

Chapter 3

PATENTS AND INTELLECTUAL PROPERTY:

A GENERAL FRAMEWORK

Chapter contents:

- 3.1 Chapter outline
- 3.2 Inventions, innovations and diffusion
 - 3.2.1 Innovations in general
 - 3.2.2 Innovation models
 - 3.2.3 Diffusion processes for innovations
- 3.3 Patents
 - 3.3.1 Introduction
 - 3.3.2 What is a patent?
 - 3.3.3 How to value a patent
 - 3.3.4 What is patentable?
 - 3.3.5 Who can patent?
 - 3.3.6 Why apply for a patent?
 - 3.3.7 When to apply for a patent
 - 3.3.8 Where to apply for a patent
 - 3.3.9 How to apply for a patent
 - 3.3.10 Pricing of licenses
- 3.4 The patent system as an economic incentive system
 - 3.4.1 What is a patent system?
 - 3.4.2 Why have a patent system at all?
- 3.5 Previous research and literature
- 3.6 Economic theory of IP and patents
- 3.7 Summary and conclusions

3.1 Chapter outline

Inventions, know-how, data software, designs, trademarks and artistic works are all potential intellectual properties and may be protected by Intellectual Property Rights (IPRs). Such rights are by and large granted in contemporary society as a stimulus to creative work and innovations. IPRs pertain to specific legal systems embedded in the respective social system that provides legislation, law adherence, policing, prosecution, court practices, and infringement sanctions or penalties. These elements comprise the institutional framework of the IPR system. The framework has evolved during centuries as described in Chapter 2 and varies across countries in the same way that social and legal systems differ. It is thus important to realize that what is perceived as legally or morally right and wrong is contingent upon cultural and historical influences.

This chapter aims to give a brief textbook introduction to intellectual property in general and patents in particular. After some common concepts and distinctions are described, the context of inventions, innovations and diffusion is presented as a general framework for patents and IP. The basic technicalities of patenting are then described, as well as the basic rationales behind a patent system as an economic incentive system. After previous research and literature on IP and patents is surveyed, the chapter finally gives a brief introduction to the economic theory of IP and patents.

3.2 Inventions, innovations and diffusion

3.2.1 Innovations in general

Common distinctions

Innovations imply novel and accepted changes in society. As such, innovations are fundamental not only to technological and economic development but to cultural development at large (see e.g. Barnett 1953). Table 3.1 gives an overview of the different types of innovations in general. Innovations are also often characterized as major (or radical) and

minor (or incremental) related to the "size" of the change the novelty implies. Technological innovations are moreover commonly subdivided into product and process (production) innovations.

A number of additional distinctions are commonly used when discussing innovations and are therefore useful to review first. Such important distinctions are:

- inventions and innovations
- inventions and discoveries
- technical/commercial/economic uncertainty/success/failure
- innovations and diffusion
- innovations and imitations

Although commonly used, these concepts and distinctions are not so clear-cut under closer scrutiny.

Innovation

The concept of innovation has numerous definitions, but is typically defined by a reference to a change in ideas, practices or objects involving some degree of (a) novelty or creation based on human ingenuity and (b) success in application. The concept is used to refer to the new idea, practice or object as well as the process leading to it, although the term 'innovation process' then is more correct. Novelty (or newness) may refer to what is novel on a global scale, to a nation, to an organization or to an individual. Typically novelty to the world is implied when talking about invention and innovation. If there is an underlying patentable invention, novelty to the world is presumed.

Technical, commercial and economic success

The concept of success is often subdivided into technical success, commercial success and economic success. In this context technical success means that technical specifications have been met and/or an invention was achieved, commercial success means that the invention has found a first commercial application, and economic success means that an acceptable return on the total investment (RoI) has been achieved.¹ In a corresponding way one may distinguish between three types of failure as well as between three types of uncertainty: technical, commercial and economic.

Inventions

In broad terms an invention is a novelty or creation based on human ingenuity, but the concept of invention does not require that success in application has been achieved. Thus, an invention becomes an innovation when it becomes commercially successful. Inventions of all sorts are generated out of a mixture of human curiosity, the search for betterment and additional factors, more or less random. Table 3.1 again gives an overview of the different types of innovations with examples of some underlying inventions. In order to stimulate the flow of useful inventions and innovations, society may provide incentives, e.g. in the form of intellectual property rights attached to the inventions. Table 3.1 also gives some IP related characteristics of the different types of innovation.

Discoveries

Inventions, being man-made novelties, are distinguished from mere discoveries of pre-existing features of nature, like an ore deposit or a natural law. Technical inventions are

¹ In other words, economic success (or break-even) occurs when the total revenues exceed the total expenditures, compounded with the relevant rate of interest.

patentable but discoveries are not. However, the distinction between inventions and discoveries is not always as easy to uphold as it may appear (see Chapter 2).

Diffusion

The concept of diffusion refers to the process through which an innovation is adopted by individuals and organizations in a population (e.g. potential buyers or users on a market, sellers or producers in an industry, departments in an organization, nations in the world etc.) Typically an innovation changes during diffusion. For a classic overview of diffusion literature, see Rogers (1995).

Imitation

Finally, the concept of imitation refers to a close reproduction or near duplication of ideas, practices or objects that were once perceived as inventions (or innovations). Typically, imitations are not exact copies of the underlying invention (or innovation).

3.2.2 Innovation models

Two types of innovation models will be considered here. The first type shows the cash flow pattern associated with an innovation, while the second shows the activity pattern. In the first category one finds the so-called product life cycle model, which is a simple and useful point of departure here.

Table 3.1 General types of innovations¹⁾

Type	IP related characteristics ³⁾	Examples
Technological/technical ²⁾ innovation	Patentable Also protectable by trade secrets, copyrights, designs and trademarks	Wheel Telephone Incandescent lamp Transistor
Service innovations (e-g- in information services, financial services, telecom services, medical services, transportation services, educational services etc.)	Non-patentable in general Supporting technologies may be patentable Protectable by trade secrets, trademarks and copyrights	Newspapers Messaging systems Heart transplantation New school courses
Financial innovation	Non-patentable Easy to imitate but diffusion may still be slow	Insurance Convertible debentures Certificates of deposit Put and call options
Managerial/organizational innovation	Non-patentable Often slow diffusion May be protected as trade secrets	Divisionalized organization (M-form) Program evaluation and review technique for project planning (PERT) Kanban, JIT, TQM, etc.
Marketing/distribution innovation	Non-patentable Supporting technologies may be patentable	Supermarket Mail order Teleshopping
Institutional innovation	Non-patentable Slow diffusion	Patent system Limited liability joint stock company
Other types (social, legal, political, cultural, etc.)	Cultural innovations are protectable by IPRs	European community Opera

Note:

1) The types are not mutually exclusive.

2) The concept of technology is usually taken to mean knowledge about techniques. Strictly speaking, technological innovation then refers to change in knowledge, while technical innovation refers to change in artefacts, embodying technological change. Technical innovations are generally subdivided into product/process (production) innovations.

3) The applicability of different IPRs is just illustrative. The degree of protection may also be weak in many cases.

Product life cycle model

In essence, the product life cycle model (the PLC model) is based on the idea that a product has a finite lifetime on the market and therefore proceeds through various stages of development, including introduction, growth, maturity, saturation and decline.² It should be noted that this is an idealized model, and not all products conform to it. Further, there are several sets of stages that apply to different products, and sometimes there are not really any clear stages at all. Regardless of whether there are any clearly recognizable stages during the lifetime of a product, it is useful to picture the cash flows of a new product in a money/time diagram. Figure 3.1 gives an idealized picture of cash flow for a hypothetical product, divided into stages.

In relation to Figure 3.1, the following should be noted:

1. Cash flows are depicted with smoothed, continuous curves. In reality, monetary transactions occur at discrete points in time and curves connecting the values are typically not smooth.
2. The sales curve is supposed to be unimodal (that is, rising to one peak and then falling), while in reality there may be a slump and a revival.
3. Outflow of capital is counted as negative only for investment expenditures, not for direct costs and indirect costs (e.g. organization costs). This is slightly inconsistent but in line with conventions to display the sales margin (or operating profit margin) more clearly as the difference between gross value of sales and the sum of direct and indirect costs.

² For an overview of PLC modelling, see Porter (1980). There is also a so-called international product life cycle model, originated by Vernon (1966), which refers to different stages in terms of domestic or foreign location of production and markets. Briefly stated, the stages following upon each other are 1) domestic production for home market; 2) domestic production for foreign markets; 3) foreign production for foreign markets; 4) foreign production for home market. This model has later been criticized by various authors and later revised by Vernon, see e.g. Cantwell (1989).

4. Gross, annual cash flows are shown. There are other variants when plotting cash-flow curves, e.g., letting the money axis in the money/time diagram denote cumulative cash flow or denote net cash flow.

Although there is a phase of more or less concurrent investment in R&D, production and marketing preceding market introduction, investment expenditures typically continue after market introduction. For example, product variants are developed, design features are adapted to mass production, production is scaled up and further automated, and further investments in international marketing are made.

A company usually develops new products continually. There are new variants, types, models, generations and families of new products being developed, and different products of a company are often related to each other in some way. It is thus too narrow to focus on a specific product and try to illustrate its life cycle. In fact, if the cash-flow diagrams of all of a company's products are added, the resulting picture would look something like Figure 3.2. The cash-flow curves for different products may be interdependent since they may share scarce resources. There may be synergies among the products as well, e.g. they may to some extent be based on similar technologies and patents or use similar production or marketing strategies, e.g. using the same company trademark. Two products in the same company may also compete to some extent with each other on the market, i.e. they "cannibalize" upon each other's sales, e.g. IBM's personal computers competing to some extent with their own mainframe computers business.

Figure 3.1 Life cycle of a hypothetical product with annual cash flows

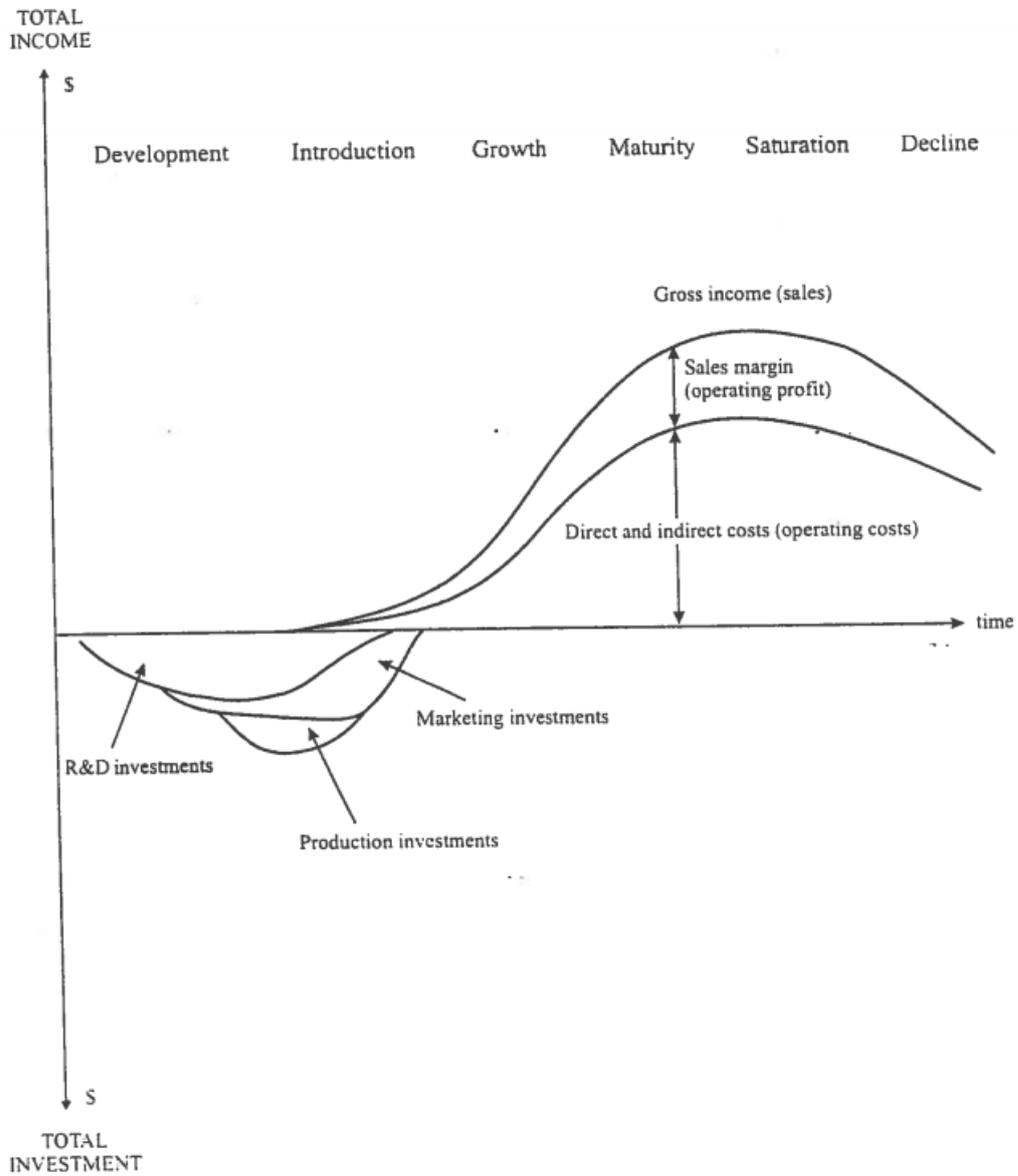


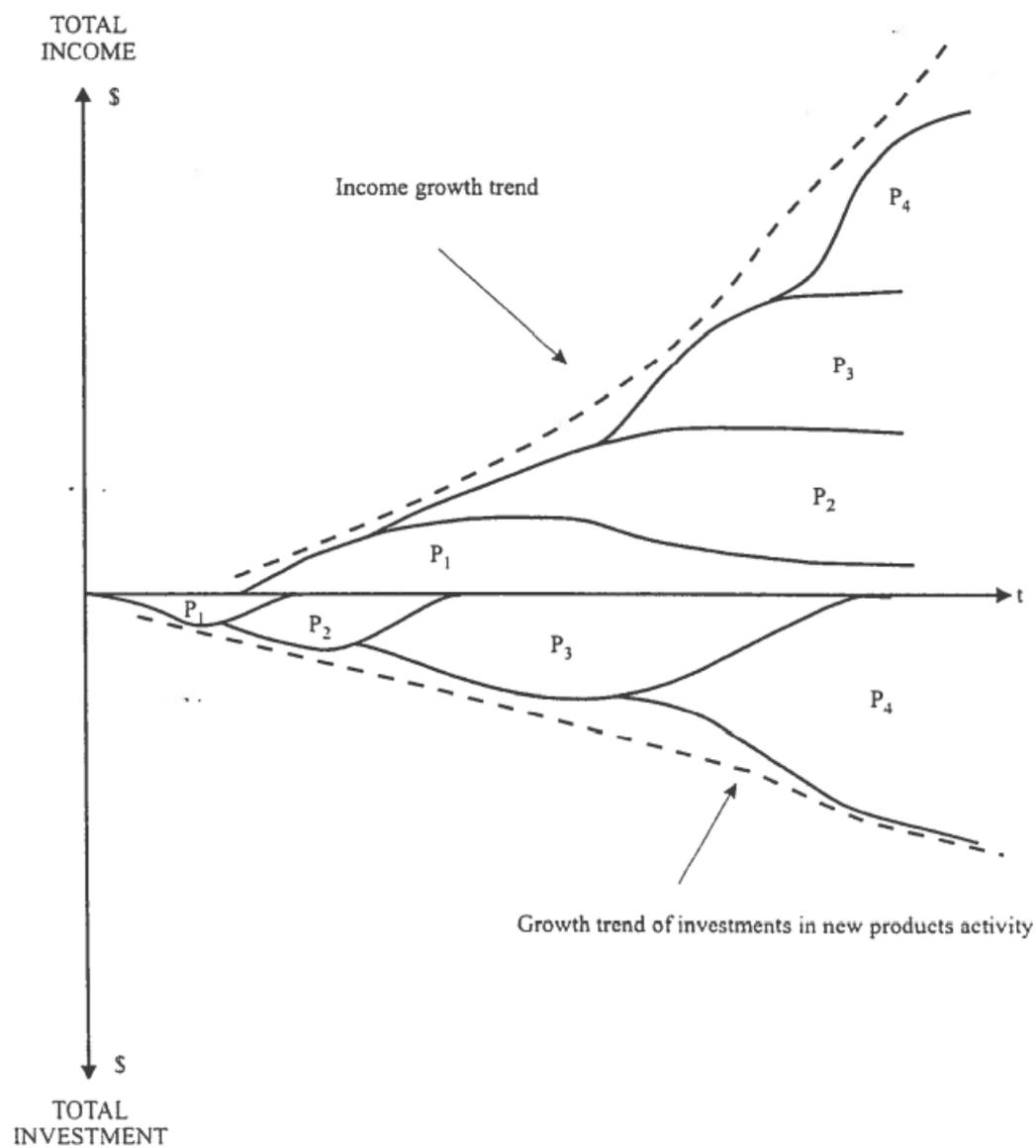
Figure 3.1 Life cycle of a hypothetical product with annual cash flows

If cash-flow curves refer to different models or generations of essentially the same product type in terms of functional characteristics, a similar type of figure as Figure 3.2 is obtained. see Figure 3.3.a. We can then observe how, e.g., R&D may be performed for different generations concurrently, just as different product generations may be marketed concurrently. In this case it is quite natural that a new generation cannibalizes on an older one, eventually outcompeting or substituting it.

