Chapter 17
Technological Change, Dynamic Efficiency, and Market Structure

To this point, most of our discussion about economic performance has focused on static efficiency and assumed that technology is fixed. Yet economic growth requires that we make investments today to develop better products or new processes that lower the cost of production. Persistent long-run economic growth has led to a continued rise in our standard of living. For example, Elwell (2006) documents that from 1980 to 2004, output per capita grew by about 2.3% per year in Great Britain and by about 2.0% in the USA, Japan, and other major European countries (Germany, France, Italy, and the Netherlands). Although these growth rates may seem inconsequential, a small increase in the growth rate can have a sizable cumulative effect. To illustrate, a 2% growth rate will double the standard of living in approximately 35 years, while a 3% growth rate doubles it in only about 23.5 years.[1]

We can see the effect of growth in our lives by comparing living standards today with a century ago. In 1900, about 40% of Americans could be classified as poor by current standards. About 60% of people lived on farms or in rural areas. The average home did not have electricity, indoor plumbing, or a telephone. Only about 7% of youth completed high school, walking was the most common form of transportation, and average life expectancy was 47 years, about 30 years lower than today.

Even if you were exceptionally wealthy in 1900, your life would still be constrained in many ways compared to today. On the plus side, you could own a large home and employ servants who could cook, clean, and launder cloths. These would be valuable services in an era without the benefit of microwave ovens, vacuum cleaners, dishwashers, and washing machines. However, you would be unable to fly to Europe or own a radio, TV, Blu-Ray player, computer, or

[1] The "rule of 72" provides an approximation. That is, if the annual growth rate is x%, then the standard of living will double in approximately $72/x$ years.

V.J. Tremblay and C.H. Tremblay, *New Perspectives on Industrial Organization*, Springer Texts in Business and Economics, DOI 10.1007/978-1-4614-3241-8_17,

cell phone. If you were a woman, you could vote in only four Western states. In addition, entertainment opportunities were very limited. There were few spectator sports, and Vaudeville[2] was the common form of entertainment.

A number of factors contribute to long-run growth in income. First, there must be an adequate infrastructure. This includes legal and other institutions and cultural attitudes that support entrepreneurial activity. Once these are in place, living standards depend on the quantity of physical capital, the level of human capital, and the level of technology. Technology or technical knowledge refers to the entire body of knowledge concerning the methods used to bring inputs together to produce goods and services. Thus, **technological change** occurs when we add to technological knowledge.

You can think of technological knowledge as information that is currently publically available and technological change as new knowledge that is created and will become publically available. Parallel to technological knowledge is human capital, which is defined as a person's level of knowledge (of publically available information); investments in education increase human capital, while investments in research and development (R&D) produce technological change. Each has a positive effect on the other, for the most part, as it takes an educated population to create new knowledge.

Economists have tried to identify the extent to which these factors contribute to economic growth in the USA. Recent evidence[3] indicates that capital accumulation contributed almost 50% to US economic growth, increases in human capital contributed about 20%, and technological change contributed 40% to 50%.[4]

Given the economic importance of technological change, public and private institutions spend a great deal of money each year on R&D. This amount varies considerably across nations, industries, and firms. Figure 17.1 plots R&D spending as a percent of gross domestic product (GDP) for a sample of countries in 2005. In general, the intensity of R&D spending is greater among developed countries. **Patents**, which give an inventor limited ownership of a new idea or method, are observable outcomes of R&D that approximate the extent of technological change. Table 17.1 lists the top ten companies that were awarded patents in the USA in 2009. As you can see, the leaders are from the high tech computer and electronics sectors of the economy. This is consistent with the evidence in Table 17.2, which identifies the most innovative industries in the USA. It shows that the chemical, computer, and electronics industries are highly innovative. In contrast, very few patents have been awarded to companies involved with simple manufacturing such as button making, needle making, typesetting, and book binding.[5]

---

[2] Vaudeville was live theater by circus entertainers, comedians, dancers, and musicians.

[3] For a review of the evidence, see Cohen and Levin (1989), Mankiw et al. (1992), Jorgenson and Stiroh (2000), DeLong et al. (2003), and Elwell (2006).

[4] Percentages exceed 100% because other factors, such as an increase in government regulation and a shorter average work week, have reduced economic growth. Early studies by authors such as Denison (1985) gave an even higher contribution to technological change (at over 60%).

[5] The US Patents and Trademark Office publishes patent information by country, industry, and company at http://www.uspto.gov.

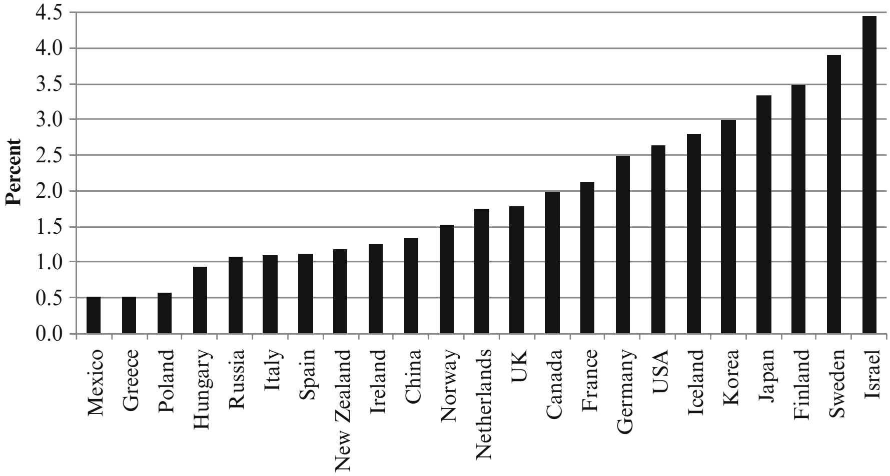

Fig. 17.1 R&D as a percent of gross domestic product by country, 2005

Table 17.1 Top ten patent-receiving firms in the USA, 2009

Company	Number of patents
International Business Machines (IBM)	4,887
Samsung Electronics	3,592
Microsoft	2,901
Cannon	2,200
Panasonic	1,759
Toshiba	1,669
Sony	1,656
Intel	1,534
Seiko-Epson	1,328
Hewlett-Packard	1,269

Source: US Patent and Trademark Office at http://www.uspto.gov

Table 17.2 Top ten patent-receiving industries in the USA, 2009

Industry	Patents granted in 2009	Cumulative patents
Multiplex communications	5,304	42,044
Semiconductor device manufacturing	4,908	68,955
Telecommunications	3,372	34,798
Chemistry: molecular biology and microbiology	2,710	57,381
Image analysis	2,625	26,176
Computer graphics	2,597	32,892
Static information storage and retrieval	2,384	33,891
Pulse and digital communications	2,280	28,005
Radiant energy	2,184	37,441
Electricity: electrical systems and devices	2,122	32,733

Source: US Patent and Trademark Office at http://www.uspto.gov

Because technological change is dynamic in nature and crucial for economic growth, we are just as interested in dynamic efficiency as static efficiency. Static efficiency ignores technological change, while **dynamic efficiency** recognizes that investments in R&D today can be socially desirable if they lead to a better life tomorrow. Dynamic efficiency occurs when there is an optimal amount of technological change from society's perspective.[6] Because one important invention can quickly outstrip the deadweight loss due to market power, maintaining static efficiency period after period need not be dynamically efficient. We will see that it may be socially desirable to give temporary market power to an inventor (which is statically inefficient) to encourage inventive effort (which is dynamically beneficial).

In the sections that follow, we discuss three issues regarding technological change. First, we describe the economics of technological change. Then we analyze a firm's motivation for investing in R&D. We pay particular attention to the connection between market structure and technological change. We will see that technological change can affect market structure, and market structure can affect technological change. A fundamental issue in the field of industrial organization is whether or not one market structure is more dynamically efficient than another. Finally, we discuss the empirical evidence regarding these issues.

17.1 Invention and Technological Change

17.1.1 The Economics of Technological Change

As discussed in the introduction, technological change produces an increase in technological knowledge. This includes the ability to conceive of a completely new product, such as the Internet in the 1980s. It also includes the discovery of new methods to produce existing products of higher quality (at the same cost) or existing products at lower cost (i.e., with fewer inputs).

The electronic calculator provides an excellent example. When first mass marketed in the early 1970s, a simple calculator that could add, subtract, multiply, and divide had a retail price of about $100. Subsequent technological change produced two major advances. First, it led to the development of a cheaper microcoprocessor, enabling simple calculators to sell for less than $5 today. Second, it produced coprocessors that could handle a wider range of calculations. As well as perform simple arithmetic, today's $100 calculator can solve calculus, trigonometric, and financial problems. Many newer calculators also have a memory and can graph a variety of functions.

---

[6] As with most economic problems, this involves tradeoffs because research and development is costly and technological change can have undesirable consequences. For example, new technologies have made mass killing more efficient and have sometimes increased the level of pollution and market power. As we discussed in Chap. 1, equity, fairness, a clean environment, and macrostability are also important to social welfare. In this chapter, we focus on technological change that is beneficial and postpone discussion of these broader concerns to Chaps. 19 and 20.

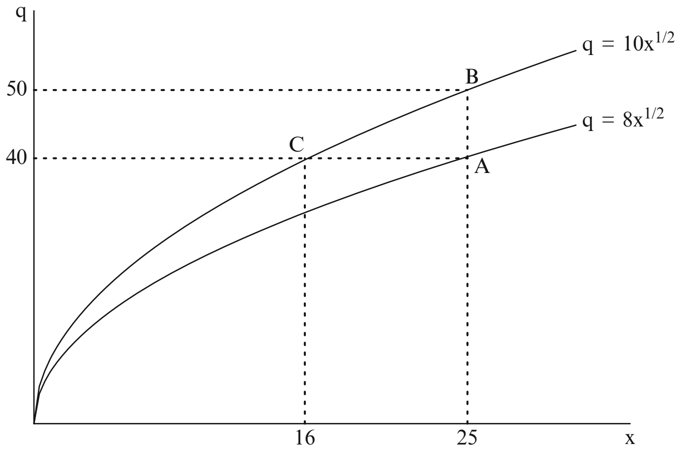

Fig. 17.2 Technological change and the production function, $q(x)$

A technological change enables a firm to produce a given output with fewer inputs, which produces an upward shift in the production function. To illustrate, consider a production function where the quantity of output (q) is a function of a single input (x) and takes the following form: $q = \alpha \cdot x^{1/2}$. Initially, α equals 8, but a technological change that makes x more productive raises α to 10. Figure 17.2 graphs these functions and shows that the production function shifts up and to the left with a technological change. That is, to produce 40 units of output, the firm uses 25 units of x with the old technology and 16 units of x with the new technology. When looking at this as an upward shift of the production function, a technological change enables the firm to produce more output (50 instead of 40) with the same quantity of input x (25).

