On the Optimal Novelty Requirement in Patent Protection

Klaus Kultti
University of Helsinki[*]

Antti Miettunen
University of Helsinki[‡]

February 13, 2008

Abstract

Novelty value is one of the legal requirements for a patentable innovation but it has been given relatively little attention in the literature. It is often abstracted away by assuming that any innovation is patentable. We study the optimal novelty requirement in a model where ideas are scarce, and where turning an idea into an innovation requires resources. We show that all innovations should not be patentable but it is optimal to have a non-zero novelty requirement. The equilibrium investment in R&D is an inverse-U-shaped function of the novelty requirement.

JEL classification: O31, O34

Keywords: Patents, novelty requirement, scarcity of ideas.

1 Introduction

Intellectual property rights have many dimensions. For instance, we can distinguish the length, the width and the strength of a patent. The number of property right holders may vary as can be seen with patents and copy-rights. There is a large body of literature, both theoretical and empirical, that studies these dimensions. Less attention has been given to the novelty requirement of innovations. For instance O'Donoghue (1998) points out that the novelty requirement is often abstracted away by "either explicitly or implicitly assuming that any innovation is patentable" (p. 6). This is somewhat surprising for at least three reasons: First, most patent laws require that an innovation has to possess sufficient novelty value in order to be patentable. Depending on the country the requirement may be that of *non-obviousness* (US) or of *involving an inventive step* (Europe). The invention also has to be *useful* (US) or *susceptible of industrial application* (Europe). Second, it is clear that the novelty

[*]Klaus.Kultti@helsinki.fi

[†]Antti.T.Miettunen@helsinki.fi

[‡]Financial support from the Center for the Analysis of Property Rights and Innovation (CAPRI), and Academy of Finland is gratefully acknowledged.

requirement affects a firm's expected returns from an innovation (Hunt (1999), O'Donoghue (1998)). Third, it is widely argued that patentability standards have declined and the quality of patents - particularly within the software industry - is too low (see e.g. Jaffe and Lerner (2004)). One might therefore imagine that more attention were given to studying the impact of the novelty requirement.

Defining novelty is not as straightforward in practice as, say, defining the length of a patent. In court cases novelty has mostly been associated with technical aspects, but some weight has also been given to commercial success (*Graham v. John Deere Co., 383 U.S. 1* (1966)). As patents provide protection along many dimensions, novelty can refer to the horizontal as well as vertical aspects, i.e., to the width of a patent or the size of technological advancement. Whatever the interpretation of novelty, it seems that the requirement is there to prevent duplication, and to secure sufficient rewards to the innovators; presumably it is much easier to come up with marginal improvements that decrease the rewards of the original innovator, than to make substantial new innovations.

We approach novelty from the point of view of scarcity. There is a limited number of ideas based on which one can make innovations. Granting a property right to an innovation precludes others from using that idea to generate other innovations. If the original innovation is not very valuable - or novel in our terminology - then more valuable future innovations are potentially lost. Consequently, it is not socially desirable to grant property rights to marginal innovations. To model this we think that innovators come up randomly with ideas. Ideas are not rendered into innovations automatically but this happens probabilistically, the probability being the higher the higher the R&D investment. The value of an innovation is also randomly determined. Each idea that is developed into an innovation, and is granted a property right, is removed from the set of available ideas. This creates a trade-off between realising the value of a new innovation immediately (and making the idea unavailable), and leaving the idea available for a potentially more profitable innovation.

We construct an infinite horizon discrete time economy where there is a stock of available ideas and a stock of innovations that provides profits, or utility. The size and the value of these stocks depend on the novelty requirement which is a policy parameter. We identify novelty with the value of an innovation so that valuable innovations have large novelty value. We determine the optimal novelty requirement both under perfect patent protection and under imperfect protection, i.e., when patents can be invented around, imitated, or when there are spill-overs. This is modelled by assuming a probabilistic patent which means that an innovation may remain the property of the innovator or it may become public. If it becomes public there is production at the competitive level and the gains for the society are higher than in monopoly prodution.

Our main objective is to determine how R&D efforts depend on the novelty requirement. If the novelty requirement is low, even low value innovations can be granted a patent, but there are not many available ideas. An individual innovator would like to adopt all innovations regardless of their value, but quite surprisingly it turns out that increasing the novelty requirement actually benefits even the individual agent and increases the R&D efforts: This is because what is lost as low value innovations is more than compensated by the larger stock of available innovations, i.e., by the increased ease of coming up with an innovation. Not too surprisingly, it also turns out that welfare is not maximised when the protection granted by a patent is maximal. This follows from the assumption that under perfect patents the combined profit and consumer surplus is less than under perfect competition.