Thus far cash-flow curves have been shown for one company with one or more products and product generations. We can also show cash flow curves for several competing companies with one or more competing product generations. Typically this is done for an innovator or leading company and a competitor (a follower or imitator) as illustrated in Figure 3.3.b. In the case of several product generations, the roles as leader and follower may switch back and forth between the companies, with one company taking the lead in one generation and another company taking the lead in the next product generation. This pattern of intermittent leadership (in the sense of being first to market) is quite common in R&D intensive oligopolies with a small number of competing companies.

In summary, the product life cycle model has many simplifying assumptions and limitations, and its division into stages is not clear. On the other hand, it is useful as a conceptual model, and cash-flow diagrams in general are highly useful as illustrations, e.g. for illustrating the timing of different product generations as in Figure 3.3.

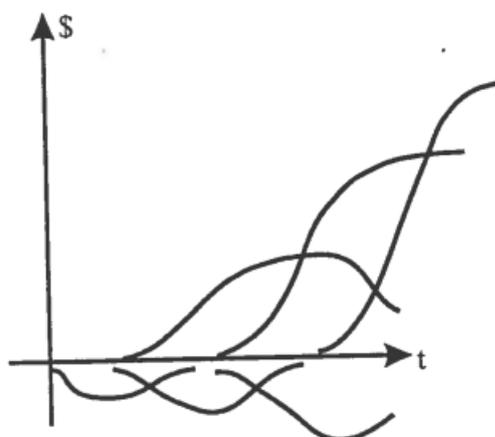
Figure 3.2 Corporate income growth through new products



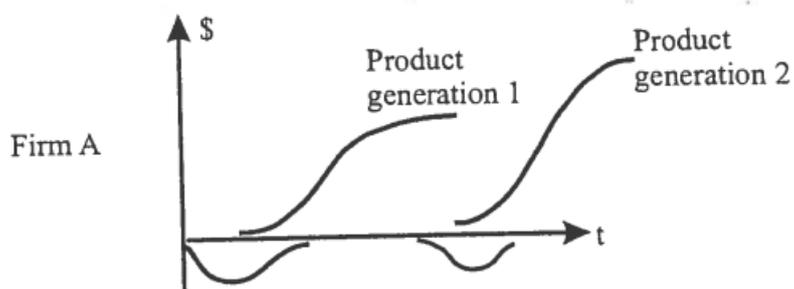
Legend: P_1, P_2, \dots denote different products with which income and investment are associated.

Figure 3.2 Corporate income growth through new products

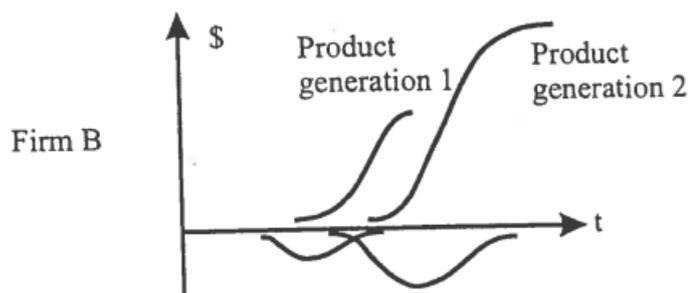
Figure 3.3 Examples of cash-flow curves for new products



a) Three product generations with overlapping R&D stages



Firm A



Firm B

b) Two firms with intermittent leadership over two product generations

Figure 3.3 Examples of cash-flow curves for new products

Activity models

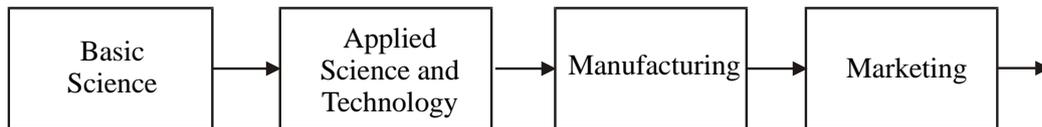
Up to now we have described cash flow curves for the whole life of an innovation, typically a new product. We will now narrow the focus and describe some activity models of the innovation process. There are several models for this, each representing different descriptive and/or normative views held by different people on how the innovation process normally looks or should look like. Figure 3.4 gives two extreme views of the innovation process as linear and one view of the innovation process as a highly interactive and iterative one.³

³ The so-called "linear model of innovation" usually refers only to the case, where innovations originate in S&T and then are developed, produced and marketed sequentially. This model, the origins of which may have been influenced by scientists seeking government support after World War II (see Bush 1945), has been strongly criticized; see e.g. Schon (1965), de Solla Price (1973), Kline and Rosenberg (1986) with an interactive type of model called the "chain-linked model", and Rosenberg (1994). Another view, articulated early on by Schmookler (1966) is that successful innovations start from recognition of a market need (or demand pull rather than technology push). See further Phillips (1971), Mowery and Rosenberg (1979), Freeman et al. (1982), Rosenberg (1994) for good reviews. As seen from Figure 3.4 the two linear views – technology push vs demand pull – pertain to what type of initial activity or institution constitutes the (major) source of innovations, while the distinction between linear and interactive (non-linear) view pertains to the order, if any, of subsequent activities in the innovation process. Authors like Schon, Rosenberg and Kline have also criticized the view that activities have to follow upon each other in a neat sequential manner in an innovation process. The practice of concurrent engineering emerging in the 1980s has also shown the possibilities to compress the innovation process through conducting activities more in parallel, thereby shortening the time to market (see further Chapter 6). For some further picturing of the innovation process, see Rosenbloom (1985) and Imai et al. (1985).

Figure 3.4a Two extreme models of the innovation process – the traditional views

(Source: Granstrand and Sigurdson 1981, p. 14)

(A) Science and Technology Push



(B) Need (Demand) Pull

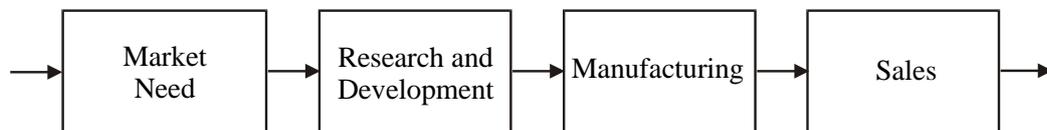
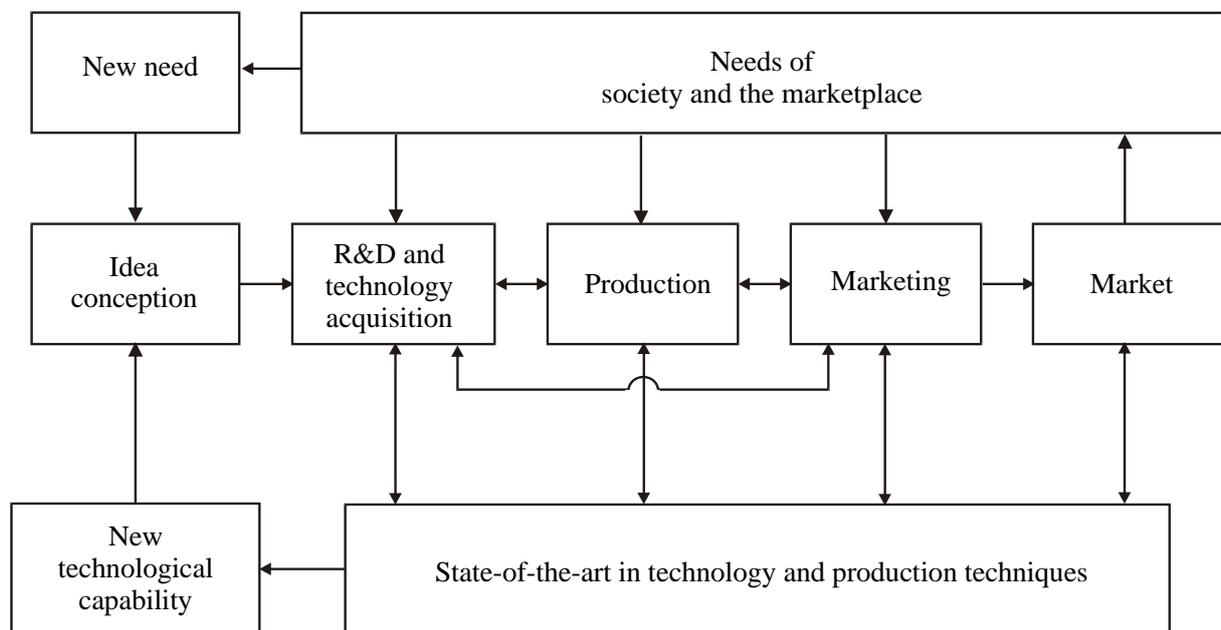


Figure 3.4b Interactive model of the innovation process



A common but far from clear cut way to structure the innovation process or the process of product development in general is to divide it into three sets of activities and events, which could be seen as constituting more or less overlapping phases (or stages):⁴

1. Idea phase
2. Research and development phase (predominantly product and process technology development, but also some commercial development)
3. Commercialization phase (typically consisting of the introduction phase in the PLC model, including manufacturing start-up)

Table 3.2 gives an overview of activities and events in these phases. Activities and sub-activities in innovation and product development typically do not proceed in neat, linear sequences but rather in "iterative spirals". The sequencing of activities also varies between products and companies due to the influence of e.g. company management and organizational behaviour as well as other factors, more or less random.

Key decision problems in the R&D and technology acquisition phase are, for instance: what to include in the technical specification; what to "make or buy"; whether, when and how to patent; how to make a prototype; and on which customers or users to perform product tests (so as to prevent information leakage, for example).

Key events, often unexpected, in this phase include becoming aware of the fact that a similar patent has been granted to a competitor; discovering that unwanted side-effects exist; or finding that some item cannot meet technical specifications. There are positive surprises as well. One might accidentally discover a new phenomenon, mechanism, compound or the like.⁵

⁴ Product innovation generally refers to a product that is new to the market, and therefore also new to the company, while product development in general refers to products that are new to the company but may not necessarily be product innovations. The line between what is new and what is not is difficult to draw.

⁵ Such discoveries are also called serendipities, that is, discoveries made while searching for something else.

Table 3.2 Illustration of key phases, activities and events in the innovation process

Idea phase	R&D and technology acquisition phase	Commercialization phase
Idea search, generation and evaluation	Technical specification	Manufacturing start-up (zero batch and so on)
Idea conception and invention	Involvement of lead users, lead suppliers and other partners	Test marketing
Preliminary business analysis	Market/user analysis	Recruitment and training of sales personnel
Technology scanning and planning	Preliminary design and invention	Marketing start-up
	External technology acquisition	Market introduction
	In-house R&D	Build-up of an international market organization
	Experimentation	Introduction to foreign markets
	Design and redesign	Broadening of application areas
	Analysis of manufacturability	Start-up of mass production and mass marketing
	Testing	Shift in management
	Patent applications	
	Tooling and pilot plant design	
	Prototyping	
	Involvement of general suppliers	
	Market planning	

Note: An innovation process is often complex and irregular; general types of separate activities, events and phases are not easily discerned. The general timing of activities may differ substantially from case to case and unexpected events occur. This table is therefore aimed at illustrating a generic case. The order of activities in the table does not represent a strict order in time. Usually iterations take place among the activities, many of which may also proceed in parallel (sometimes referred to as concurrent engineering).