Another way to see the effects of a technological change is through its influence on a firm's cost function. When a technological change enables a firm to product the same output with fewer inputs, costs fall. An example is provided in Fig. 17.3, where AC represents the long-run average cost function for the old technology and AC′ represents it for the new technology. In this case, technological change has no effect on minimum efficient scale (MES). This is a **scale-neutral technological change**. When a technological change causes MES to increase (decrease), it is said to be a **scale-increasing (scale-decreasing) technological change**. An example of a scale-increasing technological change is depicted in Fig. 17.4.

As well as affecting scale, a technological change can affect substitution possibilities between inputs. This becomes apparent by reviewing the effect of a technological change on a firm's isoquants, which is described in Fig. 17.5 for two inputs, labor (L) and capital (K). With the old technology, the firm can produce 100 units of output in an economically efficient manner (i.e., at minimum cost) with 30 units of L and 3 units of K. This occurs where the isoquant ($\bar{q}_{100}$) is tangent to the

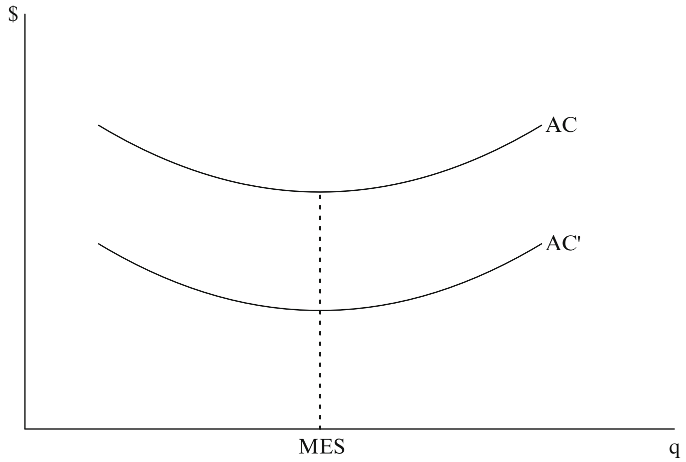

Fig. 17.3 Long-run average cost and a scale-neutral technological change

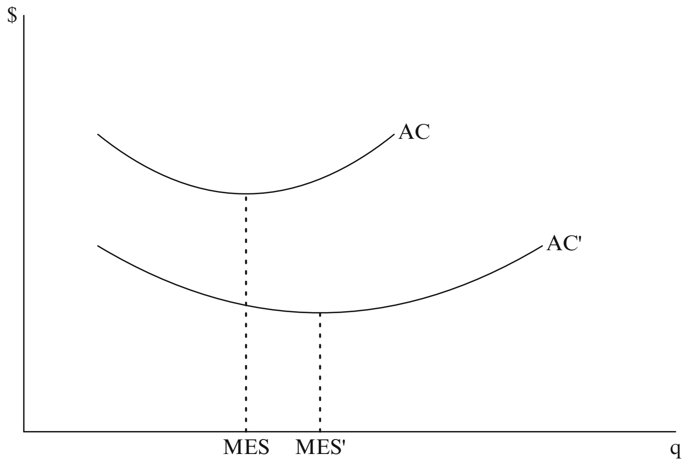

Fig. 17.4 Long-run average cost and a scale-increasing technological change

isocost function ($\bar{C}_{10}$) at point A.[7] With the new technology, the isoquant shifts towards the origin (q'_{100}), and the optimum moves from point A to point B. That is, the firm is able to use less L and K to produce the same units of output. Notice that the optimal labor–capital ratio (L/K) stays the same. With both technologies, the

[7] An isoquant maps out the minimum combinations of L and K that will produce a given level of output ($\bar{q}$). An isocost function maps out all combinations of L and K that produce a given total cost ($\bar{C}$). For further discussion, see Varian (2010).

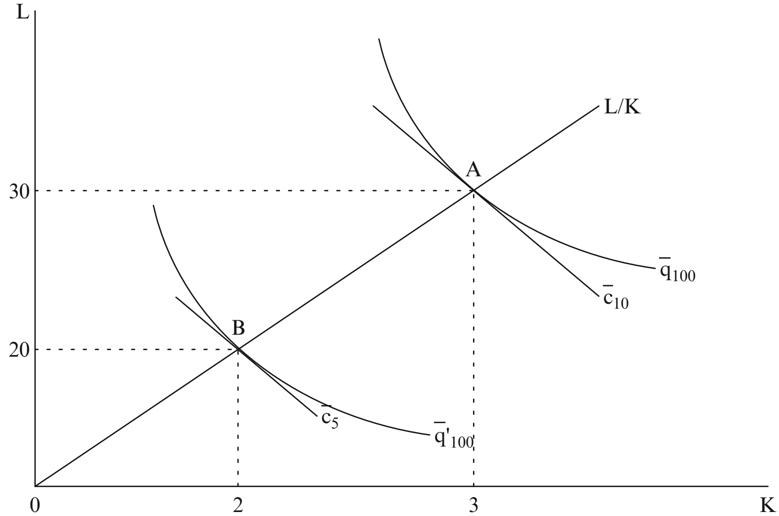

Fig. 17.5 Isoquant and isocost functions for a neutral technological change

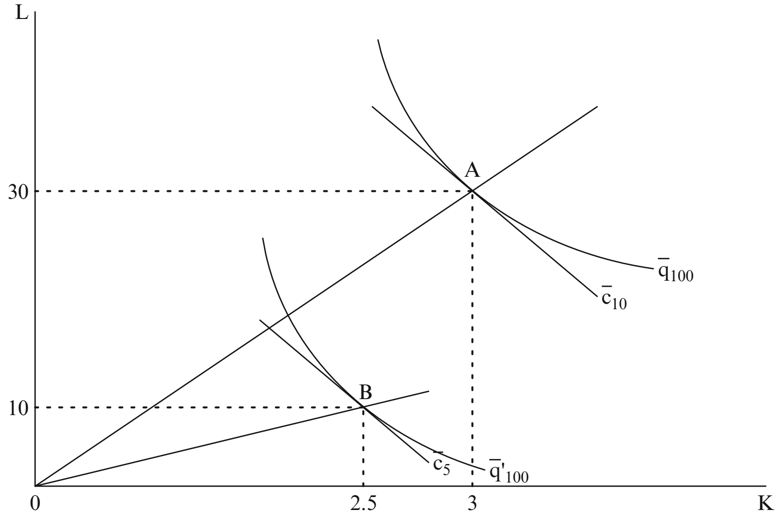

Fig. 17.6 Isoquant and isocost functions for a labor-saving technological change

optimal ratio is 10 units of labor for each unit of capital (i.e., $L/K = 30/3 = 20/2 = 10$). Because the technological change has a neutral effect on the optimal labor–capital ratio, it is called a **neutral technological change**.

There is no guarantee that technological change will be neutral, however. Figure 17.6 illustrates the case of a **labor-saving technological change**. With the old technology, the labor–capital ratio is 10, but with the new technology it is

4 (10/2.5). That is, the new technology now requires only 4 units of labor for each unit of capital. A **capital-saving technological change** occurs when the labor–capital ratio rises with a technological change, making it optimal for the firm to use more labor relative to capital.

If firms have some control over the change in technology through focused R&D efforts, then they can direct technological change in a profit-maximizing direction. For example, in the last half century US automobile producers have developed robots to replace labor on assembly lines. This would have been a rational response to the rise in the price of labor relative to capital in the USA, which may have been anticipated by the automobile industry. The point is that if a firm expects the relative price of an input to rise over time, it pays the firm to cultivate new technologies that would use relatively less of the more expensive input.

17.1.2 A Taxonomy of Research, Invention, and Technological Change

R&D programs can be divided into two types. The first is **basic research**, which involves a theoretical or experimental investigation that is designed to advance scientific knowledge without regard to a specific application. A classic example is the research by physicists on atomic structure before World War II. When first developed, this knowledge was of no apparent value, but after decades of work it led to today's Global Positioning System.[8] The second type is **applied research**, which is designed to create knowledge that has a specific practical purpose. A scientist working on a more effective allergy medicine is conducting applied research.

Basic research has a public quality, as it can benefit a wide range of users simultaneously. As a result, it is typically undertaken by researchers at major universities and research institutes and is funded by government agencies, such as the National Science Foundation, National Institute of Health, and NASA. Applied research is funded by private and public agencies. Data from the National Science Foundation show that R&D spending as a percent of Gross Domestic Product (GDP) reached a peak in 1964 and hovered between 2.2% and 2.8% since 1980 (see Fig. 17.7).[9] Government support for basic research has declined, however, equaling 0.7% of GDP in 1953 and only 0.2% in 2004 (Elwell 2006, 28).

Schumpeter (1934) described three stages of technological change. The first stage is **invention**, the act of creating a new idea or of solving a promising technical possibility. This is the initial research phase of a R&D program. The next is **innovation**, which occurs when an invention is applied for the first time and results in a new product or production process. The final step is **diffusion** (or imitation) in which the final innovation becomes widely used.

---

[8] For further discussion, see the Committee on Science, Engineering, and Public Policy (1999).

[9] The data are available at http://www.nsf.gov/statistics/nsf10314/content.cfm?pub_id=4000&id=1.

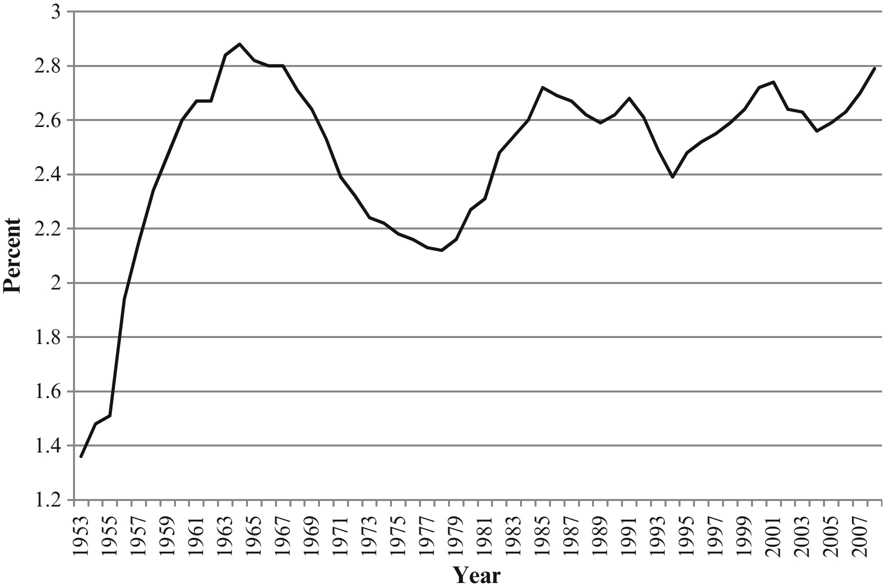

Fig. 17.7 US research and development expenditures as a percent of GDP, 1953–2008

Historical examples show that it can take a considerable amount of time to make a viable invention market worthy.[10] One example is the development of the steam engine. Although it was first invented by James Watt in 1765, it needed a considerable amount of work (i.e., time and money) to make it ready for industrial purposes. Not until Matthew Boulton stepped in with financial support was Watt able to complete the final stages of product development, which did not occur until 11 years after its initial invention. Development of photocopying by Xerography suffered similar development problems. It was first invented by Chester Carlson in 1938, but it took 21 years and $20 million before it was made available for commercial use by the Xerox Corporation.

