Juridiska institutionen
Höstterminen 2014

Examensarbete i immaterialrätt
30 högskolepoäng

Patent Scope:
- A Law and Economics Analysis

Författare: Henrik Malmsjö
Handledare: Professor Bengt Domeij
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>1.1</td>
<td>Background</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td>Aim and Purpose</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Method and Materials</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Disposition</td>
<td>6</td>
</tr>
<tr>
<td>2.</td>
<td>An Introduction to the Economic Analysis of Law</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>What is Law and Economics</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>Critique of Law and Economics</td>
<td>8</td>
</tr>
<tr>
<td>3.</td>
<td>Patents and Patent Scope</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Intellectual Property and Property</td>
<td>10</td>
</tr>
<tr>
<td>3.2</td>
<td>Patent Scope</td>
<td>10</td>
</tr>
<tr>
<td>4.</td>
<td>The Law and Economics of Patents</td>
<td>13</td>
</tr>
<tr>
<td>4.1</td>
<td>Cheap Replication, Expensive Innovation</td>
<td>13</td>
</tr>
<tr>
<td>4.2</td>
<td>Patents as Monopolies</td>
<td>14</td>
</tr>
<tr>
<td>4.3</td>
<td>From Theory to Practice</td>
<td>16</td>
</tr>
<tr>
<td>4.4</td>
<td>Optimal Patent Scope: A First Attempt</td>
<td>16</td>
</tr>
<tr>
<td>5.</td>
<td>Costs of Patents</td>
<td>19</td>
</tr>
<tr>
<td>5.1</td>
<td>Rent Seeking</td>
<td>19</td>
</tr>
<tr>
<td>5.2</td>
<td>Curtailing Rent-Seeking</td>
<td>20</td>
</tr>
<tr>
<td>6.</td>
<td>Alternative Functions of Patents</td>
<td>22</td>
</tr>
<tr>
<td>6.1</td>
<td>The Prospect Theory</td>
<td>22</td>
</tr>
<tr>
<td>6.2</td>
<td>Other Functions</td>
<td>25</td>
</tr>
<tr>
<td>7.</td>
<td>Applying the Cost Theory</td>
<td>27</td>
</tr>
<tr>
<td>7.1</td>
<td>Costs and the Current Patent System</td>
<td>27</td>
</tr>
<tr>
<td>7.2</td>
<td>Practical Problems</td>
<td>28</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Estimating Costs</td>
<td>28</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Probabilistic R&amp;D Outcomes</td>
<td>29</td>
</tr>
<tr>
<td>7.3</td>
<td>Applying the Theory</td>
<td>30</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Comviq and Europolitan</td>
<td>30</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Tele P and P-mint</td>
<td>32</td>
</tr>
<tr>
<td>7.3.3</td>
<td>Amgen and Hoechts</td>
<td>34</td>
</tr>
<tr>
<td>8.</td>
<td>Concluding Remarks</td>
<td>36</td>
</tr>
<tr>
<td>8.1</td>
<td>Possible Reforms</td>
<td>36</td>
</tr>
<tr>
<td>8.2</td>
<td>A Few Words of Caution</td>
<td>38</td>
</tr>
<tr>
<td>9.</td>
<td>References</td>
<td>40</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Background
A question of central importance in patent law is how wide the scope of patents ought to be. It is brought up in relation to individual patents in almost every single patent infringement suit. The patent holder will argue that the alleged infringer is practicing an invention which falls within the scope of his patent. The alleged infringer will often defend himself by arguing that the relevant product or process falls outside the scope of the patent.¹

Patent scope is also important from an economic perspective. Even a single patent, given a scope which is too wide, could stifle technological progress in an industry for decades. Conversely, if courts routinely assign patents an extremely narrow scope, there is a risk that patents would become a mere facade, providing no real protection for inventing firms.

1.2 Aim and Purpose
There are two main issues I will discuss in this essay. The first one is how wide the scope of patents ought to be according to economic theory, with a focus on the two classical theories; the cost theory and the prospect theory. The second issue is how the insights from the economic theory of patents can be applied by courts when determining patent scope.

1.3 Method and Materials
The method used in this essay goes under the name of law and economics. It consists of the application of microeconomics to the analysis of law. An introduction to law and economics is given in section 2.

Law and economics has traditionally been an American discipline and most of the law and economic literature and articles has been written by American scholars. The materials used in this essay is therefore mostly American.

¹ Not every instance of patent litigation will follow this pattern, the alleged infringer can for example argue that the patent does not satisfy the requirements for patentability (invalidity defense) or that the patent was obtained due to inequitable conduct during prosecution (unenforceability defense). (Ford (2013) pp. 74 & 78).
This does not mean that the conclusions are only applicable to American law. The conclusions are generalizable enough to be applicable to any patent system. Most cases discussed will be European in order to illustrate this point.

1.4 Disposition
Section 2 consists of an introduction to L&E. Section 3 consists of a brief discussion of how to think about patents and patent scope from a L&E perspective. Section 4 presents the classical economic analysis of the patent system, the cost theory, and the problem of deadweight costs caused by monopolization. Section 5 presents the second traditional cost of patents, rent-seeking, and how it can be minimized. Section 6 introduces and criticizes the classic competitor to the cost theory - Kitch's prospect theory of patents. Section 7 discusses various ways in which the law and economic insights can be applied by courts when determining patent scope. Section 8 concludes the essay.
2. An Introduction to the Economic Analysis of Law

2.1 What is Law and Economics?
L&E consists of the application of (micro)economic theory to the analysis of law. At the positive level, L&E uses the tools and models of microeconomics to analyze the incentives of different legal rules to determine their effects.\(^2\) For example, the punishment for a crime can be thought of as the price which a criminal has to (probabilistically) pay to engage in criminal activity, in the same sense that the price of apples is what an apple buyer has to pay for apples. Increasing the punishment for a crime would then, all things equal, lead to less crime – just as raising the price of apples leads to less apples being bought.

At the normative level, L&E is used to determine what kind of legal rules ought to exist. The criterion used for evaluating legal rules is economic efficiency.\(^3\) The process of determining which rules are efficient begins with calculating the "value" of different legal rules. To calculate the value of a legal rule, take everyone affected by it and ask how much they are willing to pay to have it enacted if they support it, or how much they are willing to pay to not have it enacted if they are against it. The sum of these private evaluations is called the social value. The efficient legal rule is the one which has the highest social value.\(^4\)

This is how things are done in theory. Very few works of L&E are however based on willingness to pay questionnaires.\(^5\) The way L&E is done in practice is by using economic models of human behavior to predict how people will react to changes in legal rules. The models include simplifying assumptions

---

\(^2\) Friedman (2001) p. 11.
\(^3\) Schäfer & Ott (2004) p. 3.
\(^4\) This is the efficiency criterion developed by Alfred Marshall, called "Marshall Efficiency." More commonly used is the Kaldor-Hicks efficiency criterion, which involves looking at potential changes in legal rules and a set of hypothetical side payments. The reason for preferring Marshall over Kaldor-Hicks is that it is easier to explain and most of the time leads to the same result. A notable exception is the economics of altruism (Friedman (1988) pp. 2 – 5), but this is not relevant for the discussion here.
about cognitive traits, adherence to axiological principles and preferences. The behavior of the agents in the model is then deduced by using the tools of microeconomics and the outcome evaluated using economic efficiency as the criterion.