Before going to the related literature let us point out some of the idiosyncracies of our model. First, we model the search for new innovations as an urn-ball process. The ideas are randomly contacted by the innovators. As several innovators may come up with a similar innovation, there is a possibility of simultaneous innovation. Exactly the same kind of model of simultaneous innovations has been used in Kultti and Takalo (2008) as well as in Kultti, Takalo and Toikka (2007) where the simultaneity plays a more significant role than in this model. Because of the simultaneous innovation feature we assume that the priority rule in patents is first-to-file. Second, there is a distinction between ideas and innovations. Finding an idea does not automatically lead to an innovation. It requires an R&D investment, and even this does not help if there is already a patented innovation based on the same idea. Third, novelty is associated with the magnitude of value that an innovation generates; high value implies high novelty and low value low novelty. Fourth, novelty aspects are perhaps most relevant in sequential and cumulative innovation processes; our model can be interpreted as a cumulative process where the trend has been subtracted so that it features a steady state. However, this is just interpretation and technically our model is not one of cumulative innovations.

Scotchmer and Green (1990) study the relationship between novelty and disclosure of information. They have a two-firm-model where there are two possible inventions. The second invention builds on the first one, and the first innovator has a trade-off between patenting and disclosing the first innovation, and between keeping it secret while aiming at the second, high value, innovation. The novelty requirement in their model is weak if both the first and second innovation can be patented, and strong if only the second innovation, requiring two innovative steps, can be patented. They find that weak novelty requirement is socially preferable. In our model the novelty requirement is a continuous variable, and its optimal value is more responsive to changes in the parameters of the model.

Klemperer (1990) and Gilbert and Shapiro (1990) both study the trade-off between patent width and length. In the former model the width of a patent is defined in a space of differentiated products. Klemperer determines the cases where the patent length is maximal and minimal, and this depends on the elasticity of demand. The latter model features a homogeneous good, and the concept of patent width is such that the wider the patent the higher profits it generates. Gilbert and Shapiro find that infinitely long patents are typically optimal. Both of the articles deal with a single patent and there is no interaction between the innovator, or two innovators, and the rest of the economy. In our model there is direct interaction between innovators because of the simultaneous nature of the innovation process. Further, the policy variable affects the whole economy, i.e., the steady state stocks of innovations, and also the explicitly calculable welfare.

Bessen and Maskin (2008) provide a very simple model of cumulative innovation with two firms. Each innovation builds on the previous innovation, and if an innovator gets a perfect patent he can close the competing innovator from the market. Not too surprisingly, they find that weak patents are optimal.[1]

Hopenhayn, Llobet and Mitchell (2006) study cumulative innovation with many heterogeneous innovators, or firms. The focus is on the system of optimal rewards when the profitability of innovations is private knowledge. Their basic point is that monopoly rights are scarce, and granting a strong monopoly to one innovator equals granting weak monopoly rights to future innovators. In this environment intellectual property protection has to find a balance between sufficient incentives to innovative activity for current innovators and future innovators. Hopenhayn, Llobet and Mitchell characterise the optimal intellectual property protection where a patentee commits to a menu of patents, i.e., price and duration, from which a future innovator can choose one to buy out the present patentee. This is a model of private information, and even though there are many innovators they appear in a sequence. In our model private information does not play a role, and there are many innovators whose simultaneous actions determine the state of the economy.

The rest of the article is organized as follows. The model is presented in the following Section, steady-state analysis is done in Section 3, optimal novelty requirerement condition is presented in 3.1 and comparative statics in 3.2. Different patent regimes are studied in 3.3, and finally conclusions are made in Section 4.

2 The Model

Consider the steady state of an infinite horizon economy that proceeds in discrete time. The economy is populated by a measure one of agents, innovators, who also constitute the consumers of the economy. There is also a number of potential ideas of measure I. Some of the ideas are in use and cannot be used for an innovation, while some of them are available. Each period the innovators try to develop an innovation. This is an uncertain activity and we assume that investment i into R&D allows the innovator to come up with an innovation with probability $(1 - e^{-i})$.[2] This, however, happens only if the innovation is based on an idea that is available. We thus make a distinction between ideas and innovations.

When one or several innovators find an idea, its value, or novelty value, is randomly determined.[3] If an idea and the resulting innovation is granted a

[1] Carpentier and Kultti (2003) show that all their results are completely overturned if the R&D effort is a choice variable.

[2] The rationale for this form can be found in Kultti (2003).

[3] Our first version featured a linear demand curve for each innovation, and innovations

patent then the innovators cannot develop new innovations based on that idea. This creates a trade-off as granting a patent to an innovation of small value allows realising this value immediately, while it prevents realising an innovation of large value based on the same idea in the future.The society's role is to grant intellectual property rights to innovations allowing the innovators to reap at least some monopoly rents from their innovations; this amounts to deciding the required level of novelty.