Behavioral economics provides further insight into the rate of diffusion of new technologies. As discussed in Chap. 14, Rogers (1962) theory that personality differences explain the acceptance rates of new products among consumers can also be used to explain the diffusion rate of new technologies among producers. A classic example is the 1928 introduction of hybrid seed corn in Iowa. This new seed was more profitable than traditional seed due to its greater resistance to drought and disease. Nevertheless, it took over 13 years for all farmers in Greene County, IA, to adopt it.

Figure 17.8 shows that the adoption rate of hybrid seed corn among Iowa farmers follows a pattern that is consistent with Rogers' (1962) theory that the diffusion rate

[10] See Scherer (1965) and Jewkes et al. (1969).

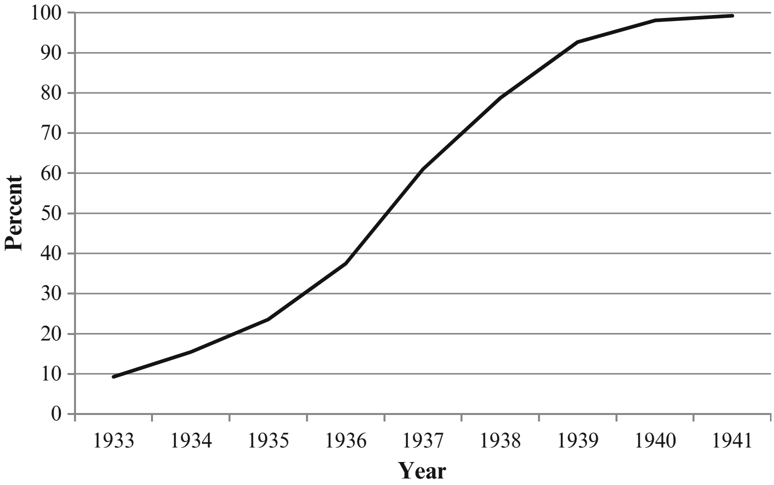

Fig. 17.8 Percent of Iowa farmers who adopted hybrid seed corn, 1933–1941

is based on personality differences among adopters. According to Rogers, a new technology is first accepted by "early adopters," those who place a high value of being first and are opinion leaders within the industry. Then it is adopted by the "early and late majority," which includes individuals who have a certain degree of skepticism about change. Finally, it is put into practice by the "late adopters," those who hold traditional values and are adverse to change.[11] Thus, personality differences among business owners explain why it can take a considerable amount of time for a new technology to become an industry norm, even for a very profitable one like hybrid seed corn.

In real-world examples such as these, it is important to distinguish between invention and innovation. Nevertheless, these terms are used interchangeably in the theoretical literature, a practice we follow in the remainder of the chapter.

17.2 Failure of the Market for Ideas

An important concern with the market for new ideas is that knowledge is a public good, and free markets undersupply public goods from society's perspective. Once created, there is nonrivalry in consumption because everyone can benefit from a new idea and put it to use at the same time.

[11] Rogers actually divided individuals into five categories: innovators, early adopters, early majority, late majority, and laggards.

In a completely free market, it will also be impossible (or uneconomic) for you to exclude others from imitating your invention or stealing your ideas. Because an inventor will be unable to recover the benefits others receive from his or her innovation, inventors will have too little incentive to create new ideas from society's perspective. For example, assume that you have an idea for a new type of battery that lasts 10 times longer but costs the same to produce as current batteries. To convince a company that your idea is sound, you must reveal your idea. But once you have done so, the company has no incentive to pay you for it. The public nature of new ideas is a classic form of market failure.

When markets fail in this way, government intervention can improve social welfare. Examples of polices designed to encourage investments in creative endeavors include patents, copyrights, trademarks, and research grants.[12] To stimulate technological change, one option is for government to subsidize R&D. In the USA, corporations are given a 20% tax credit for their R&D expenditures. In addition, the federal government provides grants to support R&D. Of the $398 billion spent on R&D in the USA in 2008, 26% was financed by the federal government. The percent of support from the federal government reached a peak at almost 67% in 1964, but has declined fairly steadily ever since. This undoubtedly has reduced total R&D spending, given Martin and Scott's (1998) finding that government-supported R&D does not crowd out privately funded R&D. Another concern is that this decline in federal support will change the composition of R&D spending, as the federal government is more likely than business to fund basic research.

An important government response to the market failure associated with ideas is a realization that it is fundamentally a property rights problem. That is, if you owned the right to a new idea or creative work, you could then sue for damages if someone stole it. Our founding fathers recognized this property rights problem as it applies to technological change, as Article I, section 8 of the US Constitution states:

> The Congress shall have power . . . to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries.

This establishes the foundation for our patent laws, which grant a property right to the inventor of a new idea or method that lasts for a limited amount of time.[13] Property rights such as these encourage innovation, because they facilitate **appropriability** (the ability of an inventor to capture the gains from an invention or new idea) and make inventive activity worthwhile. There can still be diffusion of

---

[12] Given the problem with imitation, Keller (2002) suggests that permitting joint ventures and cooperation in R&D may also increase inventive activity.

[13] This property rights issue also motivates our copyright legislation, which gives creators ownership of their artistic expression, and trademark laws, which protect a company's words or symbols used to identify a firm's particular brand or identity. Because these encourage creativity much like a patent, we focus primarily on patents in this chapter. You can learn more about patent, copyright, and trademark law from the Web page of the US Patent and Trademark Office at http://www.uspto.gov.

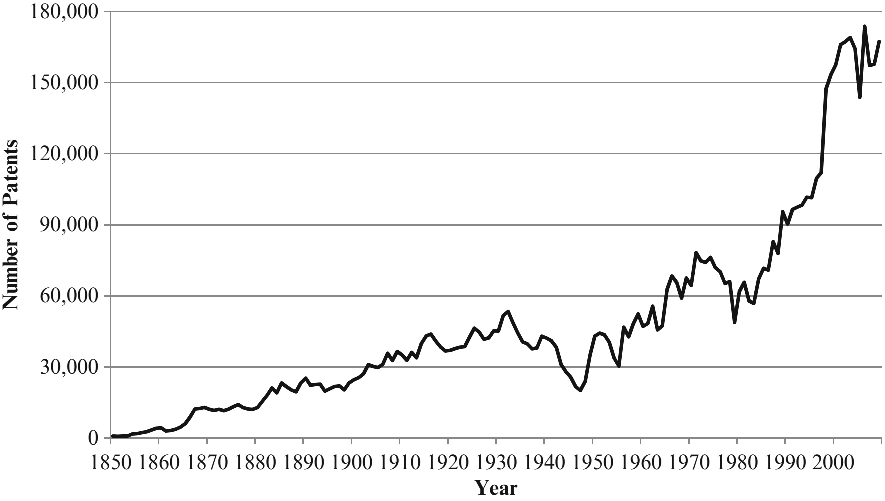

Fig. 17.9 Annual number of patents granted in the USA, 1850–2009

an invention, because clearly defined property rights facilitate the trade of a patented idea or of a licensing agreement that allows others to use the idea for a fee.

The first US patent act was approved by Congress in 1790. In exchange for disclosing the new idea to society, an inventor was granted exclusive ownership of the invention or innovation for 14 years. To receive a patent, the inventor must demonstrate that the invention or innovation is new, useful, and nonobvious.[14] A patent encourages R&D by stopping imitation and giving monopoly ownership to the inventor for a limited period of time. Since 1790, over 7.6 million patents have been awarded in the USA. Patent activity was meager until the middle of the nineteenth century when technological change began to drive the industrial revolution. You can see this from Fig. 17.9, which plots the number of patents granted each year from 1850 to 2009. In 1850, 884 patents were granted, a number that has risen steadily for most of the subsequent period. Several events undoubtedly influenced the trend in patent activity:

- 1861: Patent life was increased from 14 to 17 years, which would increase the value of a patent and encourage R&D.
- 1942–1945: US involvement in World War II diverted funds from R&D and may explain the decline in patent activity during this period.
- 1980–1981: US courts extended patentability to include genetically engineered bacteria in 1980 and software in 1981, which would encourage R&D.

[14] For a more complete description of the patent process, see Merges et al. (1997).

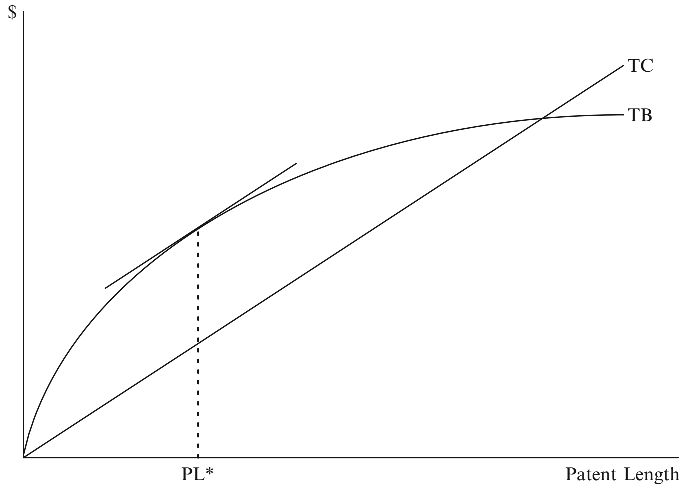

Fig. 17.10 The socially optimal patent length (PL*)

- 1995: To conform to the World Trade Organization's Agreement on intellectual property rights issues, patent life was increased from 17 to 20 years. Again, this would encourage R&D activity.
- 1998: US courts extended patentability to include business practices[15] and financial service products, which would encourage R&D.

A patent's market value increases with patent length, so a longer life should increase the incentive to invest in creative activity, which in turn would lead to greater technological change. This begs the question: why not give an inventor a patent that lasts forever?