For example, what happens if the government increases prison terms for violent crimes? A simple L&E argument goes like this: people make decisions by weighing the expected benefits and expected costs of actions. Increasing average prison terms is equivalent to increasing the expected cost of committing crimes. When the cost of an action increases, people tend to perform less of it. Therefore people will tend to commit fewer violent crimes.

It is important to note that the argument is extremely simplified. "Paying" for your crimes by going to prison is in many ways different from paying for your groceries at the grocery store. Violent criminals is also one of the groups which tend to deviate the most from the behavioral assumptions of economic theory.

The way to overcome these discrepancies between the model and reality is to add additional factors and more sophisticated assumptions to the model. However, the reason for using models in the first place, rather than just looking at the real world, is that it allows for systematization and the discovery of regularities. Even if discrepancies exist, a model can still be very useful, and a strong reason for using the assumptions of microeconomic theory is that they identify causal factors in human behavior which are both real and large.

2.2 Critique of Law and Economics

L&E has been very controversial since its inception. Two categories of criticism can be discerned: criticism of the methods used and criticism of the normative approach. The previous section discussed some of the former criticism, so the focus in this section will be on the latter. The controversy

---

6 Economic agents are assumed to form their beliefs in truth conductive ways ("Epistemic rationality"). (Huemer (2013) p. 187f).
7 For instance, their preferences are assumed to be transitive, reflexive and complete. (Schäfer & Ott (2004) p.53).
8 Usually economic agents are assumed to be selfish. (Huemer (2013) p. 189).
9 For instance, the school for crime syndrome, criminal stigmatization and human capital depreciation might lead to more crime in the longer run. (Avio (1998) p. 145).
concerns the usage of economic efficiency as the sole normative criterion. For why should law concern itself with economic efficiency as opposed to justice?

There are a few ways to respond to this criticism. First of all, there are many areas of law where appeals to justice are either not helpful or irrelevant. For example: should the term of a patent start at the day of application or the day the patent was granted? Appeals to justice seems to do little to resolve this question, for none of the alternatives seems more just than the other.

Another response is that justice and efficiency are not disjoint. Many times there is an overlap. What it means for a situation to be economically efficient is that preference satisfaction is maximized, given the resources at hand. This is typically a good thing.

One last point, which ties in to the first, is that the strength of the argument from justice varies between different areas of law. In criminal law it is many times a very reasonable one. A strong argument for making the prosecutor prove the guilt of an accused criminal "beyond a reasonable doubt" is that punishing innocent people is a great injustice.

The upshot is that economic efficiency is a reasonable criterion, as long as no strong arguments from justice exist. An area where this holds true to a large degree is patent law. Two of the most common contemporary arguments in favor of the patent system is that it creates incentives for innovation and disclosure.\(^{13}\) That is, the reason for why we have the patent system in the first place is to spur economic activity. It therefore seems reasonable to use economic efficiency as the normative criterion when analyzing patent law.

\(^{13}\) Machlup & Penrose (1958) p. 10.
3. Patents and Patent Scope

3.1 Intellectual Property and Property

As the name suggests, intellectual property law share many similarities with property law, and it is useful to think of patents as a property right. A property right has two central aspects: the right to use and the right to exclude.\(^{14}\) The owner of a piece of land has the right to use the land in certain ways. For example, he might have the right to build a house on the land or the right farm the land. He also has the right to exclude others from using the land in certain ways, for instance the right to prevent mining under it or driving over it. In a similar fashion, the owner of a patent has the right to exclude others from using his invention.\(^{15}\)

There are however some major differences between patents and more paradigmatic cases of property rights. A first difference is the type of object the property right concerns, namely, patents are on ideas, and not physical objects.\(^{16}\) Another difference is the method of acquisition. If you create a new physical object you will automatically become its owner. If you want to patent a new idea, you need to apply for a patent at the patent office. Furthermore, to be eligible for a patent, the inventor must show that the patent concerns patentable subject matter, is non-obvious, novel and possesses utility.\(^{17}\) Yet another difference concerns duration. Patents will usually expire 20 years after the filing date\(^{18}\), while ownership of a piece of land can theoretically be held in perpetuity.

3.2 Patent Scope

What all property rights have in common is that they have a certain scope. There are two relevant senses of “scope”. The first sense denotes the “bundle” of rights that the owner of some property is in possession of. An owner of a piece of land might have the right to build on his land, but not the right to farm on it. Similarly, a patent holder might have the right to prevent products

\(^{14}\) In other words, the exclusive authority to control a resource. (Alchian (“Property Rights”) 3rd paragraph).

\(^{15}\) A patent owner does however not have the affirmative right of using the invention. (Vermont (2006) p. 19).

\(^{16}\) Friedman (2001) p. 131.

\(^{17}\) Friedman (2001) p. 95.

manufactured using his invention from being imported, but not the right to prevent non-commercial use of the invention.

The second sense refers to the object of ownership and denotes what is actually owned. This is the sense used in this essay when discussing "patent scope." The scope of someone's ownership of a piece of land in this sense is the territory bounded by some geometric figure on a map. A patent can similarly be thought of as a figure drawn in an abstract "idea space". The area of this figure corresponds to the scope of the patent.

Determining the scope of a property right is done by the courts if it is disputed. In the case of real property, such a dispute can arise if a land owner claims someone has built something on his land. The court might resolve such a case by comparing the placement of the building to the relevant land registry map.

Things are often more difficult for patents. There is no equivalent "idea space" map to consult. The court therefore needs some other method of determining what the patent covers. The way this is done is by looking at the patent, or more specifically: the claims at the end of the patent. Here is an example of what a patent claim looks like:

1. A method of pest control comprising of aqueous dispersion of pyrogenically produced hydrophobic silica [...] .

This is what the inventor claims to have invented. To determine whether an infringement has taken place, the court will compare the claim(s) to the allegedly infringing device. The courts will however not blindly follow the inventors claims. The inventor has an incentive to claim more than he actually invented. As is the case with all ownership, more is better. A classic example is eighth claim in the Morse telegraph patent (US RE117 E). Morse claims to have invented "[...] the use of [...] electromagnetism, however developed, for making or printing intelligible characters, signs or letters at any distances [...].". The use of electromagnetism for this purpose had of course already been invented in the form of signal flags, semaphores and beacon fires. Friedman (2001) p. 133.

21 SE367934 T3 "Insecticides." Patents are often not as straighforward as this. For example, here is claim 1 from the patent SE429732 T3 "Sustained-release compositions containing cation exchange resins and polycarboxylic polymers": A sustained release ophthalmic pharmaceutical composition for controlling and lowering intraocular pressure, comprising a basic active which is effective to lower intraocular pressure, an anionic mucomimetic polymer, and a cation exchange resin.
22 The courts will however not blindly follow the inventors claims. The inventor has an incentive to claim more than he actually invented. As is the case with all ownership, more is better. A classic example is eighth claim in the Morse telegraph patent (US RE117 E). Morse claims to have invented "[...] the use of [...] electromagnetism, however developed, for making or printing intelligible characters, signs or letters at any distances [...].". The use of electromagnetism for this purpose had of course already been invented in the form of signal flags, semaphores and beacon fires. Friedman (2001) p. 133.
then it is infringing on the patent.
4. The Law and Economics of Patents

4.1 Cheap Replication, Expensive Innovation

Why do we have a patent system? A popular answer is that it creates incentives for innovation. And while innovation is good, it is not clear that more innovation is better. Economic theory tells us that there can be too much of a good thing. This is because more resources devoted to some area of the economy will necessarily drain resources from other areas.