3.2.3 Diffusion processes for innovations

In contemporary technology-based businesses it is more or less necessary to be innovative in order to stay competitive. From an economic point of view it is not sufficient only to be innovative, however. What is important is to create economically successful new products, that is, new products which not only are introduced and sold on a market, but are sold in sufficient numbers with sufficient profit margins for the total investment to pay off. The standard way of expressing this criterion is that the net present value of the total investment should be positive in order to be undertaken, i.e.

$$NPV = \sum_{i=1}^L (p_i q_i - c_i) / (1+r_i)^i \text{ should be } > 0$$

where (the values are expected or observed, depending upon in which period the calculation is done)

NPV = Net present value of the investment

p_i = Product price in period (year) i

q_i = Number or quantity of products sold in period i

c_i = Total cost incurred in period i for the new product (including investment expenditures)

L = Lifetime (number of periods) of the product (alternatively length of planning horizon)

r_i = Discount rate for investments of this sort for the company, reflecting the company's cost of capital and rate of interest in period i

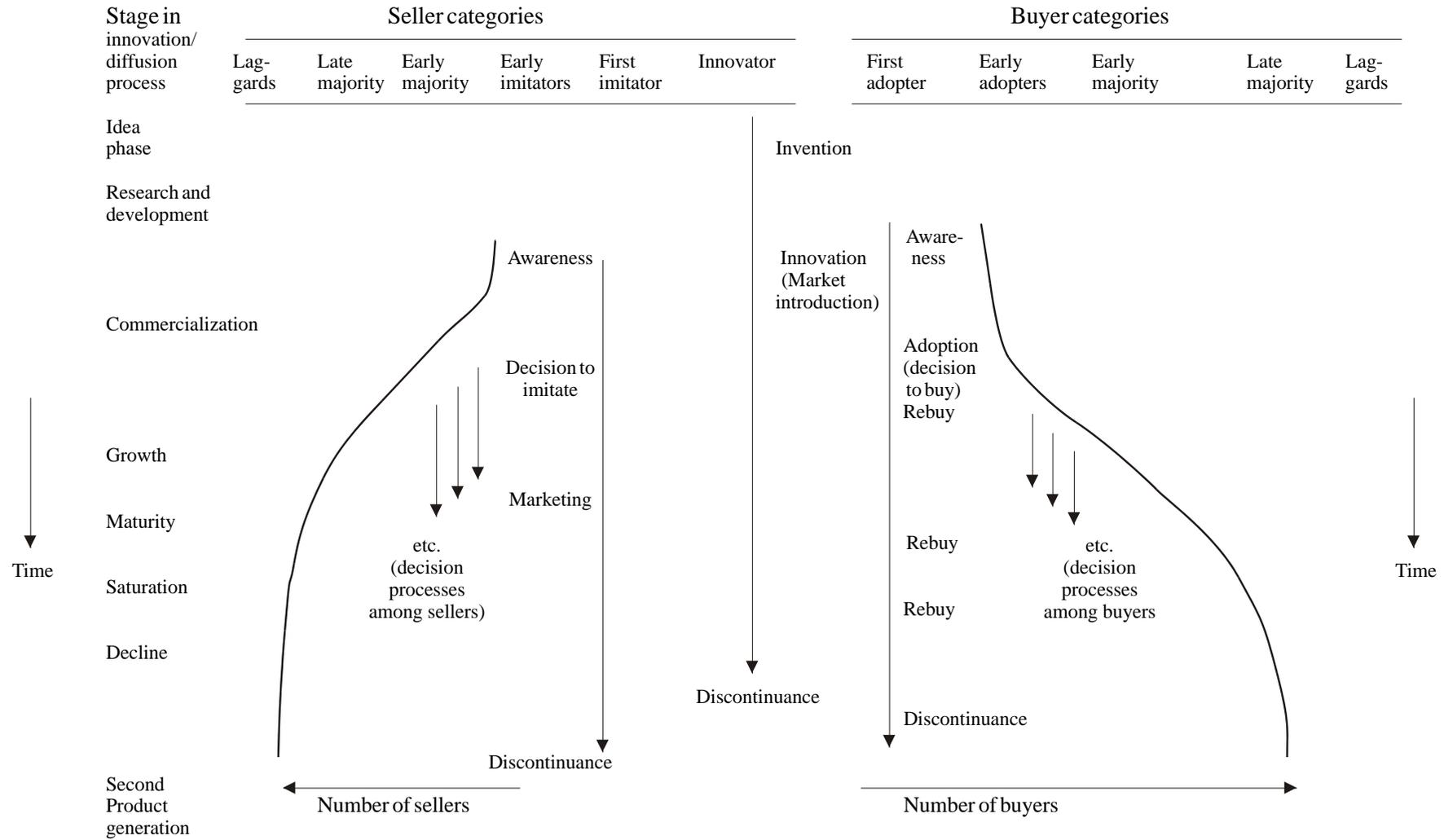
Economic calculations of this kind for judging a company's investment in the development, production and marketing of a new product will be further dealt with in Section 3.6. Here we will review qualitatively how price (p_i), sales volume (q_i), cost (c_i) and product lifetime (L) might be affected during diffusion.

The main objective of the commercialization of a new product is to build up the initial sales of the product in order to get a foothold on the market. However the main objective in the overall process is to ensure sufficient total sales at prices ensuring that the new product breaks even at some point in time, considering all investment expenditures and the relevant cost of capital. For the latter purpose, the market penetration or diffusion of the innovation is critical. It is then important to try to influence not only the diffusion or adoption of the new product among a number of buyers, but also to try to influence the diffusion or imitation of the new product and process technology among a number of sellers and producers who appear as competitors. These two different diffusion processes on the market will be called buyer diffusion and seller diffusion, respectively:

Buyer and seller diffusion

The process of buyer diffusion comprises a series of individual adoption processes, one for each buyer/user. Similarly, the process of seller diffusion comprises a series of individual imitation processes, one for each seller/producer. Depending upon when, in the buyer diffusion process, buyers adopt the innovation, the population of buyers (which in itself may be ill-defined or changing over time) may be decomposed into adopter categories, such as the standard grouping of adopters into early adopters, early majority, late majority and laggards. The first buyer to adopt a new product is simply called the first adopter. (A minority of authors prefer to call the first adopter the innovator.) Similarly the population of sellers/producers may be decomposed into innovators, early imitators, early majority etc. The innovator is the main actor behind the introduction of an innovation. Sometimes there will be several innovators appearing around the same time with similar innovations. In some cases the innovator may also be the first adopter, that is, the innovator is also the first one to use the innovation. Figure 3.5 gives an overview of the buyer and seller diffusion processes.

Figure 3.5 Overview of buyer and seller diffusion processes on the market for a new product.



What happens to the new product and technology during the innovation and diffusion processes, and how are price (p), quantity (q), cost (c) and product lifetime (L) affected? The product and its technology continue to change and develop during its diffusion. These changes and developments take place partly as a result of both buyer diffusion and seller diffusion. On the buyer side, adaptations to different users are made; new applications are found and new ideas come up, often from the users themselves (see von Hippel 1976, 1988). On the seller side, imitations are seldom true copies. Both modifications and significant changes occur as a result of adaptations to different production equipment, inventing around patents, product differentiation and new ideas. Often these changes and developments during diffusion take the form of accumulated minor improvements, but radical changes also occur. Thus an innovation is never a one-shot affair, but instead the first shot that triggers a swarm of mostly minor changes.

In combination, the subsequent changes and innovations lead to a series of minor improvements with a few major jumps in the technical performance parameters of the product (weight, efficiency, durability etc.). But entirely new functionalities may also be developed in addition (for instance portability, heat resistance etc.). Improved product performance is often correlated with cumulative production as well as with the cumulative stock of products in use, and thus can be interpreted as a result of learning. This learning effect is different from the reduction in direct unit variable cost as a result of cumulative production, which stems from improvements in the performance of the production process. An important question is which factors account for this learning effect, and whether learning takes place predominantly on the buyer or the seller side. However, it may also be argued that the really important point is not whether technological change based on learning is mainly user-driven or maker (producer)-driven, but what enables the whole system of actors to function as a learning system. Technological change may take place among makers of materials, production machinery, components, and other supplies as well as among users in different market segments in different applications and moreover among other makers and users, which are connected to different value chains.

IPR influences on diffusion

From the point of view of an innovating company, it is important to try to create rapid buyer diffusion with good profit margins while preventing a seller diffusion that reduces profit margins. Marketing efforts, in which trademarks and design play important roles, may boost buyer diffusion. Patents can also be used as a marketing tool, signalling technical superiority to prospective customers. Seller diffusion may be limited and/or delayed, for example, by keeping the new product or process technology secret, creating a strong patent situation that blocks other sellers or erecting other forms of entry barriers, such as creating costs for buyers to switch to another supplier for their rebuys (i.e. creating so called switching costs). In the case of strong patent or secret protection, the diffusion pattern may look as in Figure 3.6.

A strong patent or patent portfolio enables the patent holder to create a temporary monopoly on the market. The new product or process may also be licensed profitably to other producers in order to boost the new product. In general, studies have shown that a high market share in growing markets for new products often is associated with high profitability (at least when cost reduction due to proprietary learning is significant). But if market size can be increased by the joint efforts of many producers, it may pay off to give up a bit of the market share to other licensees.

The point to stress here is that the buyer and seller diffusion processes and their interaction jointly determine economic success and therefore must be taken into consideration in the planning and commercialization of new products. IPRs may be used to effect both types of diffusion. Trademarks, designs, and also patents typically speed up buyer diffusion, while patents, trade secrets and copyrights can be used to influence seller diffusion, typically by delaying it and thereby weakening competition for some time.

Figure 3.6 Buyer and seller diffusion with temporary monopoly

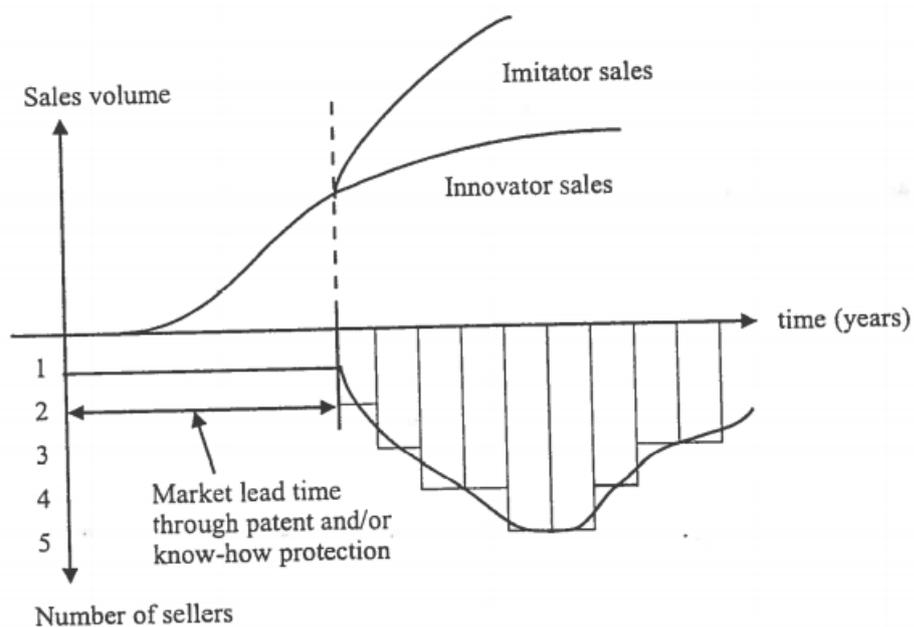


Figure 3.6 Buyer and seller diffusion with temporary monopoly

3.3 Patents

3.3.1 Introduction

The aim with this section is to give a general introduction to patents. For the interested reader there are several good manuals and textbooks providing more detailed information. One has to bear in mind, though, that IP laws, and thus patent laws, differ between countries and vary over time. For accurate information, publications from the patent offices in various countries and from the WIPO should be consulted. There are also numerous pamphlets from law firms, patent bureaus, consultants, etc. A few suggested readings are given in Section 3.5.

3.3.2 What is a patent?

Official definitions of a patent vary somewhat. The following one is from the European Patent Office (EPO):⁶

A patent is a legal title granting its holder the exclusive right to make use of an invention for a limited area and time by stopping others from, among other things, making, using or selling it without authorization.

(EPO Annual Reports)

Thus, a patent is not, strictly speaking, a technical invention or a technical document but a legal right with a possible economic value. A patent does not directly entitle the holder to exclusively sell or even manufacture the invention. In essence, it is a negative right, a right to exclude others. A patent for a drug does not entitle the holder to sell it as a medicine. For this, authorization by the national health authorities is required.

By its nature, a patent can be seen as a socio-economic contract between an inventor (or IPR holder) and society. Upon voluntary request by an inventor who fulfils certain requirements, society grants patent rights to the inventor, who in turn "pays" for the rights by

⁶ There are similar definitions for a US patent as well as for a Japanese patent.

disclosing the invention to the public. Since patent offices need to be financed, an applicant is also required to pay administrative fees. This is illustrated in Figure 3.7. One can say that there are two sides to the patent "coin":

1. Patent rights are important as competitive means for protection and commercial exploitation of new technologies.
2. Patent information is important as a means for technology and competitor intelligence.

The legal and institutional framework providing and processing patent rights (rather than inventions themselves) forms a patent system. A patent system is further characterized by:

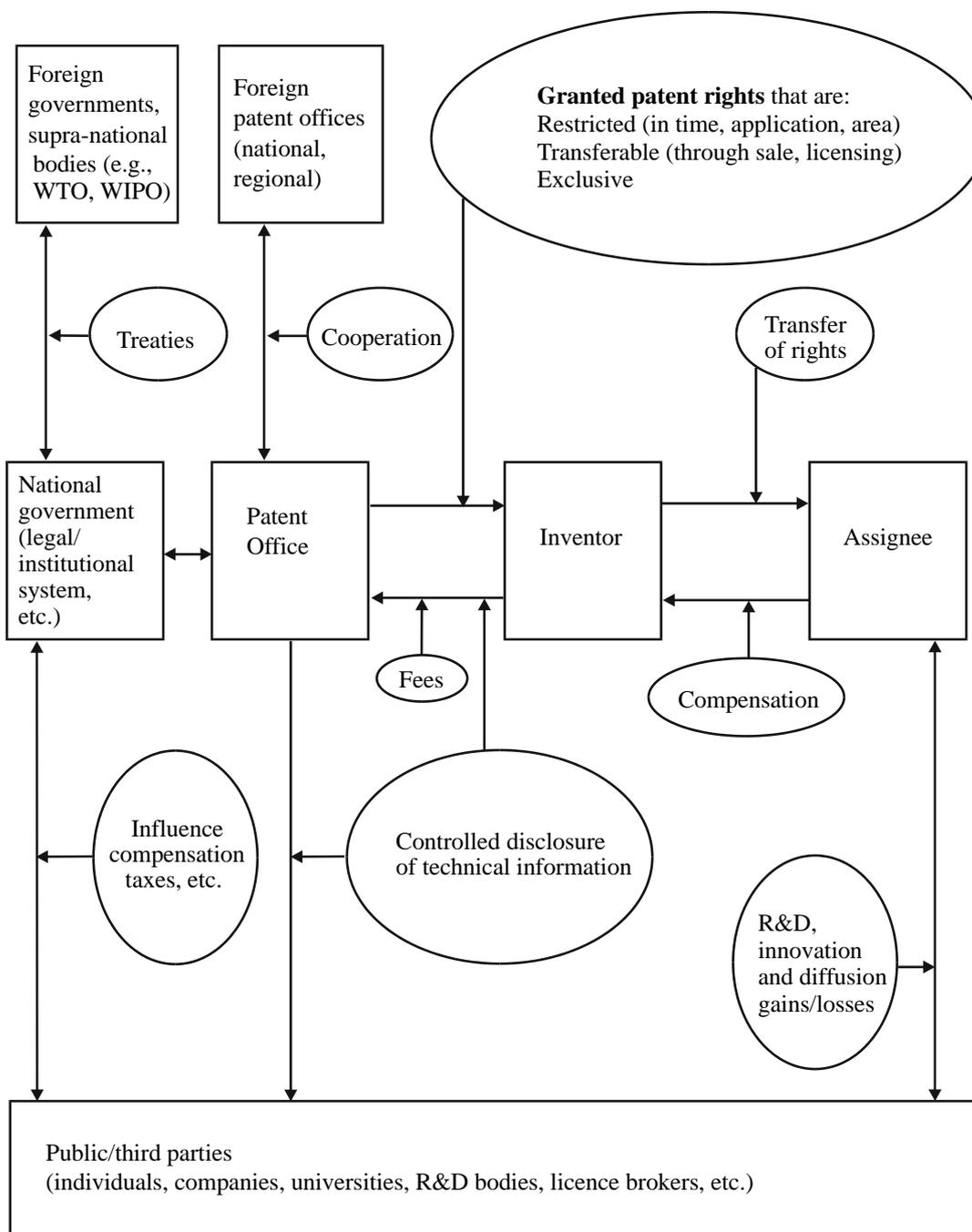
1. conditions for what is patentable
2. term of protection and territorial extension
3. protection against activities violating a patent including sanctions and their enforcement

A patent right is granted to an inventor or his/her assignee contingent upon the following conditions: (1) technical nature (industrial applicability), (2) novelty and (3) non-obviousness (see below). Novelty and non-obviousness are determined as of the first filing date (priority date). In order to have priority to an invention, the inventor must be the first one to file a patent application with a patent office. The USA is a major exception to this rule, since priority is given in the USA to the person who is (or is considered to be) the first to make the invention, that is, to conceive it and to reduce it to practice.