To identify the socially optimal patent life, we need to realize that there is a cost–benefit trade-off associated with increasing patent life.[16] On the plus side, a longer life increases the expected returns associated with an innovation, which will encourage R&D activity and promote dynamic efficiency. On the other hand, a patent gives its owner a monopoly over the use of the innovation. The longer the patent life, the greater the static inefficiency associated with the monopoly power that is created by the patent. One would expect the social benefits of a longer patent life to increase at a decreasing rate, assuming diminishing returns. Social costs should rise as well, as a longer life implies greater static inefficiency. To identify the social optimum, consider the example of total benefit (TB) and total cost (TC) functions found in Fig. 17.10. The optimum occurs where the difference between TB and TC is greatest or where the marginal principle holds (i.e., the marginal benefit equals the marginal cost). Patent life (PL) is optimal at PL^* in this example.

[15] One example is Amazon.com's one-click method of placing an order on the internet.

[16] For an early discussion of this issue, see Nordhaus (1969).

The obvious problem with this analysis is that the benefits and costs of a longer patent life vary by the type of invention. Those with exceptionally high (low) benefits warrant a longer (shorter) life, while those that produce greater (less) static inefficiency warrant a shorter (longer) life. As a practical matter, it is very difficult and costly to predict these things when an inventor first applies for a patent. Even the inventor, let alone a government agency, would be uncertain of the potential merits of an innovation. Given this uncertainty, a system with a fixed patent life that applies to all innovations may be the best we can do.

17.3 The Effect of Market Structure on Technological Change

In this section, we ignore government involvement and focus on market incentives to invest in R&D that is designed to create profitable new technologies. Although a patent provides ownership for a limited period, it also forces the firm to reveal its new idea for its competitors to see. In some cases, the firm may prefer to keep the new idea a **trade secret** rather than disclose it in a patent application. In fact Levin et al. (1987) and Moser (2005) found that secrecy can be an effective alternative to patent protection in certain industries.

The Coca-Cola Company provides one of the best examples of a successful trade secret. The company has never held a patent for the formula of its Coke brand of cola but has been very effective in protecting it from outsiders. Only a few employees know the formula, and each of them has signed a contract with the company that forbids them from disclosing the formula to others and from opening their own cola company. This strategy has allowed Coca-Cola to conceal its formula for well over a century.

Schumpeter (1942) was the first to emphasize the significance of dynamic efficiency and to analyze the possible connection between market structure and technological change. According to Schumpeter, competition for new technologies is an essential part of capitalism and is more important than price competition. Firms compete to develop new technologies that can be extremely profitable. Through a process that Schumpeter called **creative destruction**, the development of new technologies continually revolutionizes our economy by destroying old ways (i.e., methods, companies, and markets) and creating new ones.

Regarding market structure, Schumpeter went on to say that the statically efficient model of perfect competition "has no title to being set up as a model of ideal efficiency." Instead, large firms in concentrated industries are necessary for dynamic efficiency because they are more likely to invest in the R&D that drives creative destruction through technological change. One possibility is if there are economies of scale in R&D, then large firms will have an innovation advantage.

The presence of market power that is typically associated with large corporations may encourage innovative activity in two ways. First, existing market power gives firms the financial <u>means to invest</u> in R&D. Second, the potential for successful innovation to increase market power gives firms the <u>incentive to invest</u> in R&D.

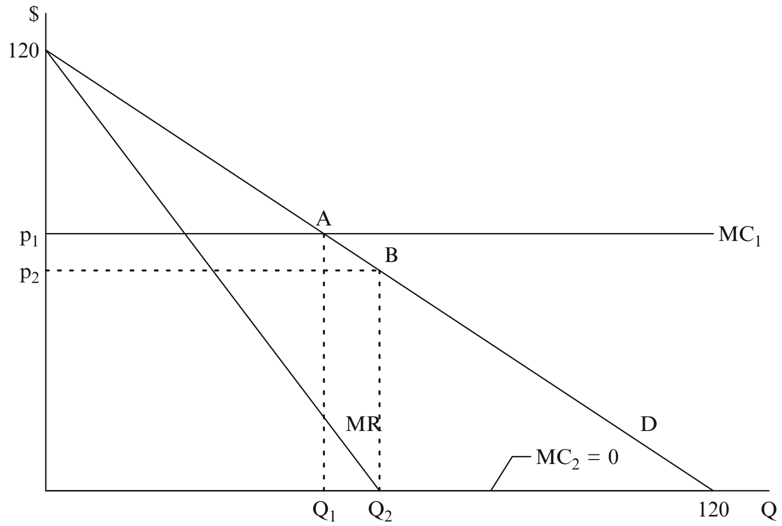

Fig. 17.11 Technological change that transforms an industry from perfect competition (at A) to monopoly (at B)

In the end, this drive to invent leads to better products and processes. Through creative destruction, it also makes the benefits of market power short lived. In Schumpeter's view, large corporations are necessary for progress and are constantly being challenged and replaced by more dynamically efficient competitors who produce better products at lower cost.

To illustrate Schumpeter's point that maintaining static efficiency period after period need not be dynamically efficient, we consider a simple two-period model. Demand and cost conditions are described in Fig. 17.11, where D is demand, MR is marginal revenue, and MC is long-run marginal cost. In period 1, marginal cost is $MC_1 = \$70$, and the market is perfectly competitive. Equilibrium price and quantity are $p_1 = \$70$ and $Q_1 = 50$. Given the lure of monopoly profits, one firm invests in R&D which pays off and lowers its marginal cost to $MC_2 = 0$. This innovation puts all other firms out of business, leaving the inventive firm with a monopoly position, and the optimal price–output pair becomes $p_2 = \$60$ and $Q_2 = 60$. Notice that the market is statically efficient in period 1 (because $p_1 = MC_1 = 70$) but is not statically efficient in period 2 (because $p_2 = 60 > MC_2 = 0$).

This outcome can be dynamically efficient, however, if the cost of R&D is not too high and if the inventive firm would not have undertaken the R&D project without the payback of monopoly profits in period 2. To demonstrate, we compare total surplus in the two alternatives (ignoring discounting):

- Without the innovation, $p_1 = p_2 = MC_1$ in equilibrium. Total surplus is the area under demand and above MC_1 at Q_1 in period 1 (i.e., consumer surplus in period 1) plus the area under demand and above MC_1 at Q_1 in period 2 (i.e., consumer surplus in period 2). This is \$2,500 (\$1,250 in each period).

- With the innovation (ignoring the cost of R&D for the moment), total surplus is the area under demand and above MC_1 at Q_1 in period 1 ($1,250) plus the area under demand and above MC_2 at Q_2 in period 2 (i.e., consumer plus producer surplus in period 2). This is $6,650 (i.e., total surplus of $1,250 in period 1 plus consumer surplus of $1,800 in period 2 and gross profit of $3,600 in period 2).

As long as the cost of R&D is no greater than the firm's gross profit of $3,600, the firm will invest and the innovation produces a higher total surplus over the two periods. This outcome is dynamically efficient because society is better off with the innovation even though it leads to static inefficiency in period 2. In other words, the benefits of innovation far outstrip the costs of static market power.

Given the importance of Schumpeter's ideas, the purpose of this section is to investigate the link between market structure and innovative activity. We begin by analyzing monopoly, perfect competition, and oligopoly models. We examine the empirical evidence at the end of the chapter. Issues of technological change are clearly dynamic, but to keep things simple we begin by ignoring time.

17.3.1 Monopoly and Technological Change

First, we consider a monopoly market where the firm invests in R&D that produces a new technology. Two cases are analyzed: (1) product innovation where technological change increases product quality; (2) process innovation where technological change lowers production costs.

17.3.1.1 Technological Change that Improves Product Quality

The firm's goal is to maximize profit with respect to price and R&D spending. Investing in R&D is costly and increases demand by improving product quality. Firm (market) demand is defined as $q(p, \text{R\&D})$, where q is quantity demanded and p is price. Demand has a negative slope and increases at a decreasing rate with spending on R&D.[17] Total cost (TC) consists of total production cost, $C(q)$, and spending on research and development, R&D. The firm's profit equation is

$$\pi = \text{TR} - \text{TC} = pq(p, \text{R\&D}) - [C(q) + \text{R\&D}], \tag{17.1}$$

where TR is total revenue.

[17] That is, $\partial q/\partial \text{R\&D} > 0$ and $\partial^2 q/\partial \text{R\&D}^2 < 0$. This last condition assures that the firm's second-order condition of profit maximization is met. To simplify the analysis, we let R&D affect demand or costs directly. That is, its effect on technology (T) is assumed to be 1 (i.e., $\partial T/\partial \text{R\&D} = 1$).

To determine the profit-maximizing price and level of R&D, we apply marginal analysis by maximizing profit with respect to p and R&D. Optimal values are obtained by solving these first order conditions simultaneously for p and R&D. It turns out that the first-order conditions are identical to those in the advertising problem in Chap. 15 (Sect. 15.3.2), except that R&D is the choice variable instead of advertising.

Because the advertising and R&D models are so similar, we ignore the first-order condition for price (or assumed that price is fixed) and focus on the firm's R&D problem. The first-order condition with respect to R&D is[18]

$$\begin{aligned}\frac{\partial \pi}{\partial \text{R\&D}} &= \frac{\partial \text{TR}}{\partial \text{R\&D}} - \frac{\partial \text{TC}}{\partial \text{R\&D}} \\ &= \text{MR}_{\text{R\&D}} - \text{MC}_{\text{R\&D}} \\ &= \left(p\frac{\partial q}{\partial \text{R\&D}}\right) - \left(\frac{\partial C}{\partial q}\frac{\partial q}{\partial \text{R\&D}} + 1\right) = 0, \end{aligned} \tag{17.2}$$

where $\text{MR}_{\text{R\&D}}$ is the marginal revenue associated with R&D and $\text{MC}_{\text{R\&D}}$ is the marginal cost of R&D. Notice that the marginal principle applies: the optimum is reached where the marginal benefit equals the marginal cost of R&D.

Equation (17.2) has a similar interpretation as the first-order condition for advertising in (15.5). Like advertising, we can think of R&D as a costly input that is designed to increase demand. The marginal benefit, $p(\partial q/\partial \text{R\&D})$, can be thought of as the value of the marginal product of R&D. The marginal cost can be decomposed into two parts. The direct effect measures the added cost of increasing R&D expenditures by one dollar (the +1 with in parentheses). The indirect effect captures the influence that R&D has on the marginal cost of production $[(\partial C/\partial q)(\partial q/\partial \text{R\&D})]$.