There is however an economic argument for why a patent system might be efficient: Research and development ("R&D") is socially beneficial and involves large costs. Firms investing in R&D will not be able to recoup their costs since other firms will just copy their results. Since it is not possible to recoup the costs, no firms will incur them in the first place. The result is an inefficiently low level of R&D.

For example: suppose Company A can spend €100 million to develop a more efficient manufacturing process. The process will save them €200 million over a number of years. However, once developed it will immediately be copied by their competitors (at zero cost). Company A will be at a competitive disadvantage after this, having had to spend resources which the replicating firms did not. Company A therefore decides against investing in the development of the process.

The solution is to give Company A a mechanism of recouping the costs it has incurred. This is the principal function of the patent system, according to the cost theory. Once Company A has been granted a patent, it can exclude other firms from using the improvement. This enables Company A to recoup its costs in one of two ways, by licensing the invention or being the only firm being able

---

23 This is one of the four traditional arguments for patents identified in Machlup and Penrose (1950) pp. 10 - 11. The other three are: natural rights in ideas, just reward for inventors and incentives for disclosure.
24 In the case of industrial research and development, the allocated human capital could alternatively be used in the production of consumer or capital goods, basic research or education. (Machlup (1958) p. 1321.)
26 In a perfectly competitive market, the inventor would be driven out of business because he would need to charge a price high enough to recoup his fixed costs, while the competitors would just charge a price equivalent to their marginal costs. (Guellec & Potterie (2007) p. 50)
to use it.

Solving the problem this way is not perfect. The invention is a non-rivalrous resource. This means that the marginal cost of using it is zero, because additional users does not debase the value of the invention. For instance, if Company A had five competitors, the social benefits of allowing them to use the improvement would be €1 billion (five times the €200 million in savings) while the costs of inventing would still be €100 million. According to economic logic then, it seems that the efficient price for an owner of a non-rivalrous resource to charge others for using it is zero.

This seems paradoxical: allowing the inventor to exclude others from using the invention is efficient, but so is allowing the competitors to use it. This analytical tension can however be resolved by comparing the situation ex ante versus ex post. Ex ante, before the invention has been made, the efficient rule is that inventors should be given the right to exclude others from using the invention. However, once the invention has actually been made, as many people as possible should be allowed to use it. The efficient patent policy, then, is one where an optimal balance between these two ends is achieved.

4.2 Patents as Monopolies

To explore the problem of patents more thoroughly, it is useful to think of a patent as a form of monopoly. The economic definition of a monopoly is a firm which is the only provider of a product for which there are no close substitutes. As the only provider they have a wide discretion with regards to the pricing their product.

The problem with monopolies is however not that they overcharge their customers per se. While consumers might become poorer because of monopoly pricing, this loss is matched by a gain for the monopoly firm.

The problem is rather the dead-weight losses caused by the goods which

---

29 This ability to raise the price above the marginal cost of production and still profit is called the market power. It can be argued that patents do not create monopolies since few patent confer much market power. However, the inefficiencies created by patents and monopolies are the same.
are not sold due to the increase in prices.\textsuperscript{30} The image below show the effects of monopoly.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{monopoly_diagram.png}
\caption{Comparative effects of a competitive market versus a monopoly.}
\end{figure}

In a competitive market, firms will produce and consumers will buy until the marginal cost of production equals the marginal utility of consumption. This is also the level at which the price of the good will be set, since all units are priced at the same level. The producer surplus is shown in green and the consumer surplus in orange. In a monopoly market, the firm will set its price at the level which maximizes the producer surplus. However at this (higher than equilibrium) level there will be deadweight losses (shown in gray) due to fewer units being sold.

Deadweight losses are caused by patents in a similar fashion, as the image below illustrates (notice that the supply curve is identical to the x-axis).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{patent_effect.png}
\caption{Effect of patents on market outcomes.}
\end{figure}

\textsuperscript{30} Arnold (2010) p. 239.
Since the invention is non-rivalrous, the cost of providing it is zero. This is also the price at which combined consumer and producer surplus is maximized. If the invention is covered by a patent, the patent holder will set a non-zero price for using the invention. This will lead to deadweight losses, due to some potential users deciding against using the invention.

4.3 From Theory to Practice
How does patent scope fit into all of this? It is one of the methods by which this balancing of ends can be achieved. Widening the scope leads to a more valuable patent. This will increase the incentives to invent, but also the deadweight losses. Conversely, narrowing the scope will decrease the incentives to invent while at the same time decreasing the deadweight losses.

While the economic argument is sound, it does not seem overly helpful. Suppose a judge presiding over a patent infringement suit tries to apply it. In order to do so he first of all needs to know how large the monopoly losses would be from siding with the plaintiff. Furthermore he needs to know how his decision will affect the incentives to invent. Determining either factors to a reasonable degree of accuracy seems hopeless.

This is a very common situation in L&E. There is an efficient rule which is fairly simple but is very difficult to implement because real life agents do not possess enough knowledge. It is however not a reason for completely abandoning the rule. As will be shown below, there are situations where it can be applied. Furthermore other relevant factors, which are easier to quantify, will be identified.

4.4 Optimal Patent Scope: A First Attempt
In order to determine the efficient scope of patents, let us look a bit closer at how firms make investment decisions. Economic agents are assumed to make a decision to invest when the private benefits (Bp) exceed the private costs (Cp). Put in another way, a decision will be taken if and only if Bp – Cp > 0. When determining whether the decision is efficient, one also needs to look at

31 For instance, the L&E criteria for if a person has been negligent is determined by whether he has taken all cost justified precautions or not. (Friedman (2001) p. 198) This can prove difficult to determine in practice.
32 This is the proverbial cost-benefit principle. (Bernanke & Frank (2008) p. 3).
the social costs and social benefits. In the case of R&D, the costs are still private but parts of the benefits are social.\textsuperscript{33} Furthermore, the inventor will not always be able to capture all the social benefits of the invention. If these things are considered, the inequality will look like this: \( A^*B_s - C_p > 0 \). \( B_s \) represents the social benefits of the invention. "\( A^* \) is the appropriability factor, the fraction of the social benefits which the inventor can capture.\textsuperscript{34} Widening the patent scope enables the inventor to capture a larger fraction of the social value of the invention. Widening the scope is, in other words, equivalent to increasing the appropriability factor.

At a first pass, the efficient policy might seem to be to construe the scope so that \( A \) equals 1. This lets the inventor capture all the social value he creates (or as much as possible).\textsuperscript{35} The positive externalities of the invention will then be internalized to the highest degree possible. If \( A \) was set lower, there would be social benefits an inventing firm could not capture and hence will not take into account when making the decision of whether to invent or not. The result would be underinvestment in R&D.

The problem with setting the appropriability factor to 1 is that, due to the deadweight losses from monopoly, each increase in the benefits to the inventor is matched by a much larger social loss.\textsuperscript{36} Increasing the appropriability factor will, in other words, decrease the social benefits of the invention.