Across most major industrialized nations (Europe, Japan, USA etc.), the maximum lifetime or period of protection by a patent is now (1998) 20 years counted from the date the application was filed.⁷ The patent right is valid in each country where it has been applied for and granted.

⁷ Extended protection may be granted for medicines covered by a patent if the period of registration for the drug consumed excessive time. (In Europe this legal title is called 'Supplementary Protection Certificate – SPC.)

Figure 3.7 The socio-economic contract nature of a patent



Patent rights are national in the sense that they refer only to the country that granted them, and they must be applied for in each country of interest. Thus, there is no such thing as an international patent. However, a patent right can be extended across nations according to various procedures. Under the Paris Convention, the inventor has a 12-month period, starting from the first filing date, to file applications in other countries. The PCT facilitates the further processing of the original patent application (the so-called priority application) in the form of an international application. Ultimately, an invention matures into a family of patents in countries where applications claiming priority from the priority applications have been filed and approved. An exception is European patent applications, which are filed with, processed at and granted by the European Patent Office (EPO), a regional patent office for Western Europe (except for Norway). To be valid in a member state, a European patent has to be filed with the national office of that state within 3 months of the publication of grant in the national language of that country.

A patent right is violated or infringed if someone exploits the invention commercially, in contrast to using it for experimentation. The patent right holder may go to court to enforce his or her rights. That is, a patent gives to the patent holder the right to sue an infringer. This may result in the court stopping the infringer temporarily through an injunction or by the sentencing of a fraud conviction. The court may decide that the infringer pay damages to the rights holder for economic losses incurred. Thus, ultimately, the value of a patent right as a competitive tool derives from its use as a means for litigation.

3.3.3 How to value a patent

The economic value of a patent derives from several sources. These sources are related to the various advantages and disadvantages of a patent, as described below. The major source of value, as traditionally conceived, is through deterred or deferred imitation and competition. This is illustrated in Figure 3.8. Another source of value arises through the licensing of a patent as shown in Figure 3.9. The licensee is thereby entitled to use the patent rights and

may, in turn, deter or defer imitation and competition. In return, the licensee (license buyer) pays a fee to the licensor (license seller), usually in the form of a down payment and royalties that are usually calculated as a fixed or variable percentage of license-based sales.

3.3.4 What is patentable?

A technical invention fulfilling minimal requirements in respect of the following criteria is considered patentable:

- 1) Industrially applicable (exploitable within certain moral limits, not excluding weapons however). Usually, in this context it is more specifically required that the invention has technical character and can be carried out.
- 2) Novel to the world;
- 3) Non-obvious to the average "person skilled in the art" (professional practitioner), that is, the invention must exceed a certain minimum inventive step (level of invention) or level of combinatorial skill.⁸

⁸ In other words, any technical invention that is industrially applicable, not known beforehand, and not obvious is patentable. All these three requirements create legal debate, e.g. over what is meant by technical. Also note that there is only a very weak requirement on usefulness. The requirement on level of invention is fairly low in practice, especially in certain countries. The novelty requirement could be seen as strong.

Figure 3.8 Time concepts in innovation and imitation

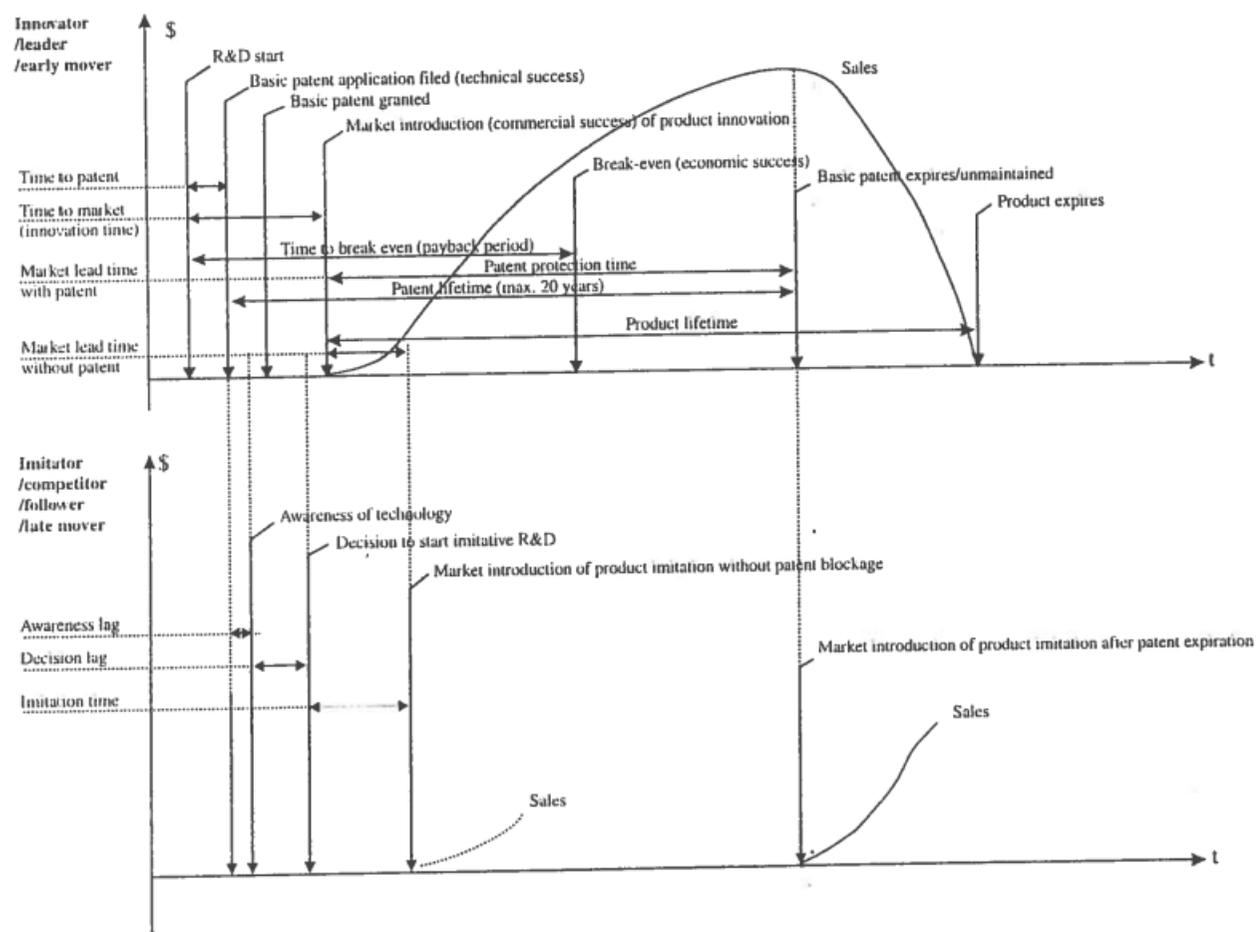


Figure 3.8 Time concepts in innovation and imitation

Figure 3.9 Economic value arising from licensing a patent

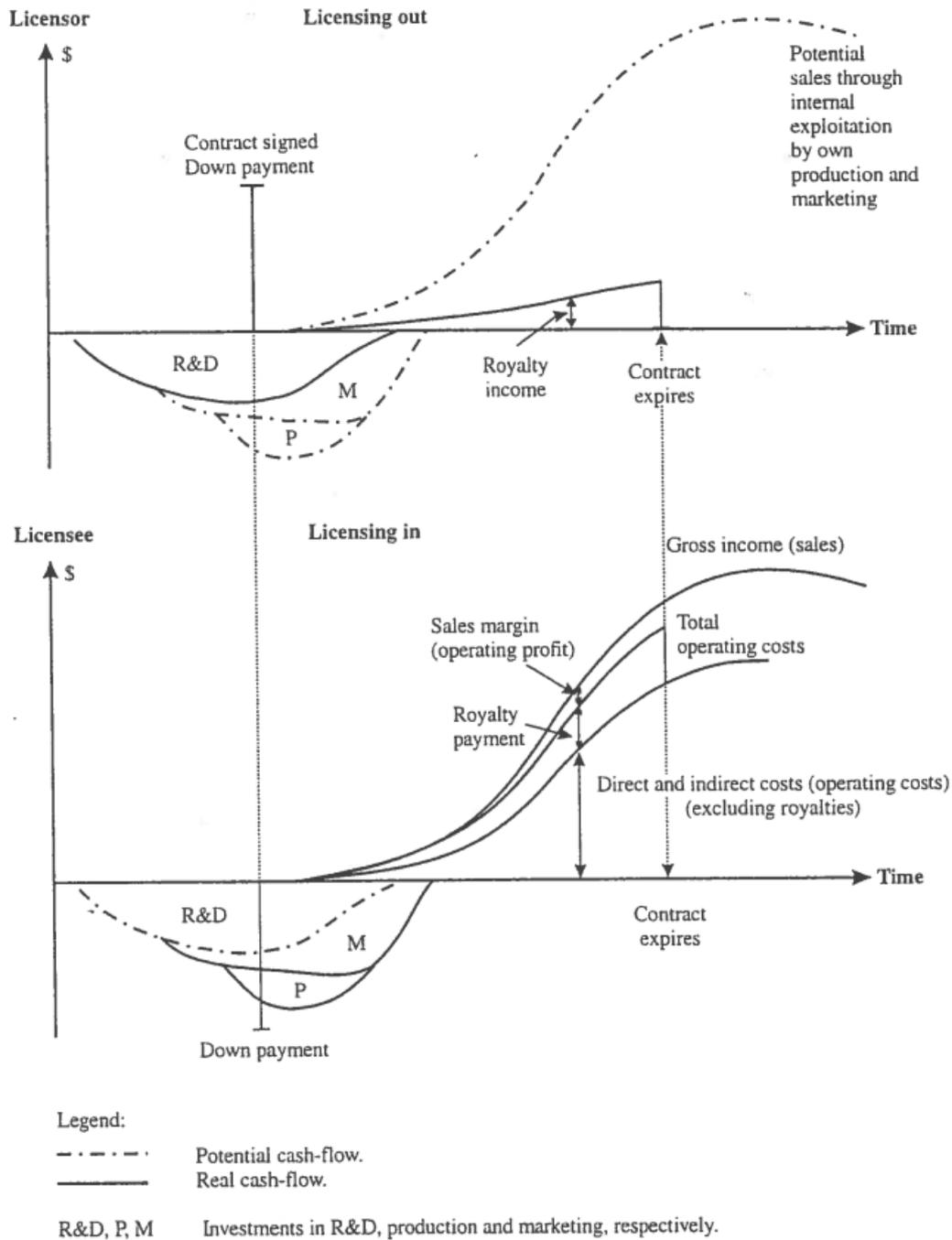


Figure 3.9 Economic value arising from licensing a patent

The minimum level of patentability requirements may vary across nations (and across patent examiners as well) and over time. There is a trend towards some convergence through international cooperation and harmonization. There is also a trend towards enlarging the patentable area, e.g. in biotechnology and computer software, which is a broadening of the first requirement.⁹ There is even variation across technologies, for instance due to new technologies presenting difficulties to patent examiners when assessing novelty and non-obviousness. It is obvious that there are not many professional practitioners skilled in a new technology. The obviousness criterion might be less stringent in a new technology, since a patent examiner might lack experience on where to draw the line between what is obvious and what is not. As a result, the minimum level of invention required by patent examiners for patentability may be lower in the early stages of an emerging technology. This translates into the scope of patents that may become too large in a new field.¹⁰

3.3.5 Who can patent?

Any individual or group of individuals can apply for a patent, and in most countries a legal person such as a company, can do so. In the USA only individuals can apply, but this is a formality since an individual can transfer his/her rights to a company. Normally an employee in industry (but not in a university) would have to assign the right for an invention to the employer. It is then the burden of the assignee to apply for a patent and to turn the invention into an innovation in the form of a marketable product or process.

3.3.6 Why apply for a patent?

Firms may benefit in several ways from applying for patent protection, since a patent may:

⁹ Neither living matter, such as genetically engineered plants or animals, nor computer codes in form of strings of symbols could readily be considered technical inventions with a narrow interpretation of "technical".

¹⁰ Criteria for assessing the proper scope of a patent are not easily established. See Nelson and Merges (1990) and NRC (1993).

- 1) Deter a competitor from introducing technologically similar inventions.
- 2) Put pressure on a competitor who introduced a technologically similar invention to withdraw from the market, if the introduction has been detected to have been made after the patent was applied for or granted (depending on the country under consideration).
- 3) Block a competitor from patenting. (This could also be achieved through prophylactic publishing, that is, publishing of information that prevents others from applying for a patent for what has been disclosed to the public, and thus cannot be considered novel anymore.)
- 4) Create an identifiable asset with certain rights attached to it, which in turn may create strength in negotiations regarding financing, licensing, cooperation, acquisition, divestment or standard setting.
- 5) Create an economic asset that could be activated on the balance sheet and used for PR, marketing, financing and also tax-planning purposes.
- 6) Create a possibility of cross-country transfer of profits within an MNC through intra-firm licensing.
- 7) Create an incentive to invent among personnel within the firm.
- 8) Create a possibility to stimulate and measure R&D productivity (even if this is a questionable and expensive measure).

On the other hand, a firm reaps benefits when other firms file patent applications, since they may:

- 9) Gain easy access to the information pertaining to a patent application since it is made publicly available.
- 10) Conduct business negotiations with the patent holder more easily with a well defined patented asset as a basis for financing, licensing, cooperation, purchasing and standard setting. For example, it is easier for a firm to buy an invention from an inventor if it is

patented, although the price may be higher for the same reason. Thus, in this respect, patents make it easier to trade technology for both the buyer and the seller, thereby reducing their transaction costs.

Benefits under the first three choices above are usually referred to as the monopoly element of patents, since the patent system is designed in such a way that "the winner gets all", that is there is only one winner in the race for a patent..

A patent offers not only advantages but a mix of advantages and disadvantages. There are several ways to list these and make the necessary trade-offs. The following is a listing used in the present study (see Table 7.1 and also Table 3.3). Thus:

With respect to the advantages to a company or an individual inventor, a patent in general offers a means for...

- 1) Protecting proprietary product technology
- 2) Protecting proprietary process technology
- 3) Creating retaliatory power against competitors
- 4) Creating better possibilities of selling licenses
- 5) Giving better possibilities of accessing technology through cross-licensing
- 6) Facilitating R&D cooperation with others
- 7) Creating a better bargaining position in standard-setting
- 8) Providing motivation for employees to invent
- 9) Providing a measure of R&D productivity
- 10) Improving the corporate image

...at the expense of the following disadvantages:

- 1) Disclosure of technical information
- 2) Direct costs of patenting

The advantages of patenting can be summarized as follows:

- 1) To block competitors
- 2) To improve the bargaining position
- 3) To stimulate and monitor R&D

Blocking competitors actually means two things: first, to block their R&D and business activities, and second, to block their possibilities to block one's own R&D and business activities. The first blocking motive has traditionally been considered less gentleman-like, while the second one is considered to be fair because it involves only protective (defensive) measures. However, as competition becomes fiercer and patenting becomes more aggressive and offensive, blocking competitors outside one's own area is becoming increasingly accepted.