Not every innovator necessarily succeeds in turning the idea into an innovation; the probability depends on the amount of R&D as described above. Of the successful innovators one is given a patent to the innovation. We allow the property right to be imperfect, and it is then characterised by a number $\alpha \in [0, 1]$. One should think of α to indicate how much of the proceeds of the innovation the innovator can retain and how much spills over to the others. For modelling purposes it is most convenient to assume that with probability α the innovator can retain the proceeds wholly, and that with probability $1 - \alpha$ the innovation becomes public domain and it is produced at the competitive level that generates zero profits. We regard α as exogenously given.[4]

We use an urn-ball model to describe the process of finding ideas. The innovators contact the ideas randomly, and the probability that an idea is contacted by k innovators is determined by a Poisson distribution with rate $\theta \equiv \frac{1}{I}$ and it is given by $e^{-\theta}\frac{\theta^k}{k!}$. To facilitate analysis we assume that the quality of inventions is uniformly distributed on the unit interval. If the novelty requirement is $a \in [0, 1]$, then only innovations whose novelty index is higher than a are granted a property right. We assume that an invention whose novelty index is a yields profit $2a$ to the innovator. If the innovation becomes public it is produced at the competitive level and yields consumer surplus $\gamma 2a$ where $\gamma > 1$.[5] We focus on the steady state and to this end we assume that each period each innovation faces a probability $1 - \lambda$ of dying, or becoming obsolete. When this happens the idea underlying the innovation again becomes available to the innovators.

The order of events within a period is as follows:

1. Innovators invest in R&D

turned out of different quality. The higher the quality the higher the intercept of the demand curve indicating that people were more willing to pay for the innovation.

[4] It is usually thought that the society can decide the strength of the patent protection, but there are many things that make this suspect. The innovators can keep their innovations secret, depending on the technology and the effectiveness of the legal system the innovations may be invented around or they may be vulnerable to reverse engineering or imitation, and there might also be spillovers due to factors like information travelling with employees who change jobs. Making the strength a decision variable also makes this model amenable only to numerical methods, which, of course, obscures the relation between the scarcity of ideas and novelty value.

[5] These are just convenient magnitudes not derivable from any primitives. If one assumes that a novelty index a corresponds to an inverse demand curve $p = a(1-q)$ then an invention yields monopoly profits $a/4$ and consumer surplus $a/8$, and when it is in public domain it yields zero profits and consumer surplus $a/2$. The chosen magnitudes yield pretty much the same insight as this more complicated model.

2. They randomly contact ideas and the value of the idea is randomly determined
3. One of the innovators is granted a patent and the innovation either remains private or becomes public
4. Profits and consumer surplus accrue[6]
5. The innovations either remain viable or become obsolete.

Here it is assumed that investment is a prerequisite for contacting an idea. The investment can be thought to include the remuneration of research personnel and research facilities like laboratories. Of course, typically there are also R&D investments after an idea has been found but we abstract from these. Our assumption about investments highlights the view that ideas are scarce, and that granting a property right to an innovation precludes others from utilising the underlying idea which is costly to the others.

3 Analysis

We focus on the steady state analysis of the model. In the steady state there are three magnitudes to be determined. Let S be the stock of private innovations (i.e. innovations that were granted a property right and which did not come to the public domain), T the stock of public innovations (i.e. innovations that were granted a property right but which became public domain), and H the stock of available ideas (i.e. ideas that are not currently used for an existing innovation). The steady state levels of these variables are determined by equating the outflow from and the inflow to the stock.

The outflow from S is just $(1-\lambda)S$. The inflow is is given by

$$\sum_{h=1}^{\infty} e^{-\theta} \frac{\theta^h}{h!} \sum_{j=1}^{h} \binom{h}{j} \left(1-e^{-i}\right)^j e^{-(h-j)i} I \frac{H}{I} (1-a)\alpha \tag{1}$$

The outer sum traces the probability that any idea is contacted by h innovators. The inner sum is the probability that at least one of these h innovators manages to turn the idea into an innovation. To become a private innovation its value has to be higher than a, this happens with probability $1-a$, and it has to remain private, this happens with probability α. There are I ideas but only H of them are available. Expression (1) simplifies to

$$H(1-a)\alpha\left(1-e^{-\widetilde{\theta}}\right) \tag{2}$$

where $\widetilde{\theta} = \theta\left(1-e^{-i}\right)$. The steady state stock of private innovations is then

$$S = \frac{H(1-a)\alpha\left(1-e^{-\widetilde{\theta}}\right)}{1-\lambda} \tag{3}$$

[6]Notice that an innovation always generates economic value at least once before it dies.

