) assumes that the decay parameter δ is the same throughout the time period. Keller (2002) recognizes, however, that increased globalization over the 25 years of his sample suggests that δ will decline over this period. He therefore estimates an alternative specification given by:

$$F_{cit} = \alpha_{ci} + \alpha_t + \lambda \ln\left[S_{cit} + \sum_{g \in G5} \gamma_G \left(1 + \psi_F I_t\right) S_{git} e^{-\delta(1+\psi_D I_t) D_{cg}}\right] + \varepsilon_{cit} \tag{22.19}$$

In this equation, I_t is a 1,0 dummy variable equal to zero over the first half of the sample to 1982 and then 1 in the 13 years thereafter. The coefficient ψ_F permits the effect of G5 R&D to have a different effect on domestic productivity in the second half of the sample than it does in the first, holding the distance between the domestic and G5 countries constant. Similarly, the coefficient ψ_D permits the extent to which spillovers decline with distance to change from the first half of the sample to the second half. Note too that this specification allows the effect of G5 R&D to differ across each G5 country by permitting a different coefficient γ_G for each one. This is reasonable as the different languages in these countries may affect the ease with which a technology can be transferred.

Because the contribution of foreign R&D to the total relevant measure of R&D depends on the parameters γ (or γ_G) and δ that are also to be estimated, equations (22.18) and (22.19) cannot be estimated by ordinary least squares. Instead, a non-linear least squares estimation is required. In this procedure, we begin with a starting value for the nonlinear parameters and then estimate the regression with OLS. We may then use these estimates to reiterate the process until the coefficient estimates stop changing and converge to stable values. Table 22.4 shows the key parameter estimates and their standard errors that Keller (2002) obtains from this maximum likelihood process for both specifications.

The estimates in specification 1 suggest that technical spillovers are strongly localized. The cumulative productivity effect of overall R&D is to raise productivity by 7.8 percent. However, foreign (G5) R&D contributes only 84 percent to the technical base that domestic research does and that is only if the domestic country is right next to the G5 source nation so that $D = 0$. The estimate of $\delta = 1.05$ though, indicates that that contribution dies out rapidly. Half of it is gone when $D = 0.69$ or at a distance of 162 kilometers (100 miles), and the rest is virtually eliminated once the source country of the foreign R&D is more than 400 miles away.

However, the results from specification 2 qualify the foregoing findings. It is useful to first note that the estimate of ψ_F is insignificantly different from zero. Hence, correcting for distance and country of origin, the contribution of a G5 country's R&D on domestic technical know-how is pretty much the same throughout the sample years and, on average, not too different from the 84 percent found in Specification 1 when the distance to the G5

Table 22.4 Regression estimates of international R&D spillovers

Parameter	*Specification 1*		*Specification 2*	
	Parameter estimate	*Standard error*	*Parameter estimate*	*Standard error*
λ	0.078	(0.013)	0.096	(0.008)
δ	1.005	(0.239)	0.384	(0.047)
γ	0.843	(0.059)	—	—
γ_J	—	—	1.000 (set)	set
γ_{US}	—	—	1.031	(0.059)
γ_{UK}	—	—	0.863	(0.060)
γ_{GER}	—	—	1.157	(0.060)
γ_F	—	—	1.011	(0.060)
ψ_D	—	—	−0.784	(0.068)
ψ_F	—	—	−0.061	(0.108)

country is $D = 0$. The real change comes in the extent to which the impact of G5 R&D declines with distance. Now the estimate of δ is a much smaller 0.384 indicating that the effect declines much more slowly with distance, even in the first half of the sample when $I_t = 0$. Over the latter half though when I_t is 1, the estimate of $\psi_D = -0.784$ indicates that this small rate of decline is even smaller from 1983 to 1995 than it was previously. Together, these estimates indicate that at least half of the effect that a G5 nation's R&D would have had if the domestic country had been right next to it ($D = 0$) is still there as far out as 424 kilometers (263 miles) from 1970 to 1982, and is felt as far out as 1,963 kilometers (1,217 miles) after 1983. This implies a larger and growing degree of technical spillovers between industries in different nations.

Because Keller's spillover estimates apply to whole sectors separated by national boundaries they may well be a lower bound for the extent of such spillovers between firms within the same domestic industry. If this is so, then these empirical estimates when taken together with our analysis of the noncooperative outcome with high spillovers, suggest that the market will likely be characterized by inefficiently low R&D. If that is the case, then the argument for permitting R&D cooperation and/or joint ventures becomes noticeably more compelling.