It is also important to judge the pros and cons of a single patent in relation to a network or a portfolio of patents, that is, building of a patent portfolio for a product area or for the company as a whole. For data on how companies put weight on the advantages above, see Chapter 7.

3.3.7 When to apply for a patent

The timing of a patent application in the innovation process is also a matter of importance. To apply early in the process provides the following advantages by:

- 1) Reducing the risk of being blocked by others
- 2) Providing for earlier revenues stemming from the patent
- 3) Increasing the likelihood of getting a patent granted
- 4) Increasing the likelihood of obtaining a broader scope of protection in case of an emerging technology

- 5) Increasing the likelihood of setting the inventive approach as a standard
- 6) Possibly deterring others from pursuing the particular approach or line of R&D

At the same time applying early provides the following disadvantages by:

- 1) Providing shorter protection time for the product/process on the market
- 2) Providing early indication to competitors of the company's R&D activities
- 3) Increasing the risk of not securing a patent or a weak patent only due to insufficient experimental support for the invention

Again a patent application must not be judged in isolation but as a part of the company's effort to build and maintain an IPR portfolio. For example if a product already has patent protection, one could wait a while with applying for an additional patent in order to prolong the total time of protection.

3.3.8 Where to apply for a patent

The basic rule is to apply in those countries where the company has a business or expects to have a business within the lifetime of the patent, and where the value of the patent for protecting one's own business exceeds the costs for obtaining and maintaining it. The latter may be high for many countries in relation to the size of the market and the competition.¹¹ The costs of international patenting may also be too high for small companies to absorb. However, consideration of the offensive motives for patenting, rather than the traditionally defensive ones mentioned above, may justify applications in a wider range of countries. The same is true if prospective licensing is considered.

¹¹ For example, an application at the EPO allows for designating essentially all industrialized countries of Western Europe for protection. However, on average only about 7 countries are being designated in European applications.

It has also been customary to file the first patent application for a particular invention (the "basic" or "priority" application) in the home country of the company where most of its R&D are usually located. However, this habit is being challenged as R&D becomes more internationalized along with the patenting operations in MNCs. For example, if an Italian company acquires a German company with a good patent department, the company may move some of its patent application work to Germany, which will then become the country in which basic applications are filed.¹² There may also be advantages for a MNC to file first in a country with a large market and a strong IP regime like the USA. Swift and cost-effective patent processing by the patent office or favourable tax conditions are other factors that may attract first filings to a country.

3.3.9 How to apply for a patent

There are many textbooks and manuals for this purpose to which the reader is referred for detailed information. See e.g. Foster (1993) or Anawalt and Enayati (1996) for the USA, Hodkinson (1987) for the UK, guides from the JPO (1991) for Japan (e.g. JPO 1994) and WIPO's "PCT Applicants' Guide". Patent offices also provide instructional material together with application forms.

3.3.10 Pricing of licenses

As mentioned in Section 3.3.3, the net value of a patent or a patent portfolio to the patenting company is derived from many sources. The market value of a patent or a patent portfolio (package) is a different matter. It is determined by the considerations of prospective buyers. In principle, the market value of a patent can be determined as the value of a corresponding exclusive, non-geographically restricted license. Since licenses can be of various types as well

¹² If Italian inventors do the inventive R&D work, the patent application will still be considered of Italian origin, although written and filed in Germany.

(exclusive, sole, non-exclusive, etc.), the valuation of a patent or a group of patents is a special case of license valuation. To make such a valuation is no easy matter. Rather, it is extremely difficult to be precise. The difficulties arise from the unique nature of a patent, the long time horizons (up to 20 years), and the technical, commercial and economic uncertainties involved. Nevertheless, numerous license agreements are being signed every day on the basis of some kind of valuation.

Figure 3.10 gives an overview of various factors that the buyer and the seller might consider. They could be translated into ceiling price levels and floor price levels respectively, when aggregated, and this assessment could be done in various ways. Often it is useful to distinguish between two floor price levels. The higher floor price is related to total cost, including a portion of fixed R&D costs, while the lower floor price is only related to operating costs. Similarly, different ceiling price levels could be related to different ranges of factors influencing value, e.g. pertaining only to a narrow commercial value in a particular area or to a higher strategic value, incorporating spill-overs to other areas in the buying company or consequences for future license deals.

Such economic calculations, as well as contractual designs, negotiation tactics, etc. can be further elaborated. The difficulties in valuation and the considerable uncertainties involved leave a great freedom for exercising negotiation skills in licensing. A few useful hints are:

- 1) Consider option agreements and letters of intent.
- 2) Distinguish between technical, commercial and economic risks.
- 3) Go for the proper jurisdiction.
- 4) Be critical of industry conventions regarding pricing etc. Try to be legally and commercially creative as well, but stay within the legal framework.
- 5) Prepare negotiations carefully. Be aware of cultural differences.
- 6) Make investment calculations but do not use them rigidly.

- 7) Utilize legal, economic, and technical expertise.
- 8) Have a profit-sharing attitude (try to play a win-win game with empathy regarding the partner).
- 9) Consider contingencies and the need for re-negotiation clauses.
- 10) Be prepared not to reach an agreement.

The model in Figure 3.10 could be further elaborated when determining damage claims in case of patent infringement, and it could also be used in cross-licensing situations. It should be repeated that accurate valuations are difficult to attain and that time horizons are usually long with uncertain prospects. This has a tendency to raise negotiation costs, which makes it attractive to negotiate the cross-licensing of broad patent packages in certain industries. In many situations, patent mappings are also useful as a bargaining tool (see Chapter 9).

Figure 3.10 A valuation and pricing model for patents and licenses

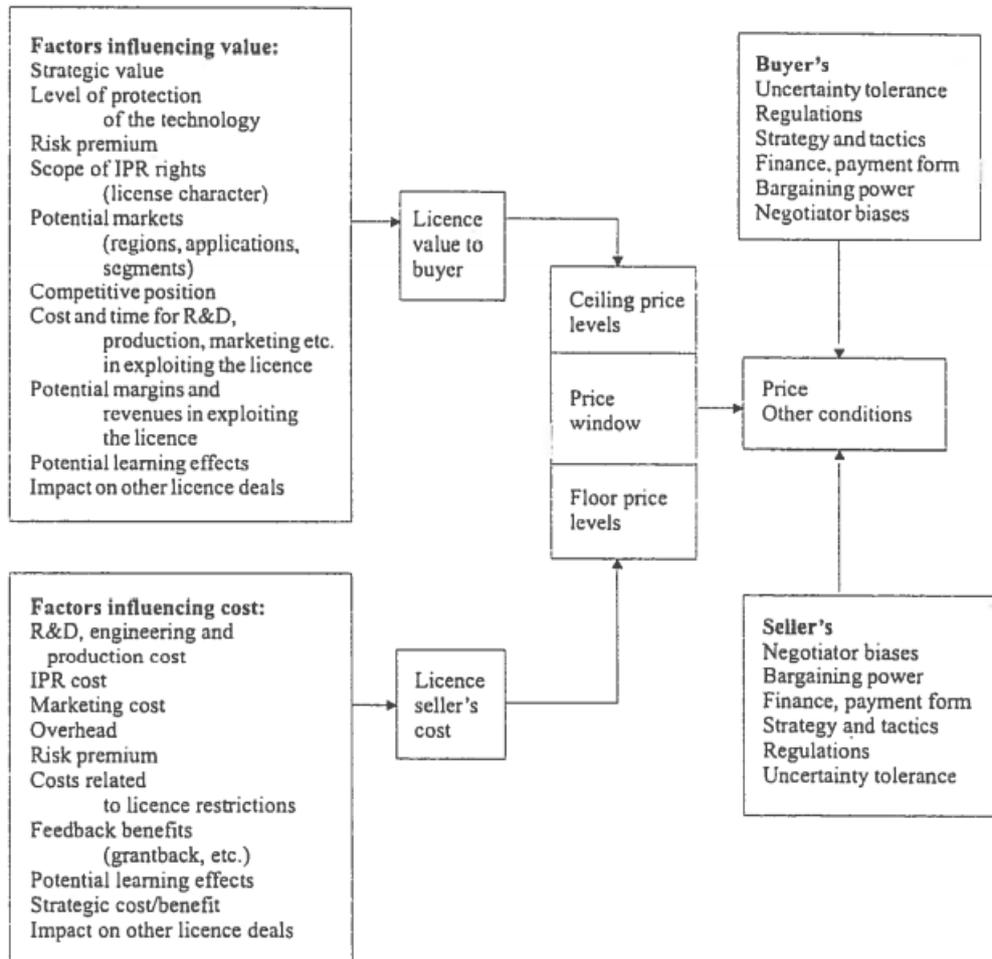


Figure 3.10 A valuation and pricing model for patents and licences

3.4 The patent system as an economic incentive system

3.4.1 What is a patent system?

The patent system in itself can be viewed as a legal or institutional innovation in the sense that its basic legal framework diffused among the countries of the world, as described in Chapter 2. The patent system will probably continue to diffuse globally. Despite long-standing public debate and concern about the positive and negative effects of the patent system, only minor modifications have taken place in recent years, which on the whole can be seen as improvements. There have been no major changes or radical legal innovations regarding the patent system in this century.¹³ The basic economic and legal ideas behind the patent system have remained much the same.

In principle the grant of a patent is a transaction between the right holder and the state. As described earlier, this transaction means that the right holder is rewarded for disclosing the information to the public by receiving the transferable, temporary, exclusive legal right to prevent others from commercially exploiting the invention. Furthermore the right holder can use legal means, which can enjoin the infringer and disgorge his profits, as a means to stop any further exploitation by an infringer who has been found guilty by a court. The traditional interpretation of this among economists is that the inventor/right holder voluntarily enters into a binding contract with society, which grants a temporary monopoly right in return for information about the invention. There is, of course, a cost of setting up and running a system with legislation, patent offices, and courts (with any patent policing costs deferred to the right holder). The fixed system costs are mostly tax-financed while operating costs are mostly fee-financed. The cost for third parties to access patent information is rather low, although it may be costly to absorb it. Figure 3.7 above gave an overview of the socioeconomic contract nature of a patent.

¹³ Some major steps towards international cooperation and harmonization are noteworthy, however, in particular the Paris convention 1883 and the implementation of PCT and EPC in 1978 (see Table 2.2).

3.4.2 Why have a patent system at all?

For a society that wants to stimulate the generation and diffusion of technical information and technical innovations, a patent system is one way.¹⁴ More specifically, the major rationale for the patent system from a traditional economist's point of view is to:

1. Stimulate invention and investments in R&D
2. Stimulate commercial exploitation of inventions through direct investments in production and marketing and/or through technology trade (licensing in and out)
3. Stimulate public disclosure of technical information

Thus, the patent system is intended to be a stimulus to investments not only in R&D but also in production and marketing. In this respect the patent system differs from an inventor prize system such as the one practised in the former USSR (with 'inventors certificates'). Public disclosure of an invention is also thought of as stimulating technological progress and competition after the patent protection has ceased. The economic benefits to society have to be weighed against economic losses due to any monopolistic behaviour of the inventor/rights holder, plus the net administrative cost of setting up and running the patent system.¹⁵ This trade-off has been at the centre of the debate about the patent system. In a country like the USA, the monopoly considerations have traditionally carried great weight. However, this has changed, as described in Chapter 2. For a more thorough discussion about the rationale behind the patent system, see Scherer (1980) and Kaufer (1989).

In addition, three things must be kept in mind when discussing the patent system as a system of incentives. First, there are different pros and cons concerning the patent system for different levels or actors in the economy; see Table 3.3. Second, there are a number of

¹⁴ There are other ways as well, e.g. through prizes and contracts (see Wright 1983 for a theoretical analysis), or through tax deduction schemes (see Mansfield 1985 for an empirical analysis). Each way has a particular set of advantages and disadvantages and there is no way that is clearly recognized as overall superior. Usually, several ways are used at the same time in the hope that they complement each other.

¹⁵ The full cost of the patent system can be seen as a kind of large transaction cost for society, that is, a cost for improving and using market mechanisms in society.

alternative policy measures for a government wanting to stimulate R&D and innovation. In this context it may be argued that the patent system is a demand side rather than a supply side measure, since it affects the output markets of a firm through its supposed benefits rather than its input markets through reduced factor costs. Third, there are numerous possibilities for modifying existing patent laws and practices, and the strengths and weaknesses of their effects as incentives depend very much on the detailed design of the patent system.

Monopolistic conditions may moreover arise in a number of other ways than through patents. If markets are regionalized with local or state monopolies – regulated or "natural" monopolies – patents do not significantly affect competition (cases in point are traditional public utility companies and old-style telephone service companies.) Monopolistic conditions also arise from lead times on the market, e.g. if imitation lags and/or imitation costs are high. A third case in which additional monopolizing effects of patents are limited is when product differentiation is so high that the competing products are essentially dissimilar.

Clearly in certain markets and in respect to certain products, the patent system may create temporary monopolistic conditions. However, the monopolistic effects are then often reduced by limiting factors such as a deficient infringement sanction systems.

Table 3.3 Principal advantages and disadvantages of the patent system

Level	Advantages	Disadvantages
Nation (society, consumers)	Stimulates rate of invention by providing an incentive for investment in R&D (also for reinvestment and for invent-around work)	Risk of monopolistic inefficiencies (including risk of hampered commercialization of new technologies) ⁴⁾
	Stimulates rate of commercialization (rate of innovation) through investment in general	Administrative costs for setting up and running the system
	Stimulates rate of diffusion and technology transfer through disclosure, marketing and licensing	Risk of R&D and investment distortion
	Provides an artificial metric (yardstick) of invention	Risk of over-investment in duplicative R&D and/or substitute inventions
Company ¹⁾	Offers restricted, transferable monopoly rights	Requires controlled disclosure ³⁾
	Provides bargaining power and a basis for buying or selling an identified piece of technology	Monopolistic over-pricing (incl. cost for acquiring technology) and/or barriers to entry induced by competitors
	Provides information about technology and industry competitors Provides motivation for employees and yardsticks for technology management	Patenting costs, direct and indirect (including e.g. litigation costs)
Individual ²⁾	Provides a basis for award negotiation of a contract or start-up of a company	Requires controlled disclosure ³⁾
	Provides a means for recognition	Monopolistic behaviour of holders of possibly interfering or complementary patents
	Provides information on technology	Patenting costs

Notes:

¹⁾ These advantages and disadvantages are of course related to a company's advantages and disadvantages of taking out patents, as described above and in Chapter 7, but they do not exactly match because the pros and cons here concern the patent system as a whole, compared to a hypothetical situation with no such system at all. Moreover, seeking to take out a patent is voluntary (as is keeping it in force, once it is granted) and a company perceives advantages of so doing in comparison with the alternative of not doing so while still having a patent system in place.

²⁾ Typically an inventor (engineer or scientist), either autonomous or employed.

³⁾ No pre-publication is allowed (i.e. before a patent application is filed) and full post-publication is enforced.

⁴⁾ The commercialization of new technologies can be hampered by the dispersion of several necessary patents (and IPRs in general) among actors who cannot agree.

The secrecy alternative

This said, a large number of cases remain where patents do have monopolistic effects. An obvious question then is what would happen if patenting possibilities ceased to exist. The main alternative open to firms is to reduce inventive activities, i.e. reduce R&D in a broad sense, and/or attempt to appropriate indigenous technology through some other means, e.g. through secrecy operations. Secrecy protection is substantially more effective for production technology than for product technology, since products may more easily be "reverse engineered". Thus, secrecy protection may to some extent substitute for patent protection regarding process inventions but less so for product inventions.