The efficient policy is instead to set the appropriability factor at a different level for each invention. It should be as small as possible, while still being large enough to offset the costs of R&D.\textsuperscript{37} With this policy, the optimal level of invention will be achieved and the deadweight losses will be minimized.\textsuperscript{38} For innovations with a high cost, the appropriability factor should be set at a high level, since there are more costs for the firm to recoup. For inventions with a low cost the appropriability factor should conversely be set at a low level. Since modifying patent scope is the way in which the appropriability factor is

\begin{itemize}
  \item \textsuperscript{33} Tabarrok (2002) p. 2.
  \item \textsuperscript{34} Note that \( A^*B_s \) is equal to the private benefits of the invention.
  \item \textsuperscript{35} Scotchmer (1991) p. 31.
  \item \textsuperscript{36} Tabarrok (2002) p. 3.
  \item \textsuperscript{37} Notice that if \( A^*B_s = C_p \) then the decision might not be made since the inventor will be indifferent to inventing. (Boldrin & Levine (2001) p. 178).
  \item \textsuperscript{38} Tabarrok (2002) pp. 2f.
\end{itemize}
set, this is equivalent to saying that high-cost patents should have a wide scope and low-cost patents should have a narrow scope.

However, patent scope is not the only determinant of the appropriability factor. Another important determinant is the costs of replicating the invention. To illustrate this point, let us return to the Company A example in part 3.1. The reason why Company A decided against inventing the process (without patent protection) is that inventing it would have put it at a competitive disadvantage. However, the higher the cost of replicating the invention is, the smaller Company A’s competitive disadvantage becomes. In other words, the higher the replication costs, the less wide the patent scope needs to be for Company A to invent.

The upshot of the cost theory is this: efficient patent scope is one which enables the inventor to recoup the costs of inventing, but nothing more. The patent scope should be wide for inventions with a high invention to replication cost and narrow for inventions with where invention to replication costs are low.

An application of the economic argument is that, holding the costs of invention and replication constant, patent scope should be wider for inventions with low social benefits and narrower for inventions for which the social benefits are high. Suppose that the cost of inventing a product is €100 million, the social benefits are €300 million and the costs of replication is zero. The firm which invents the product is granted a patent. How wide should the scope of the patent be? The answer is that it should be wide enough for the firm to recover the costs of inventing. In this case, this amounts to €100 million or one third of the value of the invention. However, even if the social benefits only amounted to €150 million, the firm should still be allowed to appropriate €100 million or two thirds of the value. In other words, the lower the social benefits of the patent are, the higher the appropriability factor should be set.

There are more applications of the cost theory. Before delving deeper into them, I will discuss the second classic cost of patents: rent seeking.

39 The need for patent protection diminishes the higher the costs of replication are. If the replication cost is equivalent to the cost of replication, no patent protection is necessary. 40 Tabarrok (2002) p. 3.
5. Costs of Patents

5.1 Rent Seeking

In the arguments presented above, there was only one firm conducting R&D. The R&D was also assumed to be instantaneous and the potential results were known with certainty. Removing these assumptions enables us to illuminate another social cost of the patent system – rent seeking.

Rent seeking is socially inefficient behavior which can occur when a group of people or firms are competing for a prize, if they can increase the probability of winning by expediting more resources.\(^\text{41}\) The inefficiency manifests itself as a collective tendency to spend too much resources. An example of a potentially rent-seeking activity is the "race" to do enough R&D to qualify for a patent.

Suppose two firms are competing for a patent which has a private value of €400 million. The companies can chose between three R&D activity levels. If a firm spends €10 million ('low activity') on R&D it will be eligible for the patent after 10 months. At €20 million ('medium activity') it will take 8 months and at €100 million('high activity') it will take 6 months. Each company also needs to invest €20 million in basic research before they can start the product R&D. If both companies finish their R&D during the same month, a coin will be flipped to determine who gets the patents.

Had there only been one firm, it would have picked the optimal activity level: €10 million per month. At this level, the R&D is finished after 10 months at a total cost of €30 million (€20 million in basic research + €10 million in R&D). Returning to the two firm situation: if both of them picks the optimal activity level, the total R&D expenditures incurred will have totaled €60 million by the time they are done.

However, things have a tendency to become more much more costly. If both firms chose the low activity level, they each have a 50 percent chance of getting the patent.\(^\text{42}\) But if one firm selects the medium activity level while the other firm sticks with the low, that firm will have a guaranteed win, finishing their

\(^{42}\) They will finish the same month and things will be determined by the coinflip.
R&D two months before the competitor. Realizing this, both firms decide to opt for the medium activity level. Each firm will again have a 50 percent chance of winning the race. Before the race starts however, they both realize that the other firm will choose the medium activity level. Since this is the case, it is better to choose the high activity level which is a guaranteed win against the medium. Both firms therefore opt for the high activity level.

The outcome: one of the firms wins the race after 6 months, during which a combined total of €240 million has been spent on basic research and R&D. Comparing the one firm and two firm situations, it is clear that the two firm race is inefficient. In both cases society ended up with a new invention, but in the second case the cost was eight times greater.

The example points at two reasons for why rent-seeking is inefficient. First, firms will spend too much money on R&D. There is a point of diminishing marginal returns for all beneficial activities. At some point, doing more of any activity will cost more than it is worth. In rent-seeking situations, the large prize for the winner will skew the private cost/benefit analysis. The parties will go beyond the point where the marginal social benefits equals the marginal social costs. The second inefficiency stem from a duplication of efforts. The costs incurred by the losing firm will have been in vain, since they will not be able to practice the invention and the invention would have been invented anyway.

5.2 Curtailing Rent-Seeking
There are ways of curtailing rent-seeking patent races. These include decreasing the 'length' of the race, decreasing the size of the price, decreasing the number of races and changing when the race starts.

Some patent rules have the effect of minimizing rent-seeking. For instance, the patentability requirement of utility has the effect of delaying patent-races until a technology has been adequately described or understood.43 Similarly, the non-obviousness requirement can be thought of as restricting patent-races to cases where the invention is beneficial.44 Furthermore, the fact that patents are granted before commercialization45 means that patent-races are shortened.

The way in which patent scope can minimize rent seeking is by decreasing the size of the prize. A narrower scope leads to, all things equal, a less valuable patent, which in turn means less rent-seeking. It seems that the patent scope offered by the cost theory will in many instances have the beneficial side effect of minimizing rent-seeking. If the patent scope is narrower than the cost theory suggests, no firm will enter the race in the first place. If it is wider, firms will spend too much on R&D. However if the patent scope is at the width which the cost theory suggests, the race will get started and as few resources as possible will be wasted. It therefore seems that the cost theory has the additional benefit of limiting resource-wasting patent races.
6. Alternative Functions of Patents

6.1 The Prospect Theory

In his seminal paper titled "The Nature and Function of the Patent System"\(^{46}\), Edmund Kitch offered a different view on the patent system than the one presented above. On this view, it is sometimes efficient for the government to reward inventors with prospect patents. A prospect patent has two properties. It is awarded very soon after a technological discovery has been made and it has a very wide scope.\(^{47}\)

Awarding prospect patents, Kitch argues, has three benefits: it reduces rent-seeking, discourages free-riding and enables the patent holder to co-ordinate the "exploitation" of the patented idea.