Another way of thinking about (17.2) is to rearrange terms so that it identifies the firm's ratio of R&D to total sales (total revenue). This is called the R&D-to-sales ratio. By rearranging terms,[19]

$$\begin{aligned} 1 &= \frac{\partial q}{\partial \text{R\&D}}(p - \text{MC}), \\ \frac{\text{R\&D}}{p \cdot q} &= \frac{\text{R\&D}}{q}\frac{\partial q}{\partial \text{R\&D}}\frac{(p - \text{MC})}{p}, \\ \frac{\text{R\&D}}{p \cdot q} &= \frac{\eta_{R\&D}}{\eta}, \end{aligned} \tag{17.3}$$

[18] This derivative involves the use of the chain rule, which is discussed in the Mathematics and Econometrics Appendix at the end of the book. According to the chain rule, if $y = f(x_1)$ and $x_1 = f(x_2)$, then a change in x_2 causes a change in x_1 which causes y to change. That is, $\text{d}y/\text{d}x_2 = (\text{d}y/\text{d}x_1)(\text{d}x_1/\text{d}x_2)$. In this case, because $C = C(q)$ and $q = q(\text{R\&D})$, $\partial C/\partial \text{R\&D} = (\partial C/\partial q)(\partial q/\partial \text{R\&D})$.

[19] This derivation is based on the fact that the Lerner index is $\mathcal{L} \equiv (p - \text{MC})/p = 1/\eta$, where MC is the long-run marginal cost of production. Recall from Chaps. 6 and 12 that the Lerner index is derived from the firm's first-order condition of profit maximization.

where η is the absolute value of the price elasticity of demand and $\eta_{\text{R\&D}}$ is the R&D elasticity of demand, which measures the percentage change in quantity demanded that is caused by a 1% increase in R&D [i.e., $(\partial q/\partial \text{R\&D})(\text{R\&D}/q)$]. This is similar to the Dorfman–Steiner condition of advertising that we discussed in Chap. 15. It implies that:

- A firm's R&D-to-sales ratio will be higher when there is a greater likelihood that R&D will improve product quality and increase demand (i.e., $\eta_{\text{R\&D}}$ is greater). In other words, firms will invest more money in R&D when it is expected to produce greater benefits. In cases such as this, R&D is said to have greater **technological opportunity**. Such opportunities for progress will vary by industry.
- R&D spending will be greater the lower the price elasticity of demand.[20] Recall from the Lerner index that a lower η implies greater monopoly power. Thus, (17.3) implies that firms that have greater market power will invest more in R&D. Later in the chapter, we investigate this issue more thoroughly.

17.3.1.2 Technological Change That Lowers Production Costs

Next, we analyze a situation where an investment in R&D lowers the firm's marginal cost of production but has no direct effect on demand. In this case, the firm's total cost is $\text{TC} = c(\text{R\&D}) \cdot q - \text{R\&D}$. In this model, c is the marginal production cost, which falls at a decreasing rate with respect to R&D.[21] The firm's profit equation is

$$\pi = \text{TR} - \text{TC} = pq(p) - c(\text{R\&D})q - \text{R\&D}. \tag{17.4}$$

To determine the profit-maximizing price and level of R&D, we follow the same procedure as before. We maximize profit with respect to p and R&D, and optimal values are obtained by simultaneously solving these first-order conditions for p and R&D. As before, we ignore the price equation so we can focus on R&D. In this model, the first-order condition with respect to R&D is

$$\begin{aligned}\frac{\partial \pi}{\partial \text{R\&D}} &= \frac{\partial \text{TR}}{\partial \text{R\&D}} - \frac{\partial \text{TC}}{\partial \text{R\&D}} \\ &= \text{MR}_{\text{R\&D}} - \text{MC}_{\text{R\&D}} \\ &= (0) - \left(\frac{\partial c}{\partial \text{R\&D}} q + 1\right) = 0. \end{aligned} \tag{17.5}$$

[20] This is consistent with Spence (1975), who showed that the gains from improving product quality will be larger as the price elasticity of demand falls. However, Kamien and Schwartz (1970) find that the gains from reducing the cost of production are larger the more elastic is demand.

[21] That is, $\partial c/\partial \text{R\&D} < 0$ and $\partial^2 c/\partial \text{R\&D}^2 > 0$. This is required for the firm's second-order condition of profit maximization to hold.

In this case, the marginal benefit results from the reduction in marginal cost due to R&D, $\partial c/\partial \text{R\&D} < 0$.[22] The model implies that:

- An increase in the effectiveness of R&D to reduce costs (i.e., an increase in the absolute value of $\partial c/\partial \text{R\&D}$) raises the marginal benefit of R&D, which causes the firm to increase its R&D spending. When this occurs, there is greater technological opportunity in terms of cost savings.
- The firm will increase its R&D spending as it produces more output. This seems reasonable, as the benefits of R&D increase with greater sales.

Both this model and the model where R&D increases product quality imply that the firm will make greater investments in R&D when it leads to a greater increase in demand or a greater decrease in costs (i.e., there is greater technological opportunity). They also imply that firms are more likely to invest more in R&D as market power and sales increase.[23]

17.3.2 Competition Versus Monopoly

Arrow (1962) is the first to formally model the effect of market structure on innovative activity. He conducted a thought experiment where a single market has either one firm or many (competitive) firms.[24] There is no product differentiation, and firms compete by simultaneously choosing price (i.e., Bertrand). Demand and cost functions are linear. Technological change leads to a decrease in the marginal cost of production. The goal of the model is to determine whether a firm will have greater incentive to invest in innovative activity in a monopoly or a more competitive market setting.

First, we consider the case of monopoly. Figure 17.12 describes the firm's demand, marginal revenue, and marginal cost functions. Marginal cost equals MC before the technological change and is MC′ after the change. At MC, the firm's profit-maximizing price–output pair is p and Q_{m}. At MC′, it is p' and Q'_{m}. This is nondrastic or **minor technological change** because $p' > \text{MC}$. A drastic or **major technological change** occurs when MC falls by so much that $p' < \text{MC}$. At output level Q_{m}, the firm's total revenue can be defined as the area under the marginal revenue function,[25] and the firm's total production cost can be defined as

[22] The second-order condition is met, because $\partial^2 c/\partial \text{R\&D}^2 > 0$.

[23] For simplicity, we have ignored the price of conducting research and development. Investment in R&D would also be expected to increase as the price of R&D falls.

[24] Arrow actually compared monopoly with perfect competition (which is the same as the Bertrand outcome when the number of firms exceeds 1). In any case, Arrow's main results are unaffected by assuming any market structure (such as Cournot) that produces an equilibrium level of market output that exceeds the cartel level of market output.

[25] We discuss this fact in Chap. 2.

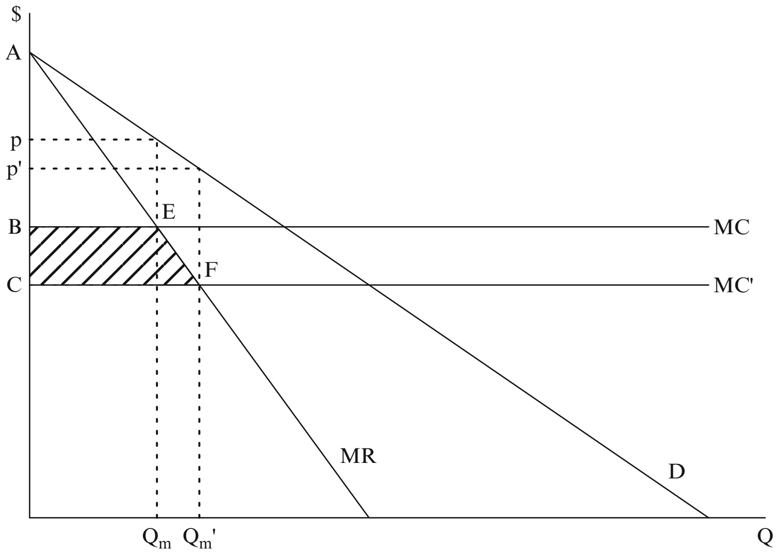

Fig. 17.12 A monopoly firm's incentive to innovate

the area under the marginal cost function as output ranges from 0 to Q_m. Note that monopoly output is considerably less than the socially optimal level of output, which occurs where MC crosses demand. When total revenue and total cost are defined in this way, the firm's total profit under the old technology is area ABE (i.e., total revenue minus total cost). For the new technology, its total gross profit is area ACF (ignoring the cost of R&D). Thus, the technological change causes gross profit to increase by area BCFE.[26] This implies that the firm will invest in R&D as long as it does not cost more than this amount, which we assume to be the case.

Next, assume that everything is the same except that there are now many firms. One of these firms invests in R&D that leads to the same cost-saving technological change as in the monopoly case. The Bertrand equilibrium for the old technology occurs where price equals marginal cost (MC) at market output level Q_c (the perfectly competitive outcome) described in Fig. 17.13. When a single firm discovers the new technology, its marginal cost falls to MC′, but its competitors' marginal cost remains at MC. In this case, the Nash equilibrium occurs where price equals MC − ε, where ε is small. Here, we assume that ε is 0 for simplicity, and the equilibrium output equals Q_c.[27] Profits for the innovative firm are 0 under the old technology and equal area BCHG for the new technology (ignoring R&D costs).

[26] If the new technology lasted for many periods, its benefits would equal the present value of the gain in future gross profits. This simply complicates the analysis without providing important new insights.

[27] Alternatively, we could assume that the firm owns the right to the new technology and licenses it out to existing firms for a royalty payment equal to MC − MC′. This produces the same result.

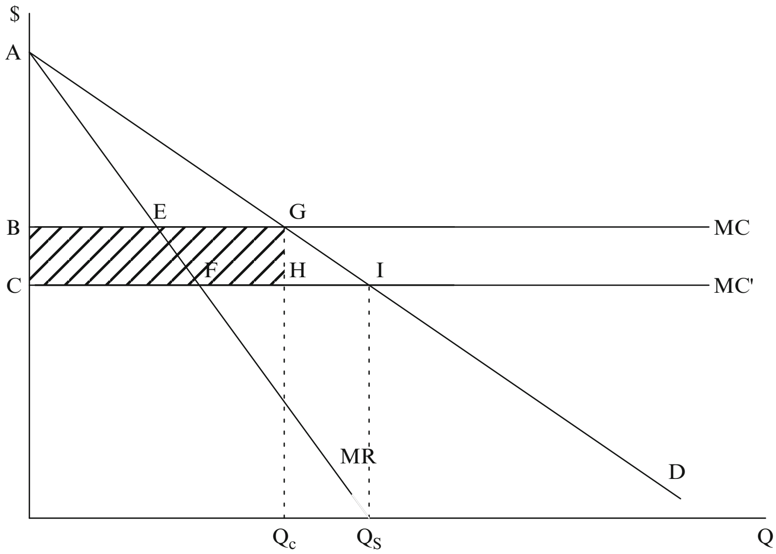

Fig. 17.13 A competitive firm's incentive to innovate

Thus, the new technology causes the innovative firm's gross profits to increase by area BCHG. Because the gain in profit in the competitive case (BCHG) exceeds the gain in profits in the monopoly case (BCFE) and BCFE is assumed to exceed the cost of R&D, the firm in the competitive market will make this investment in R&D.