To operate within secrecy agreements would make technology trade more cumbersome. It is well recognized in economics that any market for information works imperfectly. The moment one has to disclose a piece of information in order to sell it, the person is running the risk of being cheated.¹⁶ It is quite possible, at least in principle, to replace a patent with secrecy agreements in business negotiations for a piece of technological information. However, the enforcement of such agreements might become very difficult since the licensor would have to show that the secret information was divulged by the licensee. Thus, the seller has to prove ex post that his invention or know-how was uniquely new and unknown to others at the time the secrecy agreement was signed. This is a formidable burden of proof.¹⁷ Alternatively, the seller might try to establish ex ante that the invention is new. However, this is also a formidable task to perform without disclosing too much. Hence, the enforcement of a secrecy agreement leads to the necessity of establishing novelty and unique proprietorship of the information passed under the agreement. The difficulty of doing so without disclosure leads one to conceive of some kind of system where novelty and unique proprietorship are established ex ante through public disclosure. In effect, some kind of patent system would likely result as a consequence of difficulties in enforcing secrecy agreements, as well as

¹⁶ This is the so-called "information paradox" (Arrow 1962).

¹⁷ The burden of proof could of course be shifted to the other party, but to prove that one did not know something at a previous time is also difficult in general. (See also Chapter 2.)

difficulties in disseminating information in a secrecy regime.¹⁸ In summary, the transaction costs for conducting technology trade purely under secrecy agreements would be too high, and technology trade would shrink.¹⁹

Prestige system

In absence of a patent system, one alternative open to society and firms is to stimulate invention through some kind of inventor prize system. Solely from the point of view of stimulating employee inventiveness, it is likely that such a system could be designed to substitute quite effectively for a patent system. In fact, nothing prevents companies, academies or nations from running their own such systems in parallel with the patent system, and some do so even in areas not covered by patent laws (e.g. general suggestion systems in companies, or the Nobel Prize system in some sciences). However, such systems mostly provide no economic incentives for commercializing and diffusing new technologies.

It is finally important to keep in mind that forces other than those based on purely economic incentives play a role in advancing science and technology. Psychological motives such as curiosity, vanity and the need for achievement are important to scientists and inventors. Such non-monetary, socio-psychological incentives also often stimulate invention. Patents, too, are sometimes applied for by inventors not only for purely economic reasons, but since a patent provides a certain amount of recognition and prestige. It should also be noted that the prospect of a patent could be seen as a prize or a reward that could be overvalued by inventors investing in R&D. From society's point of view it would, at least in the short run, be beneficial if firms persistently expected the patent system to function to their advantage

¹⁸ Patent systems do differ in their design of what and when to disclose to whom, and numerous variants are conceivable, even a system in which secrets are filed and registered and open only to authorized agencies in case of disputes. Such a non-public "patent" system does not even have to be run by the state. However, dissemination of technical information would not be encouraged and infringement would be costly to police.

¹⁹ Not entirely, however, since a good portion of license agreements are not pure patent licenses, but are a mixture of complementary patent and know-how licenses. In addition there are pure know-how licenses. See also Anton and Yao (1994).

even if it did not. To some extent this kind of false expectations can be found among young firms and inventors. (See Dutton 1984, p. 204-205 about this ‘socially wholesome illusion’.) This is analogous to what is said about entrepreneurial hubris in Chapter 4.

3.5 Previous research and literature

Despite the practical importance of trade secrets to the business community, the law of trade secrets is a neglected orphan in economic analysis.

Friedman D., Landes W., and Posner, R. (1991, p. 61)

Trade marks are ... significant business assets; ... Yet, they have not ... been systematically studied by economic or business historians, even though much has been written by other scholars on these intangible assets.

Wilkins (1992, p. 66)

3.5.1 General structure of the literature

The literature on intellectual property has a number of interesting structural features, As a subject, intellectual property, and patents in particular, have indeed a long tradition in scholarly literature on economics, law, engineering, history, philosophy etc., as well as in miscellaneous other writings (government investigations, biographies, debate articles etc.). It has been a tiny tradition, however, although there is a recent growth in writings on patents. In general the IP-related literature is also highly fragmented, corresponding to the different IPR types (patents, copyrights, trademarks, etc.). Quite naturally the literature is also subdivided into disciplines, primarily economics, management, law and engineering. Also quite naturally there is academic literature and literature for practitioners, with a substantial separation of these two general strands. (Perhaps this separation is less in some disciplines, e.g. law.) Finally, there is a fragmentation of the literature by nation.

In summary, the IP literature has a long and thin, but growing, tradition, highly fragmented by IPR type, discipline, occupation and country. This makes surveying difficult. Below, the patent-related literature in English after World War II will be briefly surveyed in order to provide some different entries into the literature in general. More specific references are given in the rest of the book.

3.5.2 Surveys of the literature

For a long time, the patent institution and its pros and cons have stirred much controversy and given rise to innumerable publications about its economic, legal and technical aspects. In this debate and literature, the USA have held a central place. It is difficult to give more than a biased sample of literature surveys. For good such surveys, see Machlup (1958), Taylor and Silberston (1973), Scherer (1977), Scherer (1980), Scherer (1993), Griliches (1984), Kaufer (1987), MacLeod (1988), von Hippel 1988, Tirole (1988), Weil and Snapper (1989), NRC (1993), and Lanjouw and Lerner (1997). Good surveys on the history of the patent system are given in Kaufer (1987), MacLeod (1988) and David (1993). (See also Chapter 2.) Literature on international aspects can be found in Rushing and Brown (1990) and Albach and Rosenkranz (1995).

Surveys of patent literature related to management and business economics (rather than political economics) are more rare. A bibliography is provided in Clarke (1993). These sorts of writings are mostly embedded in the literature emerging since the 1950s on management of R&D, engineering, technology, and innovation. See also Granstrand (1988).

Surveys of patent literature related to law also seem rare, but some US-oriented ones are found in Besen and Raskind (1991) and NRC (1993). Others are Barton (1991) and, for IP more generally, Goldstein (1993,1994) and Merges (1992) for textbook related surveys and Gire (1992) for a bibliography.

An increasing number of patent-related articles deal with patent information and statistics as technological and economic indicators. Surveys of this type of literature, mainly

on technological indicators, are provided in Basberg (1987), Van Raan (1988), Chakrabarti (1989), Archibugi (1991), Pavitt (1985), Pavitt (1991), Grupp (1994) and OECD (1996). Outstanding surveys on patent information as economic indicators are given in Griliches (1984), Griliches (1989) and Griliches (1990). A pioneering study using patent information for economic analysis is Schmookler (1966), which also gives an overview of a few, pre-World War II studies. (See also further Chapter 9.)

Tables 3.4a,b,c give an overview of the academic writers on patents who are most cited in articles in the Social Science Citation Index for the period of 1990-95. Almost all the authors in the table are economists from the USA, although law and natural sciences are represented, and almost all are alive and active in the field, although only a few work primarily in this field. Japanese authors, and authors with mainly an engineering background, are absent even in a list extended to forty authors. Such an extended list includes only one Nobel Prize Laureate (K. J. Arrow).

Table 3.4a Academic writers of patent articles in Social Science Citation Index 1990-1995¹⁾

Author	Number of citations	Author	Number of citations
Scherer, FM	48	Gilbert, R	18
Mansfield, E	33	Katz, ML	17
Eisenberg, RS	28	Narin, F	17
Griliches, Z	28	Dasgupta, P	16
Merges, RP	25	Nelson, RR	16
Chisum, DS	24	Pavitt, K	16
MacLeod, C	24	Scotchmer, S	16
Nordhaus, WD	24	Kitch, EW	15
Schmookler, J	21	Klemperer, P	15
Levin, RC	19	Carpenter, MP	12

Table 3.4b Academic works most cited in patent articles in Social Science Citation Index in 1990-95²⁾

Work		Number of citations	Work		Number of citations
Eisenberg	(1989)	21	Dasgupta and Stiglitz	(1980)	8
MacLeod	(1988)	16	Merges	(1988)	8
Nordhaus	(1969)	15	Narin et al.	(1987)	8
Schmookler	(1966)	15	Basberg	(1987)	7
Merges and Nelson	(1990)	14	Carpenter et al.	(1981)	7
Griliches	(1990)	13	Chisum	(1986)	7
Kitch	(1977)	13	Dreyfuss	(1989b)	7
Klemperer	(1990)	12	Katz and Shapiro	(1985)	7
Eisenberg	(1987)	11	Mansfield	(1986)	7
Levin et al.	(1987)	11	Tandon	(1982)	7
Scherer	(1980)	11	Taylor and Silberston	(1973)	7
Gilbert and Shapiro	(1990)	10			

Table 3.4c Journals with most cited patent articles in Social Science Citation Index in 1990-95³⁾

Research Policy
American Economic Review
Rand Journal of Economics
Quarterly Journal of Economics
Scientometrics
World Patent Information
Yale Law Journal

Notes:

- 1) The table is based on citations of first-listed author in all articles in SSCI for the period, having "patent" in their title. There were 442 articles published in 1990-95. I thank Dr. Olle Persson, University of Umeå, Sweden for providing material for this table.
- 2) Based on same set of articles as in Table 3.4a. The full reference is given in the list of literature references.
- 3) Based on same set of articles as in Table 3.4a. The journals are ranked according to the number of articles they contain from a table of the 50 most cited works, similar to Table 3.4b.

As for literature catering more to practitioners than to academics (still English-language literature), there are numerous handbooks, manuals and the like (some of them also containing good material of academic interest). The various patent offices and agencies like the WIPO also put out a substantial amount of publications, apart from their legal publications. A practical guide for Japan is JPO (1994), Foster and Shook (1993) for the USA and Hodgkinson (1987) for the UK. OECD (1997) describes patents in the international context. There are also several journals such as "Management of Intellectual Property", "LES Nouvelles" (published by the Licensing Executive Society), "World Patent Information" and "Journal of the Patent Office Society". Several societies also publish IP-related material, such as AIPPI.

For licensing, there are a number of information services and intermediaries. As expected, these are neither rare nor difficult to find. Occasionally these sources of information contain material of more general interest, not least since the field is wide open for legal and commercial creativity at the same time as it focuses on technical and artistic creativity. For some business oriented works on licensing, see e.g. Bidault et al. (1989), Parr and Sullivan (1996) and Schlicher (1996).

3.5.3 Theoretical literature

A relatively small but recently growing number of studies of the economic theory of the patent system have emerged since the 1960s. A seminal, even monumental, work in this category is Arrow (1962) who presented a basic model of invention, R&D innovation and imitation.²⁰ Arrow essentially argues that there is a tendency in industry to underinvest in R&D from society's point of view, due to problems for a firm to appropriate the economic benefits of its R&D.²¹ Patent protection would be one way of coping with this, at least to

²⁰ Illuminating interpretations and some critique of Arrow (1962) are given in Cheung (1986) and Winter (1989).

²¹ Other subsequent works have argued that overinvestment in R&D may very well accrue, due to various factors, a.o. the tendency for firms to overly invest in R&D in order to win patent races and "fish in common

some extent.²² The principal way a patent affects invention and innovation is through its effects on the rate of imitation. In the Arrow type of patent modelling, the innovator's profits dwindle completely by competition when imitation occurs. Thus a delay in imitation through patent protection would be a stimulus for firms to invest in R&D, at the expense to society of the possible over-pricing of products by the monopolistic patent holder.

Nordhaus (1969), which is also a truly seminal work on the economic theory of patents, makes a thorough theoretical analysis of the cost and benefits to the firm and to society of the patent system in the Arrow type of framework. Nordhaus distinguishes between different types of inventions, minor and major (or "run of the mill" and "drastic" using his terms), and in particular, postulates the optimal length of patent protection time from society's point of view. (Section 3.6 further describes the Nordhaus modelling.) By increasing the length of patent protection, incentives for generating innovations are increased (i.e. dynamic efficiency is increased), while producing a longer period of monopolistic inefficiencies, (i.e. static efficiency is decreased). The Nordhaus work triggered several works in the 1980s on the optimal length of a patent under different conditions.

More recent work in the 1990's have shifted from focusing on the optimal length of a patent towards the optimal breadth or scope of a patent as well as optimal combinations of length and breadth. For works in this vein see especially Klemperer (1990) and Gilbert and Shapiro (1990). The scope of a patent is far more difficult to parameterize, however. There have been various approaches, for example, the scope of a patent could be represented by the patentor's ability to raise price (Gilbert and Shapiro 1990), the probability of infringement, the impact on close product substitutes (Klemperer 1990), the number of side classifications of a patent (Lerner 1994), and the invent around costs. The latter two approaches have the

pools" of technological opportunities. See e.g. Dasgupta and Stiglitz (1980) and Dasgupta (1988). For an empirical and rather inconclusive work on this issue in pharmaceutical industry, see Cockburn and Henderson (1994).

²² Another way is through publicly financed R&D. For a theoretical analysis of patents versus alternative incentives like prizes and contracts, see Wright (1983).

advantage that they do not rely as much on observations or estimates of post-innovation conditions.

Another shift in focus in the theoretical literature is from considering a one-stage innovation process towards a multi-stage innovation process. Building partly on Barzel (1968), Kitch (1977) introduced a new perspective on the role of patent rights, viewing them (in analogy with prospect rights in mineral extraction) not as rewards but as prospect rights, which were handed out at an early stage in the innovation process²³. Kitch's work has been highly cited but also criticized (see e.g. Beck 1983). Other works using multi-stage models of R&D and innovation include Fudenberg et al. (1983).

There is still another shift in thought from considering only a single innovation towards considering multiple innovations that build or interact upon each other. As is often the case empirically and as studied by Mansfield, a patent loses value because new and better subsequent inventions appear before the patent expires. At the same time a strong patent influences the patentability and profitability of subsequent inventions (see e.g. Scotchmer (1991) and Aghion and Howitt (1998, Ch. 14).

Game-theoretical modelling of patent races among competitors has also become popular. In many respects stylized patent races offer a theoretically appealing application for game theory. More importantly this literature throws light on how competitive races impact incentives for R&D and innovation. For further study, see e.g. Reinganum (1982), Fudenberg et al. (1983), Tirole (1988), Dasgupta (1988) and Dasgupta and David (1987).

There are many policy variables for a patent system other than patent length and breadth (e.g. regarding disclosure of patent information). More and more of these other features of a patent system have become subjected to theoretical economic analysis. See e.g. Ordoover (1991) on how different features affect diffusion of technical information. For a classic

²³ Another expression for this type of right is "license to hunt" (or fish).

qualitative review of theories of the pros and cons of patents, see Machlup (1958) and for a current review (with similar classification of theories), see Mazzoleni and Nelson (1998).

Finally there are some (but much fewer) economic analyses of IPRs other than patents. For trade secrets, see Cheung (1982), Friedman et al. (1991), for designs see BIE (1995), and for copyrights, see e.g. Palmer (1986), and Towse (1997). The issue of patentability vs copyrightability of algorithms, databases and computer software in general has generated a great deal of literature in recent decades, theoretical and empirical, and perhaps more legal than economic. See e.g. Chisum (1986), OTA (1992), NRC (1993), Reichman and Samuelson (1997) and the issue of *Columbia Law Review*, Dec. 1994.