Kitch presents the prospect theory by using an analogy with the U.S mineral claims system.\(^{48}\) Under this system, a prospector who found minerals on a piece of land could get the mineral rights by filing a claim with the government. Very little prospecting was required to get the mineral rights to a piece of land. This had the potential effect of reducing rent-seeking. Compare for instance a rule requiring commercial significance and a rule only requiring actual discovery of minerals. Under commercial significance rule there might be five prospectors drilling in parallel to determine the size of a deposit, while under the actual discovery rule the race is cut off once the first prospector finds minerals on the surface.

The scope of mineral right was also set to be wide enough for the prospector to successfully exploit his find. For if the scope of the mineral rights were to narrow, another prospector could just drill on an adjacent plot of land and free ride off the first prospector's discovery. This would decrease the incentives to find minerals.

The mineral rights also had the benefit of allowing the prospector controlling them to coordinate the efforts of others. Suppose that a prospector is given the mineral rights to a large area of land. Another prospector approaches

\(^{46}\) Kitch (1977).
\(^{47}\) Ibid., p. 266 ff.
\(^{48}\) Ibid., p. 274 – 278.
him wants to take soil samples from an area which he believes contains a high concentration of minerals. They make an agreement to share the mining profits if he is correct. However, it turns out he is wrong and no mining takes place. Later, another prospector who also believes that the same area have a high concentration of minerals approaches him with the same proposal. Knowing that there are no minerals, the mineral rights holding prospector declines the offer.

This is the efficient outcome since there are no benefits of looking for minerals where there are none. Had the original prospector not possessed the mineral rights, he would not have had any incentive to coordinate the search in this way. Had there been no mineral rights, there is a risk that soil samples would be taken from the same area multiple times, because the prospectors would have no incentives to tell others about the results of their prospecting. This ability to coordinate efforts is what Kitch calls the prospect function.

Kitch then returns to the patent system and argues that the same points apply there. Awarding patents early will reduce the costs of rent seeking, granting very wide patents will reduce free-riding and the prospect function enables the patent holder to coordinate the development of the patent.

How well does this fit with the cost theory? It is first of all important to notice that there is a difference between prospect patents and the prospect function of patents. The practice of awarding prospect patents can be in line with the cost theory. It is possible that awarding patents early will reduce the costs of rent-seeking. Making patents wide enough to prevent free-riding can also be an efficient practice according to the cost theory.

The prospect function is however something which the theory ignores. The prospect function is an ex post benefit, but only ex ante benefits exists according to the cost theory. If the prospect function is real, it must be taken into account when the determining patent scope.

Let us look at how the prospect function for a patent is purported to work. Suppose that after the invention of the automobile a very wide patent covering all "four-wheeled vehicles using an internal combustion engine" is granted. The patent holder is now in position to coordinate the development of new uses and
improvements for the automobile. Should he want a better engine, he can license the development to another firm. Should a firm have an idea for a new car, say one which transports goods instead of people, a license agreement can be reached for such development. Since it is unlikely that others would try to find new improvements without first getting a license, the inventor will have knowledge of all failed research projects. This will help him coordinate more efficiently.  

If it turns out that the social benefits of the prospect function exceeds the social costs, then it is a factor which needs to be accounted for when determining optimal patent scope. There are however a few reasons for believing this is not the case. First of all, the patent will still create deadweight costs. Kitch’s reply is that not all patents will confer strong monopoly powers, because the patent owner will face competition from other processes or products. While it is correct in that most patents do not confer very strong monopoly powers, the patents which do confer strong monopoly powers (and thus lead to large deadweight losses) are generally the ones where the prospect function is the strongest. Furthermore, monopolies often have a reduced incentive to invent. In general, a firm needs to innovate for two reasons: increased profits and retaining their position in the market. A monopoly firm which has attained its market position through superior efficiency will be motivated to invent for both reasons. For a legal monopoly however, the second reason to innovate becomes less salient.

49 It should be noted that this example is not fictional. George Selden was granted a patent in 1895 which claimed the basic automobile configuration (US549160 A). The outcome was however not what the prospect theory would predict. Rather than coordinate development of new uses and improvements, Selden was more interested in collecting royalties from car manufacturers. (Merges & Nelson (1990) p. 889).

50 Kitch (1977) p. 274.

51 Burk & Lemley (2011) p.71. Patents need to confer some monopoly power to be useful.

52 A wide patent scope is almost per definition required for the prospect function to work well. The less covered by the patent, the less there will be to “prospect.”


54 This is the Red Queen effect; firms in a competitive industry are pressured by their competitors to invent and improve as fast as they can just to retain their standing in relationship to the competitors. (Dreyfus et al (2008) p. 61 – 62).

55 The fact that firms often have to increase their performance just to retain their current position is called “the Red Queen effect.” (Dreyfus. (2008) p. 61) The constestable market hypothesis suggests that this can hold true even if the firm is a pure monopoly, as long as the cost of entering and exiting the market is relatively low. (Arnold (2010) p. 499).

An additional problem is that high transaction costs can severely diminish the prospect function since it relies on contracts as the method of coordination. A firm conducting research in an area covered by a patent must at some point negotiate with the patent holder. If this negotiation is done after substantial research has been conducted, there is a hold-up risk. However, a firm might not be able to knowledgeably negotiate with the patent holder before exploring the prospect themselves.

Finally, the prospect function does not actually seem to work as well in real life as Kitch imagined. The prospect patents granted on the automobile, the light bulb, the microprocessor and the airplane led to litigation and restricted access rather than coordinated development.

In conclusion, the prospect function is at best a minor auxiliary benefit of the patent system. In most areas, especially where the innovation process is complementary, it is very possible that it does more harm than good.

6.2 Other Functions
There are other economic theories proposing alternative benefits of the patent system. These include the disclosure theory (patents induce disclosure of information), the earlier innovation theory (patents gives us inventions earlier), the speed of innovation theory (patents speed up the rate of innovation), the transaction cost theory (patents reduce private transaction costs) and the signaling theory (patents allow firms to signal the strength of

58 Patent Breadth and Sequential innovation 453 This problem will always exist but is larger for patents with large prospects.
59 It should be noted that the Selden automobile patent was a "submarine" patent. Selden originally applied for the patent in 1879, but managed to delay the publication until 1895. However, according to Kitch's theory this should not matter, since the important thing is having someone coordinating.
60 Gilbert Hyatt was granted a patent (US4942516 A) on the basic microprocessor configuration in 1990. The only "prospecting" taking place was Hyatt partnering with Philips to collect licence fees from the rest of the industry. (Ford (2013) p. 116).
63 Machlup & Penrose (1950) p. 25.
64 Chiang (2008) p. 3.
These theories will not be discussed here, for unlike the prospect theory they do not directly contradict the cost theory.\textsuperscript{68} Furthermore, some of the theories do not have much relevance for determining patent scope.\textsuperscript{69} Rather, they show that the cost theory is not a complete account of why the existence of a patent system might be economically efficient.

\textsuperscript{68} The prospect theory implies that almost all patents should be given a broad scope. (Oddi (1995) pp. 286 - 288).  
\textsuperscript{69} For instance, the signalling theory states that an auxilliary benefit of the patent system is that firms can signal such attributes as knowledge capital through their patent portfolio (Long (2002) p. 627). This ability is not affected by the scope of patents. Even if patents were made unenforcable, the signaling benefit would be the same.
7. Applying the Cost Theory

7.1 Costs and the Current Patent System

One of the main takeaways from the preceding discussion is that one of the most important functions of a patent is the recovery of costs. How well does current the current patent system reflect this? In some areas the cost recovery function is taken into account, while in others it is largely ignored. An example of the former is the ability in some jurisdictions to extend the length of pharmaceutical patents\(^70\). The justification is that the mandatory government approval for new drugs process leads to a delayed commercialization. The time window for pharmaceutical companies to recoup investments is therefore often very short. Since pharmaceutical R&D is often very expensive\(^71\), the possibility to extend the length of such patents was created.