Summary

Research and development is the wellspring of technical advancement. Such advancement is the true source of the gain in per capita income and living standards that has characterized the developed economies for almost all of the last two centuries. It should be clear, however, that firms will only be willing to incur the heavy expenses and considerable risks associated with R&D if they can be reasonably assured that their efforts will be rewarded. Imitation by rivals has the social benefit of intensifying price competition after innovation occurs. However, it makes it less likely that the innovation will occur in the first place.

The tension between the gains from competition and the gains from innovation, i.e., the tension between the replacement effect and the efficiency effect is unavoidable. It has led economists to consider which market environment—competitive or monopolistic—will foster greater research and development. The Schumpeterian hypothesis is, broadly speaking, that oligopolistic market structures are best in this regard.

Both theory and empirical data give ambiguous evidence as to the market structure most conducive to R&D effort. Competitive markets can sometimes fail to be as innovative as their

less-competitive counterparts but a surprising number of key inventions have come from small firms. Policy has a role to play here, too. One role for policy is to encourage cooperation in research efforts. Empirical evidence suggests that we live in an increasingly interconnected world in which the benefits from one firm's R&D spill over to other firms including, in particular, its rivals. In such a world, the noncooperative outcome is likely to be one with too little R&D effort. Policy that fosters research cooperation among firms can be helpful in this setting. Yet caution is also necessary. The trick is somehow to foster cooperation on R&D without simultaneously inducing collaboration on prices and product design.

A similar tension arises in the role of patent policy. Patents can enhance the incentives for firms to pursue technological innovations. Yet, by temporarily granting monopoly power, patents can also weaken competitive forces and reduce consumers' access to those breakthroughs. We consider patents and related policy issues in the next chapter.

Problems

1. Assume that inverse demand is given by the linear function $P = A - BQ$ and that current marginal costs of production are c.
   a. By how much would an innovation have to reduce marginal cost for it to be a drastic innovation?
   b. Use your answer to derive a condition on the parameters A, B and c that determines whether a drastic innovation is feasible. (Hint: costs cannot be negative.)

   For Problems 2 through 5 assume the following: Inverse demand is given by $P = 240 - Q$. The discount factor is 0.9. Marginal production costs are initially \$120:

2. Calculate the market equilibrium price, output, and profits (if any) on the assumption that the market is currently
   a. monopolized,
   b. a Bertrand duopoly,
   c. a Cournot duopoly.

3. Suppose that a research institute develops a new technology that reduces marginal costs to \$60.
   a. Confirm that this is not a drastic innovation in either the Bertrand or Cournot cases.
   b. Calculate the new market equilibrium price, output, and profits for the monopolist and each duopolist, given that in the duopoly case the innovation is made available to only one firm.
   c. How much will the monopolist and duopolist each be willing to pay for the innovation?

4. Now assume that there is a potential entrant in the monopolized case and that the research institute is considering offering the innovation to this firm as well as to the monopolist. How does this affect the amount that the incumbent monopolist will be willing to pay for the innovation?

5. Now return to the duopoly case but assume that the research institute is considering whether it should actually sell the innovation to both firms. Will it wish to do so
   a. in the Bertrand duopoly?
   b. in the Cournot duopoly?

6. Assume that annual inverse demand for a particular product is $P = 150 - Q$. The product is offered by a pair of Bertrand competitors, each with marginal costs of \$75. The discount factor is 0.9. What is the current equilibrium price and total surplus?

7. Return to problem 6. Assume now though that if R&D is conducted at rate x, it incurs one-off costs of $r(x) = 10x^2$ and reduces marginal costs to $(75 - x)$. Suppose that one firm decides to conduct R&D at rate $x = 10$. This research will be protected by a patent of T years.
   a. What profit (ignoring the one-off costs of R&D) does the innovating firm make each year during the period of patent protection?
   b. What is the new equilibrium price and total surplus once patent protection expires?

8. Use your answers to 7(a) and (b) to write the total net surplus from the innovation as a function of the period of patent protection.

Derive (numerically) an approximation to the socially optimal period of patent protection.