3.5.4 Empirical literature

When turning to empirical studies, important studies have been made by Mansfield and others on the considerable gap between the very high social returns to innovation and the private returns to innovation, as well as the importance of the rate of imitation behind these gaps. These studies indicate empirically that underinvestment in private R&D is likely because of imitation. This in turn indicates the need for something resembling a patent system. However, the studies also show that imitation is a costly and time-consuming process, affected by many more factors than just patents (see Scherer 1980, Mansfield et al. 1981). The question then is how effective the patent system actually is in practice and what would happen if it was changed or even abolished.

A classical and wide-reaching study of patenting practices in UK industry is presented by Taylor and Silberston (1973). This study also provides some data on the impact of patent lifetimes on R&D budgeting. Their main findings, based on interviews with 27 UK firms, indicate that without effective patent protection R&D budgets would be cut marginally (by 5 per cent or less), except in certain specialty chemicals, where R&D would be cut by 25 per cent, and in pharmaceuticals, where R&D would be cut by 64 per cent. These findings can be

contrasted with the empirical results of the study presented in Chapter 5 and the theoretical calculations in Section 3.6.

Mansfield (1986) sheds further light on the impact of an abolition of the patent system on the rate of invention and innovation, as estimated by firms. In essence, the study shows that the effect would be a very small decline in most industries. As almost always in these types of studies, the exceptions are pharmaceuticals and chemicals for which the patent system is essential. However, Mansfield shows that despite this outcome the firms make frequent use of the patent system. This "patenting paradox" requires further explanation. Since the propensity to patent was found not to have declined in the US from 1960 to 1980, the observed decline must be due to a decline in the number of patentable inventions. Because this number is related to the amount of R&D investments, a decline in patenting without a corresponding decline in R&D investments could be due to the (temporary) presence of diminishing returns to R&D. These issues have been extensively studied by Griliches, Hall and others, see Griliches (1984, 1989, 1990) and Hall (1994). The main conclusion from those studies is that there are important but complex links between the benefits of patents, R&D and innovation over time. These links are possible to study and much remains to be done.

The propensity to utilize the patent system, i.e. to prefer to patent an invention in face of alternatives, has attracted a number of empirical studies, see e.g. Scherer (1983) and Arundel and Kabla (1998).

The well-known Yale study by Levin, Nelson and others (see e.g. Levin 1986, Levin et al. 1987, and Klevorich et al. 1995) investigated, through a survey of hundreds of R&D managers in more than a hundred industries, the strong sector-specific variations in appropriability conditions and the role of patents in different industrial sectors. Briefly expressed, markets are imperfect and so are patents and patent systems. Thus innovations will continue to appear even without patents, and patents will not be sufficient to recap the benefits from innovation in general. Regarding the economic value of patents, much important work has been done by Scherer, see for example Scherer (1998). The distribution of

patent values is generally found to be very skew, sometimes possessing no finite mean and variance.

As mentioned, the empirical studies of patents have grown rapidly since the 1980s. Several factors have spurred this growth. The increased availability of large, electronic data bases concerning patents and R&D and the availability of computers have enabled and lowered the cost of many types of analyses. As is well known to investigators (e.g. Schmookler), a manual analysis of the rich and varied mass of patent documents is a Herculean task. Moreover, increased international technology-based competition and the emergence of a pro-patent era in the 1980s has generally spurred the interest in patents among both practitioners and scholars.

3.6 Economic Theory of IP and Patents

One cannot argue the relative merits or demerits of various features of the patent system without analyzing the social costs and benefits involved. This is economic analysis and no amount of legal training or engineering experience or technological research will equip the 'expert' for it. Yet, most of the economic theories enunciated in discussions about patent reform have come from lawyers, engineers and technical experts with occasional contributions from business executives.

Fritz Machlup, in Foreword to Penrose (1951, pp. vii-viii)

The patent system is much older than any formalized theories and models specifically related to patents. Theory has certainly lagged behind practice in so far as patents are concerned. It is interesting to observe how many of the significant ideas and problems with patents were experienced in Renaissance Venice, while it took the development of neo-classical microeconomic theory after World War II for significant theoretical achievements to be made regarding patents.

Seminal theoretical works include Arrow (1962) and Nordhaus (1969), with a geometric interpretation by Scherer (1972). These works all use neo-classical theory and tools. Some modelling in this spirit will be illustrated below in this section.

To put the subsequent illustrations in context, it must be recalled that the major problem to be addressed by a society that wants technical and industrial progress is the tendency to underinvest in R&D and its exploitation.²⁴ To correct for this tendency, incentives must be created for additional investment and its exploitation. This was as relevant a problem in Venice as it is in the various economies of today, and derives from the nature of valuable information being expensive to produce but, once generated, being cheap to reproduce (and imitate) by others. The problem can be attacked in many ways, of which setting-up a patent system is one. Other ways are to use e.g. prizes, grants, contracts or R&D tax deductions as mentioned. All these methods have their advantages and disadvantages. The first task for economic analysis is then to compare a patent system with its alternatives.²⁵ A patent system may also come in many varieties with many decision variables for inventor firms and policy-makers. Therefore a second task for analysis is how to design a specific system. Third, there are other types of IPRs besides patents and these types also come in many varieties. Below we will focus mainly on patents but will also make comparisons to other types of IPRs.

Let us first look at the situation of a new invention from the point of view of an inventor or, rather, the firm developing the invention (the innovator). Figure 3.1 above shows the standard cash flow picture of a new product or a product with a new cost-saving process. In fact the picture applies to any product or service, including cultural items such as books, films, records, concerts, etc., as long as it generates cash flows on a market and comprises some element of investment. Thus we will refer in the sequel to any of these items as a product.

²⁴ Overinvestment in R&D may of course occur as well, but it has seldom been perceived as the dominant tendency, especially not for countries trying to catch up technologically, for which purpose the patent system has historically often been used.

²⁵ This is done in Wright (1983).

The new product is being developed during a certain period of time, which we can call time to market or innovation time and has a corresponding market lifetime. During the market life time it is subjected to price competition, which typically is weak at an early stage and stronger at a later stage (thereafter it can possibly be weakened again). Competition is weaker at an early stage due to delays in effective imitation.

These imitation delays or lags have various sources. They may be due to information and decision lags, entry barriers, imitation difficulties, and competitive disadvantages in general. In particular the innovator's property rights, physical as well as intellectual, may be used by the innovator to hinder imitation, since private property by its very nature confers rights on the property holder to exclude others from using the property. Thus, any resource subjected to private property (e.g. a mine, a plant site or a store location) has an intrinsic monopolistic element deriving from the excludability associated with the property in question. When private property is used as an input to a production process or in a business context at large, it provides a certain type of input monopoly. Albeit narrow, this input monopoly position can be used to build up a monopoly position on the output product market or at least weaken imitation-based price competition on that market. In this sense, any private property serves as a source of potential competitive advantage. In particular, intellectual property, including secrets as well as patents, designs or copyrights, can be used by the innovator during a limited time to weaken competition (or strengthen the own competitive advantage). For patents, designs and copyrights, the time limit is fixed by law and known to the players; for secrets, time runs out when the secret is revealed or independently discovered, and thus the time limit is highly uncertain. In addition, information about the spread of certain secret information may be kept secret in turn. This is what we can call "second-order secrets." Thus, a secret-holder sometimes may not know if his secret is out or not, which poses an additional uncertainty about its duration.

Trademarks are also used to weaken competition, since a well-known trademark induces a customer preference in a purchasing situation, and the exact way this is done cannot be imitated. However, a trademark does not have a definite lifetime.

Thus we can view the innovator's IP as giving rise to a temporary competitive advantage, limited to an early stage of the product's market lifetime. This competitive advantage is used to generate a stream of operating profits (i.e. profits before depreciation, financial costs and taxes). The operating profits may derive from product inventions, cost-reducing process inventions or through other advantages in production and marketing, e.g. acquired through learning by doing rather than through organized R&D. Learning advantages in turn may be viewed as deriving from minor, cumulative improvements corresponding to minor inventions, guarded as secrets, and thus constituting IP as well.

Now let us turn to some formalized illustrations. The following variables and notation will be used below:

c	cost (total variable cost $c = c(q)$)
p	price
q	quantity sold at price p, i.e. the demand function is $q(p)$ with an inverse demand function $p(q)$
r	discount rate
S	sales = $p \cdot q$
π	innovator's operating profit = $p \cdot q - c(q)$
R	R&D investment
P	production investment expenditures
M	marketing investment expenditures
I	total investment expenditures = $R + P + M$
t	time (considered continuous)
V	total value discounted to time $t = 0$, i.e. present value
D	discounting operator (continuous), discounting to time $t = 0$
E	expectation operator
T	technical performance variable
L	length of period of patent or IP protection
\hat{L}	maximum L as offered by law
k, α	constants
I, II, III	time intervals $[0, t_1)$, $[t_1, t_2)$, and $[t_2, \infty)$ corresponding to stages of competition, where t_1 is the market introduction time point and t_2 is the time point when IP protection ceases. (Note that since there is a time delay for an invention to reach the market, L is normally larger than $t_2 - t_1$, which in turn is normally larger than zero.)

All variables (except V, L) are functions of time in the general case, in which cash-flow variables correspond to instantaneous cash-flow intensities, i.e. cash flows per unit of time.

The innovator's total discounted value V derived from a new product or process invention is then:

$$V = \int_0^{\infty} \pi(t)e^{-rt} dt - \int_0^{\infty} I(t)e^{-rt} dt \quad (1)$$

If $\pi(t)$ is totally dependent upon IP protection effective in the time interval $[t_1, t_2]$ then:

$$V = \int_{t_1}^{t_2} \pi(t)e^{-rt} dt - DI \quad (2)$$

Now, the profit function over time $\pi(t)$ can assume different shapes. In the simplest case π is a positive constant during a period $[0, L]$ and zero otherwise, in which case:

$$V = \pi \cdot \frac{1}{r} (1 - e^{-rL}) - DI \quad (3)$$

Simple rules of thumb for valuing a patent could then be derived under these very restrictive assumptions. For example, if a patent allows an innovator to keep a sales margin of 15 per cent on roughly stable sales due to enhanced product performance and/or cost savings from the invention during 10 years with a discount rate of 7 per cent, then the gross value of the patent (apart from discounted investments DI) is:

$$0.15 \cdot (1/0.07) \cdot (1 - e^{-0.7}) = 10.8\% \text{ of the total 10 year sales.}$$

Numerous extensions and variations are now possible. To be able to aid R&D investment decisions in principle, the dependence of π upon R must be specified. This could be done by introducing a function $T=T(R)$, a customer utility function $U=U(T)$ and a T -specific cost function $c=c(q, T)$. (The function $T(R)$ is sometimes called an effort curve.) Considerations of T could also be bypassed by linking c directly to R through specifying

$c=c(q,R)$ instead.²⁶ This is the approach of Nordhaus (1969), who introduces an "invention possibility function" for constant unit costs c (i.e. $c(q) = c \cdot q$), being reduced from c_0 to c_1 by a cost-reducing patented invention derived from spending in total R on cost-reducing R&D, according to

$$c_1 = c_0(1 - kR^\alpha) \quad k > 0, \alpha \in (0,1), R \geq 0, c_1 > 0 \quad (4)$$

Assuming e.g. that the patent (or patent portfolio) offers a perfect monopoly on the product market during a period of length L (stage II) and perfect competition before (stage I) and after this period (stage III), Nordhaus derived conditions for the innovator's optimal spending on R , regardless of whether the innovator chooses to license out the invention fully or chooses to go alone.²⁷

Assume moreover that the invention is sufficiently minor so that the pre-invention price $p=c_0$ and output q_0 remain the same during stages I and II. In stage III competition forces the price to fall to post-invention cost c_1 with a corresponding expansion of output from $q(c_0) = q_0$ to $q(c_1) = q_1$. Then in the Nordhaus type of modelling:

$$\pi(I)=\pi(III)=0 \quad (\text{i.e. } \pi(t)=0 \text{ for } t \in I,III)$$

$$\pi(t)=q_0(c_0-c_1) \text{ for } t \in II = [t_1, t_2)$$

$$V = \pi \cdot (1/r)(e^{-rt_1} - e^{-rt_2}) - DI$$

Putting $DI = R$ and inserting c_1 according to (4) gives:

$$V = q_0 c_0 k R^\alpha (e^{-rt_1} - e^{-rt_2})/r - R \quad (5)$$

Putting $\partial V/\partial R = 0$ gives the necessary condition for any optimal R&D investment $\bar{R} > 0$:

²⁶ Here, as is common, the symbols for variables and functions are used with differing connotations. Thus, the function c is not the same in the cases $c(q)$, $c(q,T)$, and $c(q,R)$, but no misunderstandings are likely to derive from this convenient practice.

²⁷ This latter alternative typically requires more investments P and M in production and marketing. Nordhaus disregards these; as long as they are independent of the optimizing variable R , it is not essential for the way the analysis is done.

$$\bar{R} = (\alpha q_0 c_0 k (e^{-rt_1} - e^{-rt_2})/r)^{1/(1-\alpha)} \quad (6)$$

Since $\alpha \in (0,1)$ implies $\partial^2 V/\partial R^2 < 0$ for $R > 0$, \bar{R} is in fact optimal and maximizes V if $V(\bar{R}) > 0$.

This means that e.g. the length of IP protection α cannot be too short in order to generate positive optimal investments in R&D.²⁸

As an illustration, we can calculate the sensitivity of the theoretically optimal R&D investments to changes in the maximal patent lifetime. Assume therefore that the maximal patent lifetime is changed from 20 to 23 years. Assume moreover that $r = 0.1$ and $\alpha = 1/2$. Then

$$\bar{R}(t_2=t_1+23) / \bar{R}(t_2=t_1+20) = ((1-e^{-2.3}) / (1-e^{-2.0}))^2 = 1.083$$

Thus lengthening the maximal possible patent protection time from 20 years to 23 years induces companies in this model to spend 8.3 per cent more on R&D. If the change is instead from 17 to 20 years, the corresponding R&D increase is 11.9 per cent. A decrease from 20 to 10 years would decrease R&D with 46.6 per cent.²⁹

So far the Nordhaus type of modelling has given the optimal response to patent or IP protection of the inventor or the inventing firm in terms of its investments in R&D. Now comes the question: what is the optimal patent system from society's point of view? The social planner or policy-maker has a wide variety of policy variables, such as:

- 1) Maximal length of patent (design, copyright) protection. (However, length of secrecy protection has not been considered as being at the policy-maker's disposal, although this is possible.)
- 2) Starting point in time for the protection period.

²⁸ It might of course happen, depending upon the coefficients in (5), that it does not pay at all to do research, i.e. V is negative for any positive R , in which case the optimal solution is $R=V=0$.

²⁹ This simple calculation actually gives results that are of the same magnitude as the empirical results reported in Table 5.11 in Chapter 5. For example, the chemical companies (for which patents have traditionally meant most) in Japan reported in 1992 they would increase their R&D spending with 8.5% in response to an increase of patent protection time with 3 years.

- 3) Minimal level of inventive step or non-obviousness of an invention or state-of-art advance.
- 4) Maximal scope of patent protection in terms of coverage.
- 5) Compulsory licensing arrangements.
- 6) Patenting fees
- 7) To whom to grant IP protection.
- 8) For what to grant patent protection.