The stock of public innovations T is just $S\frac{1-\alpha}{\alpha}$ yielding

$$T=\frac{H(1-a)(1-\alpha)\left(1-e^{-\widetilde{\theta}}\right)}{1-\lambda} \tag{4}$$

Finally, the stock of available innovations H is determined by $H=I-S-T$ which can be solved

$$H=\frac{1-\lambda}{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}I \tag{5}$$

Inserting this into (3) and (4) yields

$$S=\frac{(1-a)\alpha\left(1-e^{-\widetilde{\theta}}\right)}{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}I \tag{6}$$

$$T=\frac{(1-a)(1-\alpha)\left(1-e^{-\widetilde{\theta}}\right)}{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}I \tag{7}$$

Next we determine the equilibrium investment level in the economy. Since the innovations die with positive probability we can assume that there is no discounting. Assume that everyone except the agent under study invests i in the R&D activity, and denote his investment level by n. He expects

$$-n+\frac{H}{I}\alpha\sum_{h=0}^{\infty}e^{-\theta}\frac{\theta^h}{h!}\sum_{j=0}^{h}\left(1-e^{-n}\right)\binom{h}{j}\left(1-e^{-i}\right)^j e^{-(h-j)i}\frac{1}{j+1}\int_a^1 2xdx\frac{1}{1-\lambda} \tag{8}$$

In (8), n is the cost of investment, and $1-e^{-n}$ is the probability that the investment is successful. The probability of contacting an available idea is $\frac{H}{I}$. Private profits are generated only if the innovation remains private which happens with probability α. The outer sum keeps track of the number of other innovators who contact the same idea as the agent under study. Given that the number is h, the inner sum keeps track of the number of successful innovators out of these h. If it is j, the probability that the agent under study is granted the a patent is $\frac{1}{j+1}$. This only happens if the innovation is more valuable than a; the expected value is $\int_a^1 2xdx$. Finally, the expected life time of the innovation is the reciprocal of the obsolescence probability $\frac{1}{1-\lambda}$. Expression (8) simplifies to

$$-n+\frac{\alpha\left(1-a^2\right)}{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}\frac{1-e^{-\widetilde{\theta}}}{\widetilde{\theta}}\left(1-e^{-n}\right) \tag{9}$$

Taking the first order condition with respect to n and evaluating it at $n=i$ yields

$$\frac{1-e^{-\widetilde{\theta}}}{\widetilde{\theta}}e^{-i}-\frac{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}{\alpha\left(1-a^2\right)}=0 \tag{10}$$

which determines the equilibrium $i^*(a,\alpha,\lambda,\theta)$. In the relevant parameter space the LHS of (10) must be positive at $i=0$. This condition gives the parameter restriction $\alpha\left(1-a^2\right)>1-\lambda$. Having specified the market equilibrium, we can turn to the optimal novelty requirement.

3.1 Welfare and the Optimal IPR

Our aim is to find the optimal level of novelty requirement a. For this purpose we have to determine how the novelty requirement affects the equilibrium investment. Totally differentiating (10) with respect to a and i, it is straightforward to derive

$$\frac{di}{da}=\frac{\left(1-e^{-\widetilde{\theta}}\right)\left(2\alpha a e^{-i}-\widetilde{\theta}\right)}{\left(\begin{array}{c}-\alpha\left(1-a^2\right)e^{-i}\left(1-e^{-\widetilde{\theta}}-\theta e^{-i}e^{-\widetilde{\theta}}\right)-\\ \theta e^{-i}\left[1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)+(1-a)\widetilde{\theta}e^{-\widetilde{\theta}}\right]\end{array}\right)} \tag{11}$$

Substituting from (10) for the brackets in the denominator we get a more convenient form

$$\frac{di}{da}=\frac{\left(1-e^{-\widetilde{\theta}}\right)\left(1-e^{-i}\right)\left(-2\alpha a e^{-i}+\widetilde{\theta}\right)}{(1-a)\,e^{-i}\left\{\alpha(1+a)\left(1-e^{-\widetilde{\theta}}\right)-\widetilde{\theta}e^{-\widetilde{\theta}}\left[\alpha(1+a)e^{-i}-\widetilde{\theta}\right]\right\}} \tag{12}$$

As we focus on a steady state our welfare measure is the combined profits and consumer surplus minus the R&D costs per period. Since the stock of private inventions as well as public inventions both consist of inventions whose novelty index is at least a, the expected profit of a private invention is $\int_a^1 2x\frac{dx}{1-a}=1+a$, and the expected consumer surplus from a public invention is $\int_a^1 \gamma 2x\frac{dx}{1-a}=\gamma(1+a)$. The periodic welfare is then $W\equiv(1+a)S+\gamma(1+a)T-i^*(a,\alpha)$ which equals[7]

$$W=\frac{\left(1-a^2\right)\left(1-e^{-\widetilde{\theta}}\right)}{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}I\left(\alpha+\gamma(1-\alpha)\right)-i^*(a,\alpha) \tag{13}$$

This can also be given a more convenient form by substituting from (10)

$$W=\frac{\left(1-e^{-i}\right)\left(\alpha+\gamma(1-\alpha)\right)}{e^{-i}\alpha}-i$$

[7] Notice that we have assumed that innovations of value less than a are not taken into use at all. This advances the tractability of the model but it would also be the case if, for instance, the first adoptor of the unprotected innovator were to pay some kind of fixed cost, say marketing, and the subsequent innovators could adopt the innovation for free.