9. How are your answers to 7(a) and (b) affected if the innovating firm conducts research at rate $x = 15$?
10. How does the net social surplus change if the innovating firm conducts research at $x = 15$?
11. What research intensity will the firms choose given that the period of patent protection is set optimally?

References

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Inventions." In R. Nelson, ed., *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton: National Bureau of Economic Research, Princeton University Press.

Barro, R. and X. Sala-i-Martin. 1995. *Economic Growth*. New York: McGraw-Hill.

Blundell, R., R. Griffith, and J. Van Reenen. 1995. "Dynamic Count Data Models of Technological Innovation." *Economic Journal* 105 (March): 333–44.

Cohen, W. and R. Levin. 1989. "Empirical Studies of Innovation and Market Structure." In R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization*. Vol. 2. Amsterdam: North-Holland, 1059–98.

Cohen, W. and S. Klepper. 1996. "A Reprise of Size and R and D." *Economic Journal* 106 (July): 925–51.

Dasgupta, P. and J. Stiglitz. 1980. "Industrial Structure and the Nature of Innovative Activity." *Economic Journal* 90 (January): 266–93.

d'Aspremont, C. and A. Jacquemin, 1988. "Cooperative and Noncooperative R&D in Duopoly with Spillovers." *American Economic Review* 78 (September): 1133–7.

Gayle, P. 2002. "Market Structure and Product Innovation." Working Paper, Department of Economics, Kansas State University.

Geroski, P., 1990. "Innovation, Technology Opportunity and Market Structure," *Oxford Economic Papers* 42 (July): 586–602.

Gilbert, R. J. 2006. "Competition and Innovation," *Journal of Industrial Organization Education* 1 (1): Article 8. Available at: http://www.bepress.com/jioe/vol1/iss1/8

Gilbert, R. J. and D. M. G. Newbery. 1982. "Preemptive Patenting and the Persistence of Monopoly." *American Economic Review* 72 (June): 514–27.

Kamien, M. I., E. Muller, and I. Zang. 1992. "Research Joint Ventures and R&D Cartels." *American Economic Review* 82 (December): 1293–1306.

Keller, W. 2002. "Geographic Localization of International Technology Transfer." *American Economic Review* 92 (March): 120–42.

Klepper, S. 2002. "Firm Survival and the Evolution of Oligopoly." *Rand Journal of Economics* 33 (Spring): 37–61.

Levin, R. and P. Reiss. 1984. "Tests of a Schumpeterian Model of R and D and Market Structure." In Z. Griliches, ed., *R and D, Patents and Productivity*. Chicago: NBER University of Chicago Press.

Levin, R., W. Cohen, and D. C. Mowery. 1985a. "R and D Appropriability, Opportunity, and Market Structure: New Evidence on Some Schumpeterian Hypotheses." *American Economic Review, Papers and Proceedings* 75 (May): 20–4.

——. 1985b. "Firm Size and R&D Intensity: A Reexamination." *American Economic Review* 75 (June): 543–65.

Levin, R., A. Klevorick, R. Nelson, and S. Winter. 1987. "Appropriating the Returns from Industrial Research and Development." *Brookings Papers on Economic Activity: Microeconomics* 2: 783–822.

Lunn, J. 1986. "An Empirical Analysis of Process and Product Patenting: A Simultaneous Equation Framework." *Journal of Industrial Economics* 34 (February): 319–30.

Markides, C. and P. Geroski. 2005. *Fast Second: How Smart Companies Bypass Radical Innovations to Enter and Dominate New Markets*. San Francisco: Jossey-Bass.

Porter, M. 1990. *The Competitive Advantage of Nations*. New York: Free Press.

Romer, D. 2006. *Advanced Macroeconomics*. 3rd edition. New York: McGraw-Hill/Irwin.

Scherer, F. M. 1965. "Firm Size, Market Structure, Opportunity and the Output of Patented Innovations." *American Economic Review* 55 (September): 1097–125.

——. 1967. "Market Structure and the Employment of Scientists and Engineers." *American Economic Review* 57 (June): 524–31.

Schumpeter, J. A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper.

Scott, J. T. 1990. "Purposeful Diversification of R&D and Technological Advancement." In A. Link, ed., *Advances in Applied Micro-economics*. Vol. 5. Greenwich, CT, and London: JAI Press.

Solow, R. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70 (February): 65–94.

von Hippel, E. 1988. *The Sources of Innovation*, New York: Oxford University Press.