Arrow's analysis has two important implications. First, because the gain in gross profit due to innovative activity is greater in the competitive market (BCHG) than in the monopoly market (BCFE), the incentive to innovate is greater with more competition. In other words, the monopoly firm is willing to invest up to area BCFE in R&D, while a firm in a more competitive setting is willing to invest up to area BCHG. This is just the opposite of Schumpeter's hypothesis.

The intuition behind Arrow's result is due to the fact that the monopolist restricts output. Thus, the increase in profit per unit (MC − MC′) is applied to a smaller quantity of output in monopoly (Q'_m) than in the competitive case (Q_c). Another issue is that for the monopolist, the new technology replaces the firm's old technology and assets, which is costly to the firm; for a more competitive firm, its new technology primarily replaces the old technology and assets of its competitors. This is a form of creative destruction that is known as the **replacement effect**. Because the replacement effect is greater for the monopolist, a monopolist will have less incentive to innovate.

The second implication of Arrow's analysis is that society values R&D more than the monopolist and the competitive innovator. To see this, recall that the social optimum occurs where price equals marginal cost in each setting (i.e., at Q_c with the old technology and Q_S with the new technology in Fig. 17.13). Thus, the gain in gross profit due to the new technology from society's perspective is area BCIG.

Because area BCIG > BCHG > BCFE, society values the innovation more than the competitive firm, but the competitive firm values it more than the monopolist. Thus, a competitive firm as well as a monopoly firm will underinvest in R&D from society's perspective.

There are at least two reasons to question the conclusion of Arrow's model that competitive firms have greater incentive to innovate. First, firms face a dynamic setting in reality, and over time new processes are more likely to spread quickly to firms within an industry than to firms in unrelated industries. Thus, competitor marginal costs will gradually fall below MC in Fig. 17.13, which reduces the firm's incentive to innovate in a competitive market. If imitation were instantaneous, for example, then MC will fall immediately to MC′, and there will be no incentive for an individual firm in a competitive market to innovate.[28] The point is that imitation discourages R&D. Second, competitive firms do not have excess profits to invest in R&D. If capital markets are imperfect, it may be easier for a monopolist to use internal funds to support R&D than for a competitive firm to raise the same investment dollars from outside sources. Thus, the cost of R&D may be higher for competitive firms.

17.3.3 Monopoly and a Potential Entrant

Another concern with the Arrow model is that it assumes that only one firm invests in R&D, whereas Schumpeter argued that firms actively compete in innovative activity. Gilbert and Newbery (1982) developed a model which is more consistent with Schumpeter because it allows firms to have a choice between investing and not investing in R&D. To focus on the main ideas, we assume just two firms: an incumbent monopolist (M) and a potential entrant (PE). They compete in developing a new process that will lower marginal cost. This innovation is then protected by a patent that cannot be circumvented. They compete in a three stage game. In the first stage, M decides whether or not to invest in R&D. In the second stage, PE decides to enter or not. In the final stage, PE decides whether or not to invest in R&D. A key feature of the model, which has an important effect on the outcome, is that there is no uncertainty regarding the effectiveness of R&D to generate a new innovation.

To make this a concrete example, assume that the inverse demand function is $p = 100 - Q$, where $Q = q_1 + q_2$. Marginal cost equals $40 before the innovation

[28] Even with patents, firms can sometimes circumvent them. Mansfield (1968) found that the time between the introduction of an innovation and when 60% of related products had imitated the innovation ranged from one month for simple production processes to several decades for more complex ones (e.g., steel production). As you might expect, Levin et al. (1987) found that it takes considerably longer to imitate a major new product that has been patented than one that has not been patented.

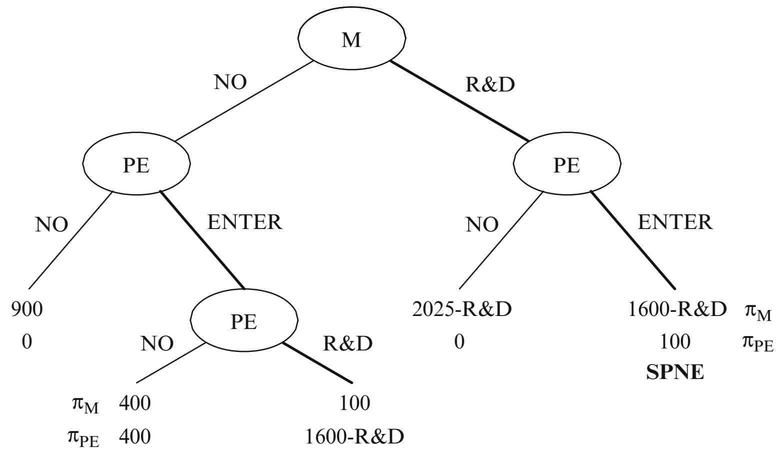

Fig. 17.14 Extensive form of an innovation game with potential entry

and \$10 after the innovation. The optimal expenditure on R&D is sufficiently low to ensure that it is a worthwhile investment. In this example, assume that \$500 < R&D < \$1,200. This makes it profitable for only one firm to invest in R&D. There are five possible outcomes, and profits for each are derived as follows:

- When M chooses not to invest in R&D and
  (1) PE chooses not to enter, M is a simple monopolist and earns a profit of \$900. PE's profit is 0.
  (2) PE chooses to enter and not invest in R&D, the outcome is that of a symmetric Cournot game as described in Chap. 10. Each firm earns a profit of \$400.
  (3) PE chooses to enter and invest in R&D, the outcome is that of a symmetric Cournot game as described in Chap. 10. M earns \$100 and PE earns \$1,600 − R&D.
- When M chooses to invest in R&D and
  (4) PE chooses not to enter, M is a simple monopolist and earns a profit of \$2,025 − R&D. PE earns 0 profit.
  (5) PE chooses to enter, the outcome is asymmetric Cournot as described in Chap. 10. M earns \$1,600 − R&D and PE earns \$100.

The extensive form of the game is described in Fig. 17.14.

Recall from Chaps. 3 and 11 that we use backwards induction to solve dynamic games. Notice that if M does not invest in R&D, PE's best reply is to enter and invest in R&D. At this outcome, M earns \$100. Alternatively, when M invests in R&D, PE's best reply is to enter. At this outcome, M earns \$1,600 − R&D. Thus, the subgame Nash equilibrium is for M to invest in R&D (because R&D < \$1,200) and for PE to enter. This outcome is more consistent with Schumpeter because it implies that a monopolist can have a strong incentive to innovate, especially when faced with a potential entrant. Given that R&D is less than \$1,200, the net returns to

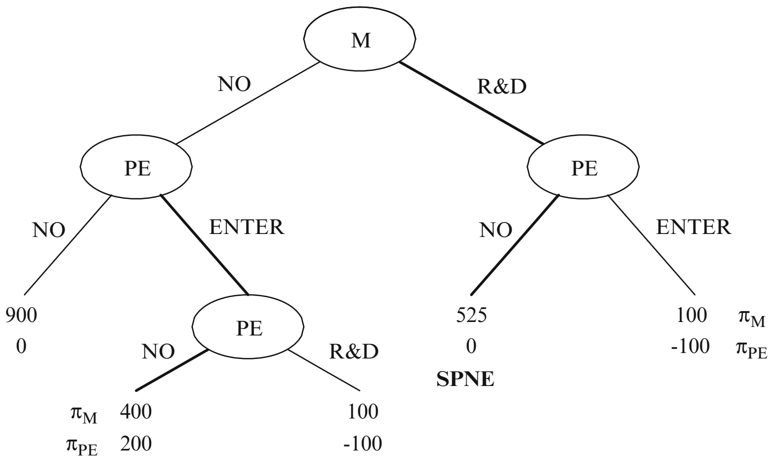

Fig. 17.15 Extensive form of an innovation game with entry costs

R&D for M are quite high (profits of 100 versus profits of more than 400 = 1,600 − R&D). In other words, M has an incentive to preempt PE by developing a new technology first.

Adding uncertainty regarding the effectiveness of R&D to produce a valuable innovation can change this result. Lee and Wilde (1980) and Reinganum (1983, 1984, 1985) showed that with uncertainty the potential entrant will enter the market and spend more on R&D than the incumbent. The reason for this is that the uncertainty of success induces firms into an R&D race. This speeds up innovation time, which more quickly lowers the value of the incumbent's old technology and reduces the incumbent's incentive to invest in R&D. A market with a large incumbent firm that is less likely to invest in R&D than a potential entrant is more in keeping with Arrow than with Schumpeter. Nevertheless, industry leaders are likely to be replaced by entrants in this model, an outcome that is consistent with Schumpeter's notion of creative destruction.

The Gilbert and Newbery (1982) model can be modified to show how R&D can serve as a strategic barrier to entry. To illustrate, assume that PE must pay an additional sunk cost (σ) to enter the market. In this example, $\sigma = 200$ and R&D = \$1,500. This game is described in Fig. 17.15. The monopolist's optimal strategy is to invest in R&D, which now eliminates entry: the SPNE is for M to invest in R&D and for PE to not enter. Thus, the monopolist has an even stronger strategic reason for investing in R&D under this scenario. The motivation for this result is that the innovation is more profitable to the monopolist because the firm uses it to preserve its monopoly position, while the potential entrant uses it to enter the market and become a duopolist. Without that threat, it would not be profitable for the monopolist to invest in R&D. Thus, R&D serves as a strategic barrier.

Patent races have similar dynamic features. By adding a time dimension to the patent problem, firms have a choice of when to patent a new product or process.

Faster development of a new technology will be more costly, but it may give a firm a strategic advantage over its competitors. In this setting, firms may be thrust into a **patent race** where each firm races to be the first to obtain a patent.[29] It can also induce an incumbent firm to start early in order to gain a head start in its R&D efforts. An early lead could encourage others to drop out of the race, unless the probability of being able to "leapfrog ahead" of the incumbent is sufficiently high. Races such as these can induce firms to innovate faster than they would prefer if they did not face competition. The rate of innovation may also be faster than is socially optimal. This is consistent with Schumpeter, because these types of races for new technologies are associated with concentrated industries.

17.3.4 Oligopoly and the Incentive to Innovate

Dasgupta and Stiglitz (1980) provided a simple oligopoly model that is consistent with Schumpeter's hypothesis that innovative activity will be greater in more highly concentrated industries. They considered an n firm oligopoly where firms simultaneously choose the level of output and R&D expenditures that lower marginal cost. Firms are symmetric, products are homogeneous, and production costs are linear. Entry is free, so that firm profits will be zero in the long run. In other words, this is a Cournot-type model with R&D and free entry.