Similarly, the United States patent system had a general patent extension rule in the 1800's which allowed an inventor to get a seven year extension on their patent if they could show that their profits had not covered the R&D costs.\(^72\)

There is also a number of American cases where the Court of Appeals explicitly took the cost recovery function into account. One example is Eli Lilly v. Generix\(^73\). Eli Lilly held a patent on a synthetic morphine. The alleged infringers argued that the patent should be limited to a use or process patent due to the existence of prior art. The court rejected this argument, writing that: "Such a niggardly patent reward for costly and painstaking research would discourage both the inspiration-perspiration process of the laboratory […]"

Another case is Panduit v. Dennison.\(^74\) Panduit claimed that Dennison had infringed on a number of their cable tie patents. Dennison argued that the patents were invalid on the ground of obviousness. The court rejected this argument partly due to

---

\(^70\) Legislation creating such "supplementary protection" was passed in the EU in 1992 (Regulation 1768/92). The U.S. Passed the Hatch-Waxman Act in 1984, which extends patent length for all pharmaceuticals.

\(^71\) Prop 1993/94:22 p. 35: "The costs of developing a new drug has been calculated to be on average around one billion crowns, and it is almost impossible to determine beforehand if the invested resources will lead to an acceptable product."

\(^72\) Tabarrok (2011) ch 9, para 16.

\(^73\) Eli Lilly and Company v. Generix Drug Sales Inc. (1971).

\(^74\) Dennison Mfg. Co. v. Panduit Corp (1986).
the fact that Panduit had invested millions of dollars in developing the cable ties.

When determining patent scope, however, the cost recovery function has largely been ignored.\(^\text{75}\) Rather, courts chiefly rely on examining the language of the claims, common sense, and the technical aspects of the invention to determine the extent of the scope.\(^\text{76}\) It is therefore very likely that the current way of determining patent scope is inefficient. In order to align patent scope decisions more closely with what is economic efficient, costs needs to be taken into account.

### 7.2 Practical Problems

One objection to using costs as a determinant of patent scope has to do with knowledge – or rather our lack of it. If we lived in a world of perfect information, the courts would have no problems applying the cost theory and determining the efficient patent scope. This is however not the case, which leads to two practical obstacles for applying the cost theory: estimating costs and taking R&D failure into account.

#### 7.2.1 Estimating Costs

The ability to use costs as a factor when determining patent scope relies on whether or not courts can estimate them accurately enough. If it turns out to be too difficult, the cost theory has to be abandoned.\(^\text{77}\) There are however reasons for believing it can be done.

First of all, courts do not need an exact value to take costs into account. Costs can be a relevant factor as long as courts are able to determine if they are low, medium or high.\(^\text{78}\) Furthermore, firms have an incentive to collect information about R&D expenditures for many reasons: tax benefits, accounting, project evaluation and for use in licensing negotiations.\(^\text{79}\) Finally, the invention itself often gives information about its associated costs. For instance, while a pharmaceutical company could plausibly argue that the R&D of a new drug have cost them hundreds of millions of dollars, the same cannot be said

---

\(^\text{75}\) Tabarrok (2011) ch 9, para 16.  
\(^\text{78}\) Tabarrok (2011) ch 9, para 16.  
for the Page-up/Page-down patent.\textsuperscript{80}

Similarly, a firm allegedly infringing on a patent should often be able to show what costs they have incurred. They have the same incentives to collect information about their R&D expenditures.

This is not to say that estimating costs is always easy. An implicit assumption made so far is that all research is conducted in research projects and that these will either succeed and yield exactly one patentable invention or fail and yield nothing. It is of course possible that a research project will yield several patentable inventions. It is also possible for a patent to be attributable to a several research projects. So even if it was possible to make very accurate estimations of research project expenditures, it does not necessarily follow that it is possible to estimate the cost of R&D for a given patent.

This is perhaps the biggest hurdle to overcome when attempting to apply the cost theory and does limit the extent of its application. The limitations to the applicability of the theory are discussed in section 8.2.

7.2.2 Probabilistic R&D Outcomes
An assumption made so far is that R&D projects have a 100 percent success rate. In real life this is far from the case. Applying the cost theory without taking this into account will lead to a "heads you break-even, tails you lose" outcome for firms investing in R&D.\textsuperscript{81} Suppose a firm engages in an R&D project which costs €10 million and has a 5 percent chance of success. A reward of €10 million upon success is analogous to a lottery where the winner gets his money back. This policy will clearly provide inadequate incentives for invention. In order to provide an adequate incentive, the reward must be a multiple of the R&D costs.

How large does this multiplier need to be? One way to think about it is by returning to the lottery analogy. In general, the multiplier required for the expected value of a lottery ticket to be positive is equal to the reciprocal of the risk. If the cost of a lottery ticket is €10 million and the chance of winning is 5

\textsuperscript{80} The patent (US7415666 B2), was applied for in 2005 and claims the method of scrolling up or down a single page in an electronic document by pressing a button. This is the function of the page up and page down button which exists on most keyboards.
\textsuperscript{81} Scotchmer (1991) p. 30.
percent, then the lowest payout a rational player would accept is €200 million.

This means that in order to determine the required multiplier for a firm to invest in an R&D project, one needs to know the chance of a successful project. Using records from the Small Business Innovation Research program, Kingston has tried to make such an estimate.\textsuperscript{82} Under the program, small firms can apply for a R&D grant from the U.S. government. The program has three phases, for which grants are given for the first two. The first phase consist of an exploration of how technically feasible the project is. The second phase consists of further R&D and evaluation of the potential for commercialization. The final phase consists of moving the product to the market.

The program shows that R&D is a risky venture. Out of 23,000 applications, only around 200 reached the market.\textsuperscript{83} One ninth of the projects failed to reach the first stage, half of those that did failed to reach the second and only a sixth of those projects reached the first stage. Using the failure rates and weighing the size of the investments at each stage, Kingston's conclusion the optimal multiplier is between 4 – 8, depending on where in the innovation process the firm is.\textsuperscript{84}

Tabarrok reaches a somewhat similar conclusion. He estimates that if there are N firms racing for a patent, then the multiplier for the winner should be roughly N.\textsuperscript{85}

\section*{7.3 Applying the Theory}
In this section I will illustrate what the application of the cost theory to an actual patent infringement case can look like.

\subsection*{7.3.1 Comviq and Europolitan}
The first case I will discuss is NJA 2000 s. 497. The parties were two telecommunications companies, Comviq and Europolitan. Comviq claimed that Europolitan infringed on one of their patents. The basic idea of the patented

\textsuperscript{82} Kingston (1994) p. 666ff.
\textsuperscript{83} And reaching the market does not mean that the project will survive.
\textsuperscript{84} The optimal multiplier shrinks as the project progresses, however the total R&D costs are larger at this point.
\textsuperscript{85} If a massive number of firms are racing, this multiplier will diverge from Kingston. This effect is however mitigated by Tabarrok's suggestion to scale down the multiplier when many firms are racing to reduce overinvestment. (Tabarrok (2002) p. 19).
invention was that a mobile phone user would be allocated two identities (IMSI numbers). The user could then select which of the identities he wanted to be active. For example, a person using a GSM mobile phone would revive a SIM card which had been allocated two phone numbers. The user would select which number was to be designated as “active” by entering a code in the mobile phone. One suggested use of the invention was having a combined business and private phone. When the user came to work, he would turn on the work number and when he left he would turn on his private number.