Now it is straightforward to formally determine the optimal form of the intellectual property right by taking the first order condition of W with respect to a. It turns out as follows

$$\frac{dW}{da} = \frac{di}{da} \frac{\alpha + \gamma(1-\alpha) - e^{-i}\alpha}{e^{-i}\alpha} = 0 \tag{14}$$

from which we see that the zero of the first order condition for welfare is at the same point as the zero of $\frac{di}{da}$. This is presented in our main result:

Proposition 1 *The optimal novelty requirement is uniquely determined by*

$$-2\alpha a e^{-i} + \widetilde{\theta} = 0. \tag{15}$$

It is bounded away from zero for all parameter values.

It follows that the equilibrium investement i is related to the novelty requirement according to the following result:

Proposition 2 *As a function of the novelty requirement, the equilibrium investment i takes an inverse-U-shaped form.*

Notice that $-2\alpha a e^{-i} + \widetilde{\theta}$ is positive for small values of a, and when a is close to unity i goes to zero and $-2\alpha a e^{-i} + \widetilde{\theta}$ becomes negative. This means that as a function of a $\frac{di}{da}$ is first increasing, peaks at some $\widehat{a}$ and is then decreasing. This is quite intuitive as when a is zero all the inventions are taken into use, and there are few ideas left, or H is small. Increasing a a little does not change an individuals payoff from an invention very much but it increases H leading to greater expected payoff as it is easier to find inventions. Increasing a further, however, makes it more difficult to find an idea with a sufficiently high novelty index, and as the expected payoffs decrease the willingness to invest in R&D also decreases.

Having specified the welfare maximizing novelty requirement and the relationship between i and a, we can turn to comparative statics.

3.2 Comparative Statics

Relation (10) determines the equilibrium investment in R&D, and relation (15) determines the socially optimal novelty requirement. We need to know how the strength of the property right α, the survival rate λ, and the ratio of inventors to potential ideas or the easiness of invention θ affect the novelty requirement. To this end we totally differentiate (15) to get

$$\begin{aligned} da\left\{2\alpha e^{-i}\right\} + d\alpha\left\{2ae^{-i} - 2\alpha a e^{-i}\frac{di}{d\alpha} - \theta e^{-i}\frac{di}{d\alpha}\right\} + \\ d\lambda\left\{-2\alpha a e^{-i}\frac{di}{d\lambda} - \theta e^{-i}\frac{di}{d\lambda}\right\} + \\ d\theta\left\{-2\alpha a e^{-i}\frac{di}{d\theta} - \left(1-e^{-i}\right) - \theta e^{-i}\frac{di}{d\theta}\right\} = 0 \end{aligned} \tag{16}$$

Notice that we have used the fact that at the optimum $\frac{di}{da} = 0$.

To find out the signs of the various partials we need to determine how the parameters affect the equilibrium investment.This is achieved by totally differentiating (10) (see Appendix A). After substituting for $\left[1 - \lambda + (1-a)\left(1 - e^{-\widetilde{\theta}}\right)\right]$ from (10), we can see that the partial derivative of the investment, w.r.t. α, is positive:

$$\frac{di}{d\alpha} = \frac{(1+a)\left(1-e^{-i}\right)\left(1-e^{-\widetilde{\theta}}\right)}{\alpha(1+a)\left(1-e^{-\widetilde{\theta}}\right) - \widetilde{\theta}e^{-\widetilde{\theta}}\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right]} > 0$$

Similarly, we get the comparative statics of investment w.r.t. the survival rate λ. It can be easily seen that it is also positive:

$$\frac{di}{d\lambda} = \frac{\left(1-e^{-i}\right)\widetilde{\theta}}{\alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}(1-a)\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right]} > 0$$

As one would expect, the partial derivative of i with respect to $\theta = \frac{1}{I}$ turns out to be negative (See Appendix A for the proof):

$$\frac{di}{d\theta} = \frac{\left(1-e^{-i}\right)\left\{\widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}\alpha(1-a^2) - \alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}^2(1-a)e^{-\widetilde{\theta}}\right\}}{\theta\left\{\alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}(1-a)\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right]\right\}} < 0$$

Inserting the above partial derivatives of investment and solving from (16) we get the following signs for the comparative statics of a:[8]

$$\frac{da}{d\alpha} = \frac{\left(2\alpha a e^{-i} + \theta e^{-i}\right)\frac{di}{d\alpha} - 2ae^{-i}}{2\alpha e^{-i}} > 0 \tag{17}$$

$$\frac{da}{d\lambda} = \frac{\left(2\alpha a e^{-i} + \theta e^{-i}\right)\frac{di}{d\lambda}}{2\alpha e^{-i}} > 0 \tag{18}$$

$$\frac{da}{d\theta} = \frac{\left(2\alpha a e^{-i} + \theta e^{-i}\right)\frac{di}{d\theta} + \left(1-e^{-i}\right)}{2\alpha e^{-i}} > 0 \tag{19}$$

The above results are gathered in the following two propositions:

Proposition 3 *The equilibrium investment is increasing in the strength of the property right and in the number of ideas, and decreasing in the probability of obsolescence.*

This is merely what one would expect and confirms that the model is well-behaved.