Two equilibrium conditions provide a connection between industry concentration and R&D activity. From Chap. 12 (Sect. 12.1.1) we saw the firm's first-order condition of profit-maximization produces the following Lerner index ($\mathcal{L}$) in an oligopoly setting:

$$\mathcal{L} \equiv \frac{p - \text{MC}}{p} = \frac{1}{n \cdot \eta} = \frac{\text{HHI}}{\eta}. \tag{17.6}$$

From symmetry, the Herfindahl–Hirschman index (HHI) equals $1/n$, which is also the market share of each firm when the market is in equilibrium. Given free entry, we also know that long-run equilibrium profits for firm i will be zero. That is,

$$\pi_i = \text{TR} - \text{TC} = p(Q)q_i - [c(\text{R\&D})q_i + \text{R\&D}] = 0, \tag{17.7}$$

where Q is the industry level of output. Aggregating (17.7) over the equilibrium number of firms in the industry (n^*) gives

$$p(Q)Q - c(\text{R\&D})Q - n^* \cdot \text{R\&D} = 0, \tag{17.8}$$

where $n^*\cdot$ R&D is the expenditures on R&D by the industry

[29] See Kamien and Schwartz (1982) and Reinganum (1989) for a more complete discussion.

Solving for n^* · R&D and dividing both sides of the equality by $p(Q)Q$ gives

$$\frac{n^* \cdot \text{R\&D}}{p(Q)Q} = \frac{p(Q)Q - c(\text{R\&D})Q}{p(Q)Q} = \frac{p(Q) - c(\text{R\&D})}{p(Q)} = \mathcal{L}. \tag{17.9}$$

Note that the first term on the left is the ratio of R&D to total revenue at the industry level, the industry R&D-to-sales ratio. Given the definition of the Lerner index in (17.6), this ratio can be written as

$$\frac{n^* \cdot \text{R\&D}}{p(Q)Q} = \frac{\text{HHI}}{\eta}. \tag{17.10}$$

It says that the R&D-to-sales ratio at the industry level will be greater as the price elasticity of demand falls and as industry concentration (HHI) increases. This is consistent with Schumpeter, as the intensity of R&D spending increases with industry concentration.[30]

17.4 The Effect of Technological Change on Market Structure

Although market structure can affect the degree of technological change, a new technology can also have a dramatic effect on market structure. An entrepreneur who creates a new product may seize a market from existing firms and increase concentration. In addition, a technological change that affects MES will also lead to a change in concentration. For example, Demsetz' (1973) superior efficiency hypothesis indicates that if a superior firm discovers a lower cost technology that increases (decreases) MES, an increase (decrease) in average firm size and industry concentration will follow.

We have also seen that investments that increase sunk costs, such as expenditures on advertising and R&D, can serve as a strategic barrier to entry. As Sutton (1999) points out, however, this relationship can be rather complex for R&D. He argues that the link between R&D and concentration depends upon the nature of the technological change. For example, improved technological opportunities that are vertical in nature (i.e., greater opportunity to raise product quality or lower marginal cost) will lead to greater R&D. They will also cause industry concentration to rise, as a firm that produces a better product at lower cost will take market share away from competitors.[31] Only the best and lowest cost products survive in the long run. This implies that in the vertical case, R&D and concentration will be high (low) when technological opportunities are high (low).

[30] This assumes that η is constant, an assumption that may not hold in reality. For further review of the theoretical literature, see Dasgupta (1988) and van Cayseele (1998).

[31] Sutton argues that this is because the degree of substitutability between products is high in the vertical case.

This relationship breaks down when technological opportunities are horizontal in nature, such as when an innovation leads to a new product that is differentiated horizontally. In this case, improved technological opportunities increase R&D but may have a negligible effect on concentration. This is because the degree of substitutability is limited when differentiation is horizontal, and a new product will not generally replace existing products. For example, a firm that develops a new coconut flavored breakfast cereal need not cause other brands such as Cheerios to be squeezed out of the market. Thus, Sutton shows how a third cause, the type of technological opportunity, can influence both R&D spending and industry concentration.

17.5 The Empirical Evidence

Discussion in previous sections indicates that many forces influence R&D expenditures, patent activity, and technological change. These can be organized into three broad categories that are summarized below:

1. **Government incentives:** Given the public nature of information, government uses the patent system and research grants to encourage technological change through investments in R&D.
2. **Private firm incentives:** Economic theory predicts that private firms will be more likely to invest in R&D when they are better able to appropriate the benefits of their innovation. Firms will also be more likely to invest in R&D when there is greater technological opportunity, that is, when such an investment is more likely to successfully lead to a new product, a better quality product, or lower production costs.
3. **The role of market structure:** Economic theory provides no clear link between market structure and technological change. Arrow's model predicts that competitive firms are more likely to invest in R&D. In contrast, Schumpeter's theory predicts that innovative activity is more likely to come from large firms in highly concentrated industries. Finally, causality can flow in the other direction: a technological change may influence market structure.

As discussed in Sect. 17.2, the US government uses a number of methods to encourage R&D activity and create new technological knowledge. Private companies receive a 20% tax credit on their R&D expenditures. The federal government supplies 26% of grant funding in the USA. Finally, the patent system encourages inventive activity by giving an inventor exclusive ownership of an invention or innovation for 20 years (appropriability). In their survey of the evidence, Cohen and Levin (1989) concluded that these government policies have reduced the cost of innovation, especially in agricultural, aircraft, and electronics industries. Mansfield (1968) found that patent protection contributed to technological change in the petroleum, machinery, and metal products industries and was especially important in the pharmaceutical and chemical industries. He estimated that 65% of pharmaceutical inventions and 30% of chemical inventions would not have occurred without patent protection. Finally, Blumenthal et al. (1986) found that university research contributes to technological progress.

It is reasonable to assume that R&D spending follows technological opportunity. After all, a profit-maximizing firm is more likely to make a risky R&D investment when it has a higher probability of success. To test this hypothesis, we need an accurate measure of technological opportunity. Because no such measure exists, most studies use industry dummy variables to control for industry-specific differences in technological opportunity.[32] Consistent with the hypothesis that greater technological opportunity increases R&D spending, these studies find greater R&D spending and patent activity in "high tech" industries that are associated with the scientific or technical fields.[33] These include chemical, computer, and electronics industries, as described in Table 17.2.

Most of the empirical literature on the subject of technological change has focused on Schumpeter's theory that innovative activity increases with firm size and market power or industry concentration. The literature is too vast to summarize here, but those who have surveyed it conclude that the early studies produce evidence that is rather weak and somewhat fragile.[34] This evidence shows that the intensity of R&D spending increases with firm size but only for a limited number of industries, such as chemicals, automobiles, and steel. In addition, many studies find that R&D intensity increases with industry concentration. Compared to other variables such as technological opportunities, however, the overall influence of firm size and concentration on R&D spending is quite small.

It may be too soon to conclude that market structure has little effect on technological change though, as these weak findings may be due to problems associated with estimating a model of technological change. First, there is a measurement problem. Because it is impossible to measure inventive output, most studies use either the intensity of R&D spending or patent counts as a proxy variable. But not all R&D projects are successful, and not all patents are of equal value.[35] Nevertheless, even if inventive output could be accurately measured, it is still difficult to control for the many other factors that influence technological change.

Another concern is that the inventive process may be substantially more complex than Schumpeter's work suggests, making it difficult to accurately capture it in an

---

[32] See, for example, MacLaurin (1954), Scherer (1965), Pakes and Schankerman (1984), Jaffe (1986), Cohen and Levin (1989), Geroski (1990), and Blundell et al. (1999).

[33] An explanation for this is provided by Nelson (1982a), who argues that advances in scientific knowledge increase technological opportunities by lowering the cost of applied research in scientific and technical fields.

[34] For a review of the literature, see Kamien and Schwartz (1982), Cohen and Levin (1989), Geroski (1990), van Cayseele (1988), and Blundell et al. (1999).

[35] The one exception is Gayle (2005), who attempted to control for the relative importance of a patent by using a citation-weighted patent count to measure innovative output. With this measure, Gayle found a stronger positive relationship between patent counts and industry concentration.

empirical model. For example, Henderson (1993) found that large and small firms contribute in different ways to technological progress in the photolithographic alignment equipment industry. She found that larger established firms are more likely to invest in incremental innovation, while small entrants are more likely to invest in radically new inventions. Looking at a series of case studies over time, Jewkes et al. (1969) analyzed the interaction of firms when a small firm comes up with a new invention. They found that in many cases a large firm acquires the smaller inventive firm, as larger firms are better at innovation and at bringing an invention to market. Contrary to both Schumpeter and Arrow, this implies that both small and large firms have an important role to play regarding technological progress.

To further complicate matters, causality between technological change and market structure runs in both directions. As indicated above, the creative destruction of technological change can have a dramatic effect on market structure. However, Sutton (1999) pointed out that this relationship can be complex as well. In Sutton’s view, when a technological change is vertical in nature (i.e., it improves product quality or lowers production costs), it causes concentration to rise; when technological change is horizontal (i.e., it creates horizontal product differentiation), it has little effect on industry concentration.

Sutton (1999) and van Cayseele (1998) investigated a number of industries and found general support for Sutton’s theory. Industries with products that are more vertically differentiated, have products that are close substitutes, and have high R&D-to-sales ratios are generally highly concentrated. These include digital watches, aircraft engines, glass processing, and photographic film. In industries with horizontal characteristics, such as the market for liquid flow meters, concentration is low even though R&D spending is high.

An alternative way of evaluating the effect of competition on technological change is to investigate the behavior of firms that are involved in patent or R&D races. In a certain world, Gilbert and Newbery (1982) showed that incumbent firms are more likely to invest in R&D than potential entrants. With uncertainty, however, Lee and Wilde (1980) and Reinganum (1983, 1984, 1985) showed that challengers are more likely to invest in R&D. The only test of these competing hypotheses was conducted by Czarnitzki and Kraft (2004), who used data from German corporations. They found strong support for the uncertainty model in which competition contributes to innovation.