Nordhaus (1969) focuses on society's optimal, maximal length of a patent protection period. In order to illustrate how the optimal \bar{E} can be calculated in a simple case, assume first a general (but declining) demand function $p(q)$ and that the invention (or innovation) reduces a constant marginal cost from c_0 to c_1 as above and as shown in Figure 3.11. The invention is minor in the sense that pre-invention marginal cost is lower than the monopolistic price for the invention. Further, assume that R occurs instantaneously at $t=0$ and put $t_1 = 0$ and $t_2 = L$.

The benefit to society of this invention then equals the producer net surplus, which is realized as the innovator's discounted value $V(\bar{R})$ of the profit stream in stage II (assuming R&D pays off at all), plus the additional consumer surplus V^c generated in stage III by the cost savings from the invention. Then (assuming the same discount rate for consumers and producers):

$$V^c = \int_L^{\infty} q_0 (c_0 - c_1) e^{-rt} dt + \int_L^{\infty} \left(\int_{q_0}^{q_1} (p(q) - c_1) dq \right) e^{-rt} dt \quad (7)$$

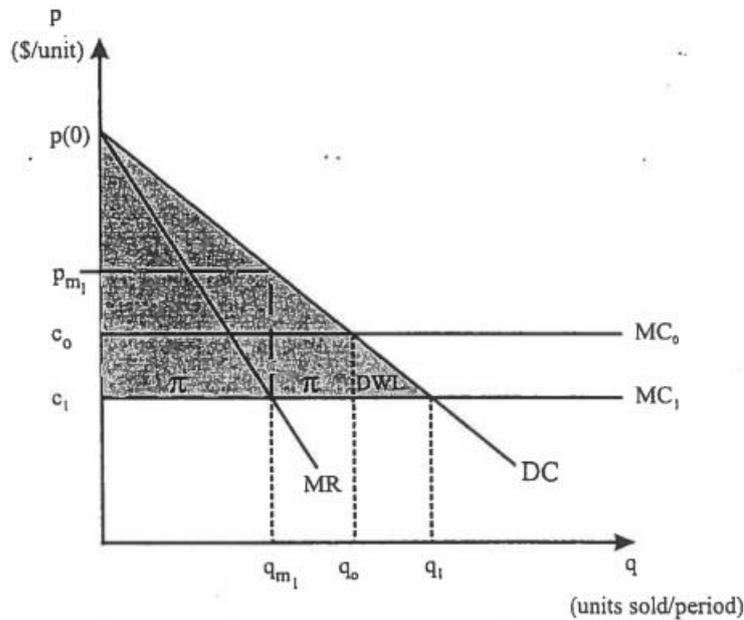
The second term then corresponds to the dead-weight loss in stage II being turned into consumer surplus in stage III when competition forces price down to c_1 . Now for $q(p)$ linear in p :

$$V^c = (c_0 - c_1)(q_0 + q_1)e^{-rL} / 2r \quad (8)$$

The producer's (innovator's) net surplus V^p , when $t_1 = 0$ and $t_2 = L$ and $R = \bar{R}$ is (from (5)):

$$V^p = q_0 c_0 k \bar{R}^\alpha (1 - e^{-rL}) / r - \bar{R} \quad \text{where (from (6))} \quad (9)$$

Figure 3.11 Value of patent protection of an innovation
 (simple case of a minor cost-reducing innovation with linear demand and constant marginal cost)



Legend:

- MC Marginal cost, being reduced from c_0 to c_1 by the innovation
- MR Marginal revenue
- DC Demand curve corresponding to $p = p(q) = -aq + b; a, b > 0$
- p Price per unit
- p_m Monopolistic price for the innovation (impossible to charge in this case since it exceeds c_0)
- p_c Competitive price = MC
- q Quantity sold per period. $q = q(p)$
- q_c Quantity sold under competitive pricing = q_0 and q_1 , corresponding to c_0 and c_1
- q_m Quantity sold under monopolistic pricing
- π Producer (gross) surplus per period under patent protection = $q_0(c_0 - c_1)$
- DWL Dead-weight loss = $\frac{1}{2}(q_1 - q_0)(c_0 - c_1)$, being lost to consumers under patent protection, allowing price to remain at c_0
- $\pi + \text{DWL}$ Consumer surplus per period after patent expiration and price drop to c_1

Figure 3.11 Value of patent protection of an innovation (simple case of a minor cost-reducing innovation with linear demand curve and constant marginal cost)

$$\bar{R} = (\alpha q_0 c_0 k (1 - e^{-rL}) / r)^{1/(1-\alpha)} \quad (10)$$

Thus (after some simplification):

$$V^p = (\alpha q_0 c_0 k (1 - e^{-rL}) / r)^{1/(1-\alpha)} (\alpha^{-1} - 1) = \bar{R} (1/\alpha - 1) \quad (11)$$

Thus total net surplus or welfare to society $V^s = V^c + V^p$ is a function in L explicitly. The first-order necessary condition for any $L = \bar{E} > 0$ to maximize V^s is then given by $\partial V^s / \partial L = 0$, so that a marginal change of \bar{E} gives equal but opposing changes in discounted consumer and producer net surplus, i.e. at \bar{E} a balancing trade-off is made between the innovator and the rest of society.

The calculation of \bar{E} is in the general case not a straightforward sequence of algebraic operations. It turns out that with $p = -aq + b$ (see Nordhaus (1969), p. 78):

$$\bar{E} = -\frac{1}{r} \cdot \ln \left(1 - \left[\frac{c_0 k \bar{R}(\bar{E})^\alpha + a q_0}{c_0 k \bar{R}(\bar{E})^\alpha \cdot (\alpha + 1) / 2\alpha + a q_0} \right] \right) \quad (12)$$

Using this expression one can analyze how \bar{E} varies with the different relevant parameters.

However a patent system does not come free of charge to society, so the costs to society for installing and operating a patent system should also be considered in addition to consumer and producer surplus. These costs include, first, the direct costs for setting up and operating patent offices and a system for enforcement of patent laws (courts, etc.). Second, there are the direct costs of the patenting operations for the inventors and inventor firms. These costs are neglected in the equations above but may amount to several percent of R&D costs. Third, there are various indirect costs as well as benefits, as shown in Table 3.3 above. Among the latter are the costs for the patent holder and the corresponding benefits to other inventors from the stipulated disclosure of information in return for patent rights. For a single patent these costs and benefits might be neglected (as in the Nordhaus model), but in total they are considerable. Just as a firm might decide that the optimal level of R&D investment is to conduct no R&D at all, the IP policy-maker might decide that the best policy is to have no

patent system at all, even if there is a positive length of patent protection time that gives a local maximum to V^s , the total discounted net surplus to society.

One can analyze various other cases of inventions, patent protection and consumer and producer behaviour both from the point of view of the inventor firm and the patent policy-maker. Some examples of such cases are given below to put the preceding analysis in context:

- 1) Major output-expanding inventions (in his original work Nordhaus also analyzed cost-reducing major inventions, which he called "drastic" inventions)
- 2) Price discrimination by the patent holder
- 3) Compulsory licensing (see e.g. Scherer (1977) and Tandon (1982), for economic analysis as well as Julian-Arnold (1992) and Merges (1996) for legal perspectives) and optimal licensing (see e.g. Gallini and Winter (1985), Shepard (1987) and Kamien et al. (1992))
- 4) Patent races in stage I (see e.g. Tirole (1988))
- 5) Other market structures in stages I and III than perfect competition
- 6) Other market structures than perfect monopoly in stage II (see e.g. DeBroch (1985))
- 7) Different shapes of market responses to a new product
- 8) Different price and output responses when patent protection ceases
- 9) Sequences of patents and innovations over time (see in particular works by S. Scotchmer, e.g. Scotchmer (1991), Green and Scotchmer (1995) and also Gallini (1992))
- 10) Clusters of interdependent patents and innovations
- 11) Welfare distribution considerations by differentiating the weights put on producer and consumer surpluses

- 12) Patent term restorations making up for increases in R&D time or time to market (see e.g. Grabowski and Vernon (1986))
- 13) Patent renewal fees (see e.g. Pakes (1986) and Cornelli and Schankerman (1995))
- 14) Optimal breadth or scope of a patent (see e.g. Klemperer (1990), Gallini (1992) and Gilbert and Shapiro (1990) for optimization models, the latter also considering jointly optimal breadth and length of a patent for a given size of the resulting reward to the inventor. For a thorough qualitative analysis, see Merges and Nelson (1990)).
- 15) Patent litigation (see e.g. Meurer (1989))
- 16) Patent valuation (see e.g. Schankerman (1991))

As can easily be imagined regarding technical inventions, the variety of behaviours and firm strategies, including both uses and abuses, which may be legally and economically inventive as well, make any comprehensive analysis exceedingly complex. As mentioned in the beginning of this section, practice has run ahead of theory as far as patents and IP are concerned. Considering the variety of firm behaviours and patent strategies exposed in the subsequent chapters, it is a safe bet to say that practice will continue to be ahead of theory for a long time, if not forever. This is not a reason for abandoning formal economic analysis altogether. On the contrary, in order for theory and practice to go hand in hand (though sometimes requiring long arms), theory has to catch up further. Apart from giving approximate guidelines in specific situations, one of several additional advantages of such analysis is that it exposes relevant assumptions and conditions, which in principle can be used to predict further practices before they arise.

A few model variants from the inventor firm's point of view will finally be outlined. First consider a genuine product invention with perfect patent protection. As the product diffuses on the market, the firm's sales S build up with a stream of operating profits π . These profits are then gradually competed away, as IP protection weakens and competition builds

up, e.g. through successful invent-around activities, secrecy leakages and expiration of patents.

Now assume that π is a constant fraction ρ of sales S , i.e. $\pi = \rho S$, as in the case where the product invention is licensed out with a constant royalty rate $\rho \cdot 100$ per cent, and let $S(t)$ correspond to a common pattern for diffusion and sales growth of innovations, e.g. $S(t) = \bar{S}(1 - e^{-\lambda t})$ where $\lambda > 0$ is a variable determining the rate of diffusion and \bar{S} is maximum sales. λ can be interpreted as the instantaneous probability that a consumer learns about the new product and adopts it. λ can be influenced by marketing investments and in particular by advertising, including the build-up of a trademark.

Assume moreover that ρ drops to zero after a period of L years of effective IP protection on the market, or the license contract expires after L years. (A gradual decline of ρ and reduction of sales after L years can also be modelled, but is left aside here.) Then:

$$V = \int_0^L \rho S(t) e^{-rt} dt - DI = (1/r) \cdot \rho \bar{S} (1 - e^{-rL}) - \frac{1}{(\lambda+r)} \rho \bar{S} (1 - e^{-(\lambda+r)L}) - DI \quad (13)$$

The second term is a loss of discounted profits due to the limited market-penetration rate λ (in the Nordhaus type of modelling $\lambda = \infty$.)

Now with a specification of how the market-penetration rate λ depends upon marketing investments M , the optimal M can be found, just as the optimal R can be determined from assuming an invention possibility function or an effort curve. Learning-curve reductions in variable unit costs could also be introduced into the model, which would correspond to a progressive royalty rate ρ .

Finally the value of secrets could be calculated. Suppose for simplicity that a cost-reducing invention, like the one in the Nordhaus model above, is guarded by a secret instead of a patent. The secret leaks out or diffuses in a manner that makes the secret-protection time uncertain. Thus, assume L is a random variable, e.g. exponentially distributed with parameter $\mu > 0$, i.e. the probability $P(L \leq x) = 1 - e^{-\mu x}$ and the expected value $E(L) = 1/\mu$. Thus, using

(3), the expected discounted value of the secret (disregarding the investments as already sunk at the time of evaluation of the secret) is:

$$E(V) = \pi/r \cdot (1-\mu/(\mu+r)) = \pi/(\mu+r) \quad (14)$$

This is equivalent to discounting with an added risk premium to the discounted rate. When $\mu \rightarrow 0$, the expected lifetime of the secret goes to infinity and $E(V)$ approaches π/r which equals the value of a patent with infinite lifetime. Now μ could be influenced by technology intelligence efforts made on both the innovator and the imitator sides (see Chapter 7). If the innovator's investment in secrecy protection is Y , one simple conceivable specification is $\mu = 1/(aY + b)$, where a and b are parameters, reflecting e.g. secrecy breaking efforts among competitors. If the product invention is covered by a patent and its process by a secret and the secret is still in force when the patent expires (an event with probability $e^{-\mu\bar{E}}$ of occurrence), its expected discounted value at time $t = \bar{E}$ conditional upon this event, is still $\pi/(\mu+r)$. Thus a secret under these assumptions, which is an "exponentially lived secret" (or a secret without 'ageing'), does not lose its value over time as long as the secret is not lost. Hence, an exponentially lived trade secret should not be depreciated. At the time the invention is made ($t=0$) and patent protection with length \bar{E} starts to cover the product and secrecy protection covers the process, the total expected discounted value to the inventor firm of the total protection is

$$EV = EV^{\text{pat}} + EV^{\text{sec}} - EDI \quad (15)$$

The value of the patent V^{pat} can be calculated as earlier, and it can be assumed to be unaffected by the secrecy protection. The secrecy, on the other hand, has a value only insofar as it gives additional protection after the patent has expired. Thus

$$EV^{\text{sec}} = E[V^{\text{sec}} | L > \bar{E}] \cdot P(L > \bar{E}) = (\pi \cdot e^{-\mu\bar{E}} / (\mu + r)) \cdot e^{-r\bar{E}} \quad (16)$$

The value of both patents, trade secrets and trademarks may be calculated in principle by using this modelling approach. Copyrights and utility models ("petty patents", "Gebrauchsmuster") can also be treated in this way in principle.

In summation, this section has illustrated, although briefly, how the economic value of the various IPRs can be calculated in principle, and also how their combined value through “multi-protection” (see Chapter 7) can be calculated at least in simple cases.

3.7 Summary and conclusions

A general framework for analyzing patents and intellectual property has been presented as a textbook introduction to the topic. (A complementary framework for analysing patent information will be described in Chapter 9.) Innovations are fundamental not only to economic development but to cultural development at large. Various types of innovations in general were described (technological, managerial, financial, legal etc.) together with concepts (invention, innovation, diffusion etc.) and models related more specifically to technological innovations. Thus the product life cycle model, the interactive innovation model and the buyer/seller-diffusion model were described.

In order to achieve economic development in a market economy it is crucial to ensure an adequate rate of innovation with an adequate rate of buyer diffusion producing an adequate rate of returns to innovators, buyers and society at large. Imitation in the form of seller diffusion and the resulting competition may hamper the rate of innovation, since innovators then may not be able to capture sufficient returns to cover their investments in innovation. At the same time seller diffusion may increase buyer diffusion and vice versa so that the resulting rate of returns to buyers and society for a given innovation may increase. Thus there is a problem when balancing or trading off the static efficiency for a given innovation and the dynamic efficiency for a stream of innovations. The patent system has been designed as an instrument or institution to deal with this balancing problem by offering restricted monopoly rights as an incentive to the innovator. The innovator can use these rights in turn to restrict seller diffusion and competition or to sell licenses on the patented technology in order to make sufficient profits (capture sufficient returns).

The chapter described the nature and functioning of the patent system and dealt with the issue of who can patent what as well as why, when, where and how to patent from the point of view of an individual or a company. A simple, qualitative model for the valuation and pricing of patents and licenses was presented as well. The pros and cons of a patent system from society's point of view were then discussed.

The chapter finally gave an overview of the growing literature on patents as well as an overview of the evolving economic theory of patents and IP. Some examples of economic modelling were also presented.