[8] See Appendix for the proof for $\frac{da}{d\alpha} > 0$.

Proposition 4 *The optimal novelty requirement is increasing in the strength of the property right and decreasing in the probability of obsolescence and in the number of ideas.*

We can see that the model behaves very intuitively. First, (19) is simply our starting point: if the ideas are scarce the novelty requirement should play an important role, and if the number of ideas I is increased (and thus scarcity decreased), the need for the novelty requiment should diminish. Second, (18) states that the longer the expected lifetime of innovations, the more strict the novelty requirement should be. Obviously there is a trade-off between the expected time that innovations generate utility, and the treshold level we set for the minimum accepted utility. Third, (17) formalizes the relation between the direct and the indirect incentives: if the direct incentives are increased, i.e. α goes up (and thus i also goes up), then indirect incentives affecting through the equilibrium stock of free ideas have to be decreased in order to maintain the optimum. To put it differently, increasing the probability of obsolescence or the number of ideas does not make anybody worse off, and thus the novelty requirement can be decreased. Increasing α, on the other hand, makes the non-patent holders worse off and this has to be compensated by increasing the novelty requirement.

3.3 Different Patent Regimes

In most of the literature it is assumed that patents provide perfect protection, which corresponds to the case where $\alpha = 1$. This is an implausible assumption for many reasons, but we take a look at the effects of the novelty requirement under different patent regimes: one, where there are perfect patents, and the other, where $\alpha < 1$.

With $\alpha = 1$, the optimal level of the novelty requirement is determined by

$$-2ae^{-i} + \widetilde{\theta} = 0$$

One would expect that at $\alpha = 1$ the derivative of the welfare with respect to α, given the optimal value of a in (15), would be negative. This turns out to be true;

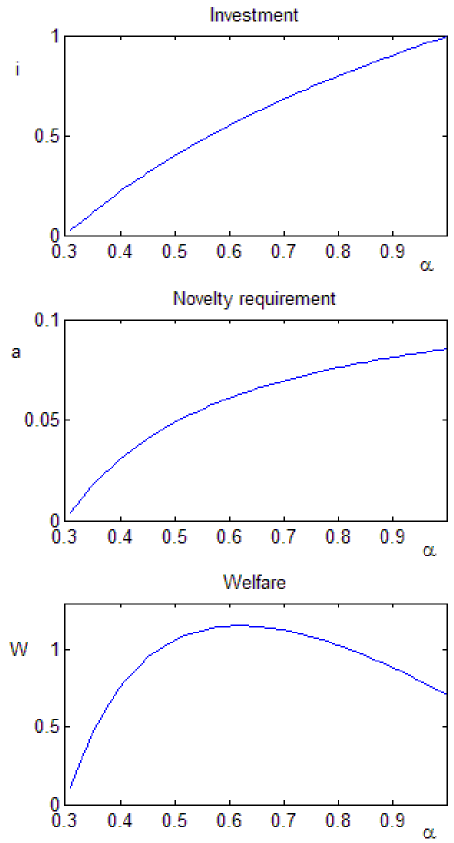
Proposition 5 *The partial of welfare with respect to α is negative at $\alpha = 1$*

$$\frac{dW}{d\alpha} |_{\alpha=1} = \frac{\frac{\partial i}{\partial \alpha}\left[1 - e^{-i}\right] - \gamma\left(1 - e^{-i}\right)}{e^{-i}} |_{\alpha=1} < 0$$

Proof. In the Appendix A. ■

When γ is close to unity, i.e. the public use of an innovation does not generate more utility than it generates private profit, and when θ is small, i.e. it is very easy to find new innovations, the perfect patent strength is optimal. This shows that the welfare is seldom maximised with perfect patent protection.

With imperfect patents we look at the optimum as a function of $\alpha < 1$. For numerical analysis we set $\gamma = 2$, $I = 10$, and $\lambda = 0.7$. Taking into account the parameter restriction $\alpha\left(1-a^2\right)>1-\lambda$ from (3), the patent strength varies between $\alpha\in[0.3,1]$. Plotting the investment, novelty requirement and welfare as a function of α gives the following results:



We can see that there exists a welfare-maximizing $\alpha=\widehat{\alpha}<1$. Insofar as the patent strength is a controllable policy variable, it would seem interesting to endogenise it simultaneously with the novelty requirement, but this makes the model analytically intractable and amenable only to numerical methdos. However, numerical analysis always produces a unique welfare-maximizing pair (a,α).