The allegedly infringing action was the sale of a product called “Twin Cards”. The product consisted of two different SIM cards, one full size and one mini. Both of these cards were assigned the same phone number. The user could use either of them for calling, but could only revive calls with the SIM card currently designated as “active.” Activation was done by entering a code into the phone. The reason for creating “Twin Cards” was increased compatibility. Some phones, such as car phones, requires a full size SIM card to work, while most mobile phones requires the smaller mini SIM. Giving the user one of each enabled him to use more types of phones.

The question of whether or not infringement had occurred hinged on how the expression “a subscriber identity module (SIM)” in the first claim was to be interpreted. Comviq argued that a system utilizing multiple physical SIM cards fell under the extension of “a subscriber identity module (SIM)”. Europolitan, on the other hand, argued that the claim only covered systems using one physical unit, for instance a two-sided SIM card.

The Supreme Court ultimately sided with Europolitan’s interpretation. It was specified in the the patent summary that the system was for “telephone systems, especially of the GSM type”. Furthermore, it was written in the description that no description of the used concepts were required since they were the subject of extensive standardisation (this also refered to the GSM standard). The Supreme Court decided that ”a subscriber identity module (SIM)” should therefore be given the definition it had been given in the GSM standard. According to that standard, a ”subscriber identity module (SIM)” refers to a single physical SIM card.

As can be seen, the deciding factor in the case was how a term in the
patent claims was to be interpreted. This is a common situation in patent infringement cases. This creates a problem when trying to determine what the result would have been if the cost theory had been applied to the case. Namely, the factors which are relevant for making such a determination are ignored by the courts and the parties. Nowhere in the case are R&D costs, replication costs or alternative means of appropriation mentioned. Without these factors, it can be difficult to tell which outcome is more efficient than the other.

As mentioned above however, the invention itself will sometimes give information about its associated R&D cost. In this case it seems that these costs were low. What had been invented was not the SIM card nor an actual method of allocating two IMSI numbers to a SIM card, but rather the practice of allocating two IMSI numbers to a SIM card. It might of course have been a really clever idea, but the function of the patent system from an economic point of view is not to reward firms for having clever ideas, but rather allowing them to recover R&D costs.

Furthermore, it seems that cost of replication is comparatively high. The patent does not describe, for example, how to store two IMSI numbers on a SIM card nor how to build two-sided SIM cards. A competitor who wants to copy the system needs to spend money implementing it, and it seems that the patent itself is of very little help when doing so. These factors point in the direction of a narrower patent scope. Thus the Supreme Courts decision seems to be in line with what the cost theory would suggest.

7.3.2 Tele P and P-mint
The second case I will discuss was decided by a Swedish Court of Appeal in 2006. The parties were two companies called Tele P and P-mint. Tele P claimed that P-mint had been infringing on a patent named “Parking System.” The invention claimed in the patent was a system for collecting parking fees without the need for pay display machines, parking meters or other similar devices. The system consisted of a database which contained three categories of codes, a code for each registered user, the VRN (“vehicle registration number”) for the registered vehicle and a code for each parking place. After

86 HovR T. 10333-06.
registering to the system, a user is able to pay for parking by transmitting his personal code, the code of the vehicle he wishes to park (the VRN number) and the code of the parking place he wishes to use. The method of data transmission between the user and the system is left open, but using the telephone network is given as an example. The user ends the parking session by transmitting his personal code to the system, at which point he will be billed. Parking attendants can check if the parking of a particular vehicle is being paid for by comparing the vehicle’s VRN to the database.

The allegedly infringing system, called "P-mint", works in a similar fashion. Utilization of the system requires that the user registers a phone number and a VRN to the database. The user can then start a parking session by parking the vehicle with the registered VRN and calling the car park operator using the phone with the registered phone number. The user is then prompted to enter a code specific to the area where the car has been parked. The user can then terminate the parking session by calling car park operator again.

The outcome of the case hinged how claim 1D was to be understood. The phrase of contention was “parking place” in the sentence “every parking place is assigned a […] code”. Tele P argued that it should be understood to mean “parking area”, i.e. an area consisting of multiple parking spots. P-mint on the other hand argued that it meant “parking spot”. The court sided with P-mint, citing three main reasons. First of all, certain claimed benefits of the invention (such as the ability to book a parking spot) required that each individual parking space was given a code. Second, the parking place code is later described as “the unique code assigned to that parking space which is occupied by said vehicle.” A vehicle can, of course, only occupy a parking spot and not a parking area. Finally, in some of the translations, the term “every parking space of the car park” is used instead of “parking place”. This too implies that “parking place” is to be understood as “parking spot”.

The problem when trying to determine what outcome an application of the cost theory would be is, again, that many relevant factors were ignored by the parties and the court. It does however seem likely that the courts decision is in line with the cost theory. There are reasons to believe that the inventor specific costs were low. Much of what the patent seems to cover had actually
already been invented. For instance, there existed a parking system covered by an American patent which consisted of scanners connected to a database which would store information about registered vehicles and users. The novel aspects of the Tele P system was the database could be accessed from the phone network and that each parking space was given a unique code.

Furthermore, it seems unlikely that the patent itself is of much benefit to a competitor who wants to copy the system. Besides the generic idea, not much relevant information for implementing the parking system is revealed. The cost of replication therefore seems somewhat similar to the cost of invention. Again, the idea itself might be very valuable and ingenious, but it seems unlikely that the inventor specific were very large.

7.3.3 Amgen and Hoechts
The final case I will discuss is the English Amgen case. The parties were two biotechnology companies, Amgen and Hoechts. Amgen argued that Hoechts was infringing on one of their patents. The patent in question claimed a certain manufacturing process for a hormone named EPO. Instead of attempting to purify natural EPO, Amgen was using genetic engineering to manufacture it.

The process involved inserting human DNA into Chinese hamster ovary cells. These genetically modified cells would then produce EPO. Hoechts method was different in that human cells were used instead of hamster cells. Furthermore, the sequence introduced to the cell was not the one which coded for EPO, but rather one which activated another gene which coded for EPO (but is typically dormant). Amgen argued that the process infringed on claim 1 by using a "DNA sequence [...] in securing expression (of EPO) in a...host cell". The court sided with Hoechts. One of the stated reasons was that the genes which expressed EPO in the Amgen process were exogenous, compared to the Hoechts process where they were endogenous.

Applying the cost theory in this case is more difficult than in the previous cases. One the one hand, it seems like there were substantial costs involved when inventing the Amgen process. On the other hand, there also seems to have been substantial costs involved in developing the Hoechts technique.

Furthermore, it is difficult to determine how large share of Amgen's costs were inventor specific.