Ultimately, we are interested in knowing the extent to which our mixed political-economic system, which provides government support for inventive effort, produces a dynamically efficient outcome. Because it is very difficult to estimate the expected future benefits and costs of our current system and compare them with alternatives, it is not surprising that there are no empirical studies on this subject. The lone study that touches on this topic is by Hughes et al. (2002), who conducted a counterfactual study of the pharmaceutical industry. That is, they estimated the net present value of the benefit to consumers of eliminating all patent protection in this industry. By substantially reducing market power, consumers will be better off today. Consumers would be worst off in the future though, as this policy would reduce the flow of new drugs in the future. Hughes et al. found that for every dollar gain in consumer welfare today, future consumers would lose 3 dollars due to a reduction in innovative activity. Although we might question the accuracy of their estimate, the magnitude of lost future benefits is substantial. Thus, it appears that consumers benefit from patents applied to pharmaceuticals. Of course, this need not be true in other industries, because the amount of lost future benefits would vary by industry.

In summary, technological opportunity and government policies clearly encourage R&D spending and technological change. There is insufficient evidence to know whether or not our current political-economic system is dynamically efficient, however. There is evidence that technological change can dramatically affect market structure, but empirical studies fail to obtain clear results regarding the effect of market structure on technological change. A close inspection of the literature suggests that the process that drives technological change is industry specific. Given this fact and other problems associated with empirical work in this area, historical case studies of technological change may be a fruitful avenue for future research (as suggested by Cohen and Levin 1989).

17.6 Summary

1. **Technological change** occurs when we add to our knowledge about technology. This leads to the creation of new products, the production of better quality products (without an increase in cost), or the invention of a new process (i.e., production of a given output with fewer inputs). Technological change is important because it contributes to economic growth and improves our standard of living.[36]
2. The concept of static efficiency ignores the time dimension that is associated with technological change, where consumption is reduced today in order to invest in research and development (R&D) and create a better life tomorrow. **Dynamic efficiency** is reached when there is a socially optimal amount of technological change. An economy that is statically efficient need not be dynamically efficient. For example, it may be dynamically efficient to allow an inventor to have temporary market power (which is statically inefficient) to encourage inventive effort. The social benefits of a new invention can quickly outstrip the deadweight loss due to market power.
3. A technological change can be scale increasing, decreasing, or neutral. **Scale-neutral technological change** has no effect on minimum efficient scale (MES). **Scale-increasing (-decreasing) technological change** leads to an increase (decrease) in MES.

[36] Again, this ignores the possible negative consequences of technological change, an issue that we address in Chaps. 19 and 20.

4. A technological change can affect the cost-minimizing combination of inputs. For example, a **labor-saving technological change** occurs when the cost-minimizing amount of labor falls relative to that of other inputs.
5. There are two types of R&D programs: basic research and applied research. **Basic research** is theoretical or experimental and is designed to create general scientific knowledge. **Applied research** creates knowledge that has a specific practical purpose.
6. The process of technological change can be divided into three stages. **Invention** is the act of conceiving a technical possibility. **Innovation** occurs when an invention is made operational. **Diffusion** (or imitation) occurs when an innovation becomes widely used.
7. Given that ideas are public goods, free markets may produce too few new and useful ideas. This motivates government policies to encourage inventive activity and creative pursuits.
8. Patents, copyrights, trademarks, and research grants are used by the government to encourage inventive activity. A **patent** facilitates **appropriability** of the benefits to the inventor, as it gives an inventor monopoly ownership of an idea or method for a limited length of time. Patent life in the USA is 20 years. A copyright gives creators ownership of their artistic expressions and computer programs. A trademark gives a company protection of a symbol or brand name that is important to the company's identity and reputation. These are designed to give creators and inventors a (sometimes temporary) property right to their inventive and creative works.
9. There are social benefits and costs associated with patents. On the benefit side, they encourage R&D and technological progress. On the cost side, they create static inefficiency by giving an inventor monopoly power. The socially optimal patent life occurs where the social marginal benefits equal the social marginal costs of lengthening patent life. If we had perfect foresight, it would be optimal to have a different patent life for each innovation, with a longer (shorter) life for more (less) valuable innovations. Given uncertainty, a practical solution is to set a single length for all patents.
10. A company can appropriate the benefits of its new products or process by obtaining a patent or by keeping the details of its innovation out of the hands of its competitors (called a **trade secret**). Coca-Cola has kept its formula for Coke a trade secret for over a century.
11. Schumpeter argued that market structure and technological change are closely linked. According to Schumpeter:

    - Capitalist markets are dynamic, as a process of **creative destruction** by which firms create new technologies to replace old technologies (i.e., methods, companies, and markets) is constantly at work.
    - Large firms in concentrated markets are necessary for dynamic efficiency, as they are more likely to invest in R&D that generates technological change. They may benefit from economies of scale in R&D and may have greater means (due to market power) and greater motive to invest in R&D in order to create or preserve market power.

12. A monopolist engages in a **minor (major) technological change** that lowers marginal cost when the equilibrium price with the new technology is above (below) the marginal cost with the old technology.
13. A monopolist is more likely to invest in R&D when:
    - There are greater **technological opportunities**. This occurs when such an investment is more likely to produce greater benefits (i.e., it is better able to increase demand or to lower costs).
    - Demand is more inelastic.
    - The firm produces a greater amount of output.
14. Economic theory provides conflicting predictions regarding the effect of market structure on technological change. Important predictions are listed below:
    - Schumpeter predicts that large firms in concentrated industries are more likely to invent than firms in more competitive markets. This prediction is supported by the Dasgupta and Stiglitz (1980) model.
    - Arrow predicts that the incentive to invent is greater with more competition. This is attributable to the **replacement effect**, which means that an innovation is less valuable to a monopolist because it has more to lose or less to gain from an innovation that replaces its current technology. However, both monopoly and competitive markets underinvest in inventive effort from society's perspective.
    - When a monopoly firm competes with a potential entrant for a new patent, the monopolist will have a stronger incentive to obtain the patent.
    - When firms compete in a **patent race**, where each firm races to be first to obtain a patent, firms may innovate faster than if they were not competitors and faster than is socially optimal.
15. Technological change can also influence market structure. When a change in technology increases (decreases) minimum efficient scale, industry concentration will increase (decrease).
16. Sutton (1999) points out that the relationship between technological change and market structure depends on the type of innovation. According to Sutton:
    - When a technological change is vertical in nature (i.e., it improves product quality or lowers marginal cost), it will cause industry concentration to increase.
    - When technological change is horizontal in nature (i.e., it changes horizontal differentiation or adds a new product that is not a close substitute with competing brands), it will have little if any effect on industry concentration.
17. There are a considerable number of empirical studies on the economics of technological change. A summary of the main results is provided below:
    - The evidence shows that government programs, such as patents and research grants, encourage inventive activity.

- R&D activity follows technological opportunity. Although it is impossible to precisely measure technological opportunity, the evidence shows that the intensity of R&D spending is greater in high tech industries that are associated with scientific and technological fields.
- There is weak and sometimes conflicting support for Schumpeter's view that large firms in concentrated industries are needed to generate technological change. This may be due to data limitations and methodological weaknesses of previous studies, such as the use of R&D spending or patent counts that may be poor proxies for inventive output. In addition, case studies show that the inventive process may be more complex than Schumpeter envisioned. In many cases the inventive process varies by industry, and in some cases small firms invent, while large firms innovate.
- There is also evidence that technological change influences market structure. Although available evidence is limited, there is support for Sutton's (1999) view that the effect of R&D on concentration depends on the type of innovation.
- Given the difficulty of accurately estimating expected future costs and benefits of government and market incentives to promote progress, there is insufficient evidence to know whether or not our current political-economic system is dynamically efficient.

17.7 Review Questions

1. Explain how technological change affects our standard of living. What would the world be like if technological change were to cease? Are there any negative consequences associated with technological change?
2. Compare and contrast the concepts of static and dynamic efficiency.
3. Consider a market with two periods (ignore discounting). Demand is $Q = 120 - p$ in both periods. In period 1, marginal cost (MC_1) is 70, and the market is perfectly competitive. One firm invests in R&D which pays off by lowering its marginal cost to $MC_2 = 20$ in the second period. The marginal cost of all other firms remains at 70.

   A. For each period, determine the equilibrium price (p^*), equilibrium market output (Q^*), consumer surplus (CS), and gross producer surplus (PS).
   B. Is the market statically efficient in each period?
   C. Under what condition(s) will this innovation be dynamically efficient?

4. Explain how the market for ideas may fail without some government involvement.
5. Explain how patents, copyrights, and trademarks may promote technological progress. Why not give creative individuals and firms unlimited ownership of the outcomes of their work?

6. In the USA, corporations are given 20% tax credit for their R&D expenditures. How will this encourage R&D?
7. If over the next decade a firm expects the price of labor to rise substantially compared to the price of capital, would the firm be more likely to invest in R&D that leads to expected labor-saving, capital-saving, or input-neutral technological change? Explain.
8. The Coca-Cola Company has kept its recipe for Coke a trade secret for over a century. Given that a patent lasts for only 20 years, why is it that all firms do not follow Coke and simply keep their new products and processes a secret?
9. Explain what is meant by technological opportunity as it applies to a firm's demand function and its cost function.
10. Schumpeter theorizes that there will be greater inventive activity in concentrated markets, while Arrow argues that there will be greater inventive activity in competitive markets.

    A. Under what set of theoretical conditions is Schumpeter more likely to be correct, and under what set of conditions will Arrow more likely be correct?
    B. Given the summary of the empirical literature, will inventive activity be greater with more or less competition, *ceteris paribus*.

11. In Figs. 17.12 and 17.13 we used a minor technological change to prove Arrow's proposition that competitive firms have greater incentive to innovate than monopoly firms. Prove or disprove Arrow's proposition for a major technological change.
12. Assume a market where the inverse demand is $p = 120 - Q$, p is price, and Q is industry output. Firm i's total cost is $\text{TC}_i = cq_i$, where q_i is firm i's output level. A single firm can invest in R&D that leads to a patentable innovation that lowers marginal cost, c. With the old technology $c = 40$, and with the new technology $c = 20$.

    A. Calculate the maximum amount that firm i is willing to invest in R&D to produce this innovation if the innovative firm competes in a

        (i) Monopoly market.
        (ii) Bertrand duopoly.
        (iii) Cournot duopoly.

        Assuming that innovation is profitable, in which market structure will firm i have greater incentive to innovate?

    B. Calculate the maximum amount that society would be willing to pay for this innovation.

13. Explain how the act of creating new technologies and obtaining patents can serve as a strategic barrier to entry.
14. Technological change through creative destruction can have a dramatic effect on market structure.

   A. Explain how technological change can either raise or lower industry concentration.
   B. How will a technological change that increases vertical versus horizontal product differentiation affect industry concentration?

15. What does the empirical evidence suggest regarding the effect of firm size and industry concentration on technological change? What does your answer imply about the dynamically efficient market structure (i.e., perfect competition, oligopoly with a relatively low or high level of concentration, and monopoly)?
16. Will an overconfident management team invest too much or too little in R&D from the firm's perspective? What about from society's perspective?