4 Conclusions

The patent law is particularly unclear about what the required novelty of an innovation means even though it is quite specific that an innovation must possess sufficient novelty to be patentable. There are two ways of thinking about novelty.

One is technical, and to proceed this way one would need to specify the available and potential technologies. This looks a difficult route to advance. The other way is to think about the economic value of an innovation, and in this article we have equated novelty with the economic profit an innovation yields. This is convenient as our viewpoint is that ideas are scarce, and granting a right to a low valued innovation precludes a high valued innovation based on the same idea in the future.

We determine the optimal novelty requirement, as well as the comparative statics, in a general equilibrium random matching model where we can track the changes in all parameters to the economywide magnitudes. In addition to being intuitive and tractable, the model successfully captures the tension between the interests of individual innovators and those of the society. Despite these desirable properties, there are still several development directions which invite further theoretical work.

First, one interpretation of our model is that of cumulative innovation where we have eliminated the trend; the value of an innovation is then an incremental value above the previous innovation. Explicitly modelling the cumulative innovation process where patented innovations become obsolete, or die, because a sufficiently large improvement has been invented seems very worthwhile. It would have the additional advantage that one of the exogenous parameters of our model, the obsolescence rate, could be endogenised.

Second, our model is unsatisfactory in that we cannot address simultaneously all the relevant dimensions of intellectual property protection. To determine the optimal policy one should be able to address the length, strength, width, novelty requirement, the number of property right holders and so on. Especially if our models are used for policy purposes we should be able to address the relationship between the different dimensions of intellectual property protection.

Third, we have focused on the innovation process at the level of the economy, and the strategic interactions between innovators have been assumed away. But it is clear that innovative activity is often economically significant in oligopolistic or developing industries where the strategic behaviour of the firms is the rule rather than an exception. Strategic behaviour may manifest in keeping an innovation secret (Kultti et al. (2007); Scotchmer and Green (1990)), patenting innovations in order to prevent competitors making advances, and in research co-operation.

Fourth, our model is one of perfect information, and it is clear that many aspects of R&D involve private information. As seen in, say, Hopenhayn et al. (2006) an environment where the optimal intellectual property protection under perfect information is easy to determine turns out non-trivial under private information.

References

Bessen, J. and Maskin, E.: 2008, Sequential innovation, patents, and imitation, *The RAND Journal of Economics (forthcoming)* .

Carpentier, X. and Kultti, K.: 2003, Cumulative innovations: Intellectual property regimes and incentives to innovate, *Discussion Paper, Department of Economics, University of Helsinki* **563/15.5.2003**. ISBN 952-10-0706-0.

Gilbert, R. and Shapiro, C.: 1990, Optimal patent length and breadth, *The RAND Journal of Economics* **21**(1), 106–112.

Graham v. John Deere Co., 383 U.S. 1: 1966, United States Supreme Court.

Hopenhayn, H., Llobet, G. and Mitchell, M.: 2006, Rewarding sequential innovators: Prizes, patents, and buyouts, *Journal of Political Economy* **114**(6), 1041–1068.

Hunt, R. M.: 1999, Nonobviousness and the incentive to innovate: An economic analysis of intellectual property reform, *Federal Reserve Bank of Philadelphia Working Papers* **99-3**.

Jaffe, A. B. and Lerner, J.: 2004, *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It*, Princeton University Press.

Klemperer, P.: 1990, How broad should the scope of patent protection be?, *The RAND Journal of Economics* **21**(1), 113–130.

Kultti, K.: 2003, p(x+y)=p(x)+p(y)-p(x)p(y), *Discussion Paper, Department of Economics, University of Helsinki* **559/April 2003**. ISBN 952-10-0700-1.

Kultti, K. and Takalo, T.: 2008, Optimal fragmentation of intellectual property rights, *International Journal of Industrial Organization* **26**, 137–149.

Kultti, K., Takalo, T. and Toikka, J.: 2007, Secrecy versus patenting, *The RAND Journal of Economics* **38**(1), 22–42.

O'Donoghue, T.: 1998, A patentability requirement for sequential innovation, *RAND Journal of Economics* **29**(4), 654–679.

Scotchmer, S. and Green, J.: 1990, Novelty and disclosure in patent law, *The RAND Journal of Economics* **21**(1), 131–146.

A Appendix

A.1 Proposition: $\frac{di}{d\theta} < 0$

Proof. The equation for $\frac{di}{d\theta}$ takes the following form:

$$\frac{di}{d\theta} = \frac{\left(1-e^{-i}\right)\left\{\widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}\alpha(1-a^2) - \alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}^2(1-a)e^{-\widetilde{\theta}}\right\}}{\theta\left\{\alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}(1-a)\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right]\right\}} \tag{20}$$

The denominator is positive since

$$\alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}(1-a)\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right] > \tag{21}$$

$$(1-a^2)\alpha e^{-i}\left\{\left(1-e^{-\widetilde{\theta}}\right) - \widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}\right\} > 0 \tag{22}$$

The nominator is negative since

$$\widetilde{\theta}e^{-i}e^{-\widetilde{\theta}}\alpha(1-a^2) - \alpha(1-a^2)\left(1-e^{-\widetilde{\theta}}\right)e^{-i} - \widetilde{\theta}^2(1-a)e^{-\widetilde{\theta}} < 0 \tag{23}$$

as already the first two terms are negative. ■

A.2 Market Equilibrium Condition

Market equilibrium condition is

$$\frac{1-e^{-\widetilde{\theta}}}{\widetilde{\theta}}e^{-i} - \frac{1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)}{\alpha\left(1-a^2\right)} = 0 \tag{24}$$

Totally differentiating it yields

$$\begin{aligned} di&\left\{\begin{array}{c} e^{-\widetilde{\theta}}\theta e^{-i}e^{-i}\alpha(1-a^2) - \left(1-e^{-\widetilde{\theta}}\right)e^{-i}\alpha(1-a^2) - \\ \theta e^{-i}\left[1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)\right] - \widetilde{\theta}(1-a)e^{-\widetilde{\theta}}\theta e^{-i} \end{array}\right\} + \\ &d\alpha\left\{\left(1-e^{-\widetilde{\theta}}\right)e^{-i}(1-a^2)\right\} + d\lambda\left\{\widetilde{\theta}\right\} + \\ d\theta&\left\{\begin{array}{c} e^{-\widetilde{\theta}}\left(1-e^{-i}\right)e^{-i}\alpha(1-a^2) - \\ \left(1-e^{-i}\right)\left[1-\lambda+(1-a)\left(1-e^{-\widetilde{\theta}}\right)\right] - \widetilde{\theta}(1-a)e^{-\widetilde{\theta}}\left(1-e^{-i}\right) \end{array}\right\} = 0 \end{aligned} \tag{25}$$

A.3 Proposition: $\frac{\partial W}{\partial \alpha}\mid_{\alpha=1} < 0$

Proof. The derivative takes the following form

$$\frac{dW}{d\alpha}\mid_{\alpha=1} = \frac{\frac{\partial i}{\partial \alpha}\left[1-e^{-i}\right] - \gamma\left(1-e^{-i}\right)}{e^{-i}}\mid_{\alpha=1} \tag{26}$$

After inserting $\frac{\partial i}{\partial \alpha}$, we can that this is negative if and only if

$$(1+a)\left(1-e^{-i}\right)\left(1-e^{-\widetilde{\theta}}\right) < \gamma\left[(1+a)\left(1-e^{-\widetilde{\theta}}\right) - \widetilde{\theta}e^{-\widetilde{\theta}}\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right]\right] \tag{27}$$

$$= \gamma\left[(1+a)\left(1-e^{-\widetilde{\theta}}\right) - \widetilde{\theta}e^{-\widetilde{\theta}}(1-a)e^{-i}\right] \tag{28}$$

by (14). As γ is larger than unity this holds for certain if

$$(1+a)\left(1-e^{-i}\right)\left(1-e^{-\widetilde{\theta}}\right) < (1+a)\left(1-e^{-\widetilde{\theta}}\right) - \widetilde{\theta}e^{-\widetilde{\theta}}(1+a)e^{-i} \tag{29}$$

which is equivalent to

$$\left(1-e^{-i}\right)\left(1-e^{-\widetilde{\theta}}\right) < 1-e^{-\widetilde{\theta}} - \widetilde{\theta}e^{-\widetilde{\theta}}e^{-i} \tag{30}$$

This is equivalent to

$$1-e^{-\widetilde{\theta}} - \widetilde{\theta}e^{-\widetilde{\theta}} > 0 \tag{31}$$

which holds always. ■

A.4 $\frac{da}{d\alpha} > 0$

Proof. $\frac{da}{d\alpha}$ is positive if and only if $\left(2\alpha a e^{-i} + \theta e^{-i}\right)\frac{di}{d\alpha} - 2ae^{-i} > 0$. This is equivalent to

$$\left(2\alpha a e^{-i} + \theta e^{-i}\right)(1+a)\left(1-e^{-i}\right)\left(1-e^{-\widetilde{\theta}}\right) > \tag{32}$$

$$2ae^{-i}\alpha(1+a)\left(1-e^{-\widetilde{\theta}}\right) - 2ae^{-i}\widetilde{\theta}e^{-\widetilde{\theta}}\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right] \tag{33}$$

This is equivalent to

$$(1+a)\left(1-e^{-\widetilde{\theta}}\right)\left[\widetilde{\theta} - 2\alpha a e^{-i}\right] > -2ae^{-i}\widetilde{\theta}e^{-\widetilde{\theta}}\left[\alpha(1+a)e^{-i} - \widetilde{\theta}\right] \tag{34}$$

but by (15) the LHS is zero. ■