This is not an uncommon situation. It is often difficult to apply the cost theory. This is partly due to the fact that the relevant factors are not discussed by the parties or the court. Another reason is that the relevant factors, even if known, might not enable us to provide a clear cut answer. Suppose that things are as they seem: both Amgen's and Hoecht's R&D costs are high. Amgen's high R&D costs point in the direction of a wide patent scope while the opposite is true for Hoecht's costs. There might not be enough additional information to determine which factor is to be given more weight. In these cases it seems reasonable to not use the cost theory and rely on traditional methods instead.
8. Concluding Remarks

8.1 Possible Reforms

In this section I will discuss some possible reforms to make the determination of patent scope more in line with the cost theory.

(1) Avoid wide interpretations when R&D costs are low.

An implication of the cost theory is that a patent should not allow the inventor to capture more benefits than the costs of R&D. If the R&D costs are low, then the amount needed to recoup the R&D costs are also low. Therefore low R&D costs implies a narrow patent scope.

(2) Use a wide scope when R&D costs are high.

This is the inverse of (1). Since high R&D costs means that the inventor has more losses to recoup, choosing a wider scope can be warranted. An example of an industry with high R&D costs is the pharmaceutical industry. The cost of a new drug is on average several hundreds of millions of dollars. The costs of replication tends to be comparatively low, thus a wide patent scope can be necessary for an efficient level of R&D investment.

(3) Looking at other means of appropriability.

Another factor relevant for determining efficient patent scope is the ability of the inventing firm to recoup its costs without a patent. If the inventing firm can recoup a large fraction of its costs without a patent, then a narrower patent scope can be warranted.

On such factor is the speed of replication. If replication is slow, the inventing firm will be the only provider of the invention for a period of time. This 'market exclusivity' has the same effects of a patent; the firm will be able to reap supra-normal profits while the competitors are trying to replicate the invention. If the inventing firm has a large enough 'first mover' advantage, patent protection is not necessary.

The appropriability factor is also affected by the type of firm. Some firms

---

89 Most of the costs are caused by testing and regulatory approval. (Burk (2012) p. 407). A firm replicating the invention do not face these costs.
engage in both R&D and the production of goods, while other engage purely in R&D or technology licensing. The appropriability factor will generally be lower for pure R&D firms, since they will lack the means of recouping their costs by embodying the invention in a product and selling it. Often, their only way of recouping the R&D costs is through licensing. They will therefore on average need stronger patent protection in order to recoup their costs.

An interesting implication is that non-practicing entities (also known as "patent trolls") should receive stronger patent protection, since have no means of recouping their costs other than licensing.

(4) "Inverting" the Pioneer Doctrine

The current patent system has designated a class of inventions as deserving extra wide protection. These are the so called pioneer inventions. What pioneer inventions have in common is that are "covering a function never before performed, a wholly novel device, or one of such novelty and importance as to mark a distinct step in the progress of the art". Examples include the laser, the microscope, the light bulb, the radio, television and the thermometer. In more economic terms, what pioneer patents have in common is that they tend to have very high social benefits.

This goes against the cost theory. Just because an invention has high social benefits does not mean it deserves a wide patent scope. If the costs of R&D were really low or the inventing firms has the ability to recoup the costs in other ways, a narrow patent scope is often sufficient.

Rather than strong protection for inventions with a high social value, the cost theory suggests that inventions with a high R&D cost and a comparatively low social value are the ones most deserving of a wide patent scope. These are the inventions which are least likely to be invented without the patent system. In order to recoup the costs of such an invention, the inventing firm needs to be able to capture a very large percentage of the social value of the invention. This might not be possible without a wide patent scope.

---

8.2 A Few Words of Caution

Even though the cost theory can be used to make the process of determining patent scope more economically efficient, it is wise not to go overboard and give the theory too great of a weight.

There are a few reasons for this. First and foremost, traditional interpretation is required for the cost theory to fulfill its role in an economically efficient way. A patent can roughly be thought of as a payment from the government for inventing something socially useful. The payment consists of the right to exclude others from using the invention. It is not however necessary for the payment to take this form. An inventor might for instance be paid directly from the government. Another possible type of payment is the right to exclude others from using some other invention.

There are reasons to believe that the right to exclude is efficient. The scope of this right does however need to have at least some sort of overlap with what has been invented. Thus the need for traditional interpretation in order to determine what has actually been invented.

A similar reason to avoid too much divergence from the patent claims is that doing so will create uncertainty as to whether or not a product is infringing on a patent. Suppose a firm is contemplating whether or not to develop a new product. A very important factor when making this decision is how many patents the product will infringe. If it turns out that the product will infringe on a lot of patents, the firm might decide against developing the product due to the costs of licensing and licensing negotiations. Similarly, if the firm is unable to determine how many patents the product will infringe it might decide against investing due to the potential costs. Since the way in which firms can determine if they are infringing on a patent is by studying the claims, a policy

---

93 I.e one which the inventor did not invent.
95 For instance, giving the right for a pharmaceutical to exclude semiconductor manufacturers from using a semiconductor patent as "payment" for inventing a new pharmaceutical hardly seems efficient, no matter how large the costs are.
96 Another way of thinking about this is to return to the "idea space" analogy in section 2.2. The scope of a patent can be thought of as the area of the figure. The "overlap" is determined not only by the scope but also where the figure is placed. The cost theory can help in determining the former but not the latter.
97 This is known as the fuzzy boundaries problem. (Bessen & Meurer (2009) p. 53).
which diverges from the claims will create such uncertainty.

Furthermore, there are instances where the cost theory does not provide much guidance. Courts do not delineate patent scope at an abstract level, but rather in relation to a specific product or process by answering the question if infringement has occurred or not. This means that the efficient solution might not be in the court's choice set. Rather, the court will have to chose between an inefficiently low reward or an inefficiently high one.

Finally, combining the cost theory and more traditional methods of determining scope can lead to strange results. For instance, the number of firms infringing on a patent will in some cases depend the order in which they are sued. In these situations, not using the cost theory seems reasonable.

It is however worth underscoring that the existence of costs are a central aspect of the L&E justification for the existance of the patent system. It therefore seems reasonable that they are taken into account on those occasions when it is possible.

---

98 Suppose that two firms are infringing on a patent. One of the firm's infringement is more obvious than the infringement of the other. The firm which is more obviously infringing is sued first and the patent holder wins. The second firm is then sued. This time the court sides with the defendent, arguing that the plantiff has another way of appropriating the cost (the damages from the first trial) and that there is therefore no need to give the patent a very wide scope. Now, suppose the order of suits is reversed. In the first trial the court sides with the plantiff against the less infringing firm because the patent holder has no other means of appropriation. In the second trail, the court also sides with the plaintiff because the infringement is so obvious.
9. References

Official Publications:

Prop 1993/94:22.

Articles:


Machlup, F. (1958) *Can There Be Too Much Research?* With competing demands on scarce resources, more applied research may curtail education and progress. Science, 128(3335), 1320.


**Books:**


Cases:

**England:**

Kirin-Amgen Inc & Ors v Hoechst Marion Roussel Ltd & Ors [2004] UKHL 46 (21 October 2004)
Sweden:
HovR T 10333-06.
NJA 2000 s. 497.

United States:

Patents:
SE367934 T3 "Insecticides."
SE429732 T3 "Sustained-release compositions containing cation exchange resins and polycarboxylic polymers"
US RE117 E "Improvement in electro-magnetic telegraphs"
US549160 A "Road-engine"
US4942516 A "Single chip integrated circuit computer architecture"
US7415666 B2 "Method and system for navigating paginated content in page-based increments."

Other: