



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Three Essays in Economics of Innovation

**Autor:
Stefano Trento**

**Director:
Michele Boldrin**

Departamento de Economía, Universidad Carlos III de Madrid

Getafe, Noviembre 2009

TESIS DOCTORAL

Three Essays in Economics of Innovation

Autor: Stefano Trento

Director/es: Michele Boldrin

Firma del Tribunal Calificador:

Firma

Presidente: Marco Celentani

Vocal: Antonio Ciccone

Vocal: Andres Erosa Etchebehere

Vocal: Gerard Llobet Codina

Secretario: Maria Belen Jerez Garcia-Vaquero

Calificación:

Getafe, de de

Three Essays in Economics of Innovation

Stefano Trento

Advisor: Michele Boldrin

Contents

1	Anticommons in Sequential Innovation	7
1.1	Introduction.	8
1.1.1	Related literature.	9
1.2	The model.	10
1.2.1	Technology.	11
1.2.2	Value of the innovation and structure of the information.	12
1.2.3	Market interaction.	13
1.2.4	Interpretation of the innovation technology	13
1.3	Equilibrium.	14
1.3.1	Elasticity of substitution.	16
1.3.2	Complements and Substitutes.	16
1.3.3	Increasing complexity.	18
1.3.4	High complexity and Monopolistic Competition.	20
1.3.5	The Tragedy of the Anticommons Revisited.	22
1.4	Patent Policy	23
1.4.1	Patent Policy with Imperfect Substitutability	26
1.4.2	Symmetric Patent Policy.	27
1.5	Patent Pools	28
1.6	Extensions.	30
1.6.1	Social value and cost depend on complexity.	30
1.6.2	No Asymmetric Information.	31
1.6.3	No price discrimination.	32
1.6.4	Uncertain return of the innovation.	32
1.7	Conclusions.	34
2	Accumulation of Claims in Sequential Innovation	35
2.1	Introduction	36
2.2	Innovation with patents	39
2.3	Innovation without patents	42
2.4	Patent pools	42
2.5	Comparison	43
2.6	Complete sale of patent rights	43
2.7	Endogenous patent pool formation	45
2.8	Socially optimal innovation	46
2.8.1	Static versus dynamic incentives	48
2.9	Dynamic externalities and optimal transfers	49
2.9.1	First best	49
2.9.2	Second best: optimal transfers	50

2.10	Ongoing innovation	50
2.11	Finite patents	51
2.11.1	Revenues depend on patent length	53
2.12	Conclusion	54
3	IPRs in a Model of Economic Growth	57
3.1	Introduction	58
3.2	The Model	60
3.2.1	Environment.	60
3.2.2	Preferences.	61
3.2.3	Technology.	61
3.2.4	R&D	61
3.3	Social Planner	62
3.4	Decentralized Steady-State Equilibrium With Infinitely Lived Patents	63
3.4.1	Representative Household.	63
3.4.2	R&D firms.	64
3.4.3	Producers.	65
3.4.4	Balanced Growth Path.	66
3.5	Decentralized Steady-State Equilibrium With Finite Patent Length	67
3.5.1	Balanced Growth Path.	68
3.6	Conclusions and Policy Implications	74
3.7	Appendix	76
3.7.1	Derivation of expected utility	76
3.7.2	Proof of Proposition (12)	76
3.7.3	Derivation of equation (3.22)	77
3.7.4	Derivation of equation (3.26)	78
3.7.5	Proof of Proposition (13)	78

List of Tables

3.1	Parameters values	69
3.2	Robustness Check for different values of σ Baseline parameters values: $\lambda = 0.4, \beta = 0.04, n = 10, \gamma = 1.25, L = 1$	72
3.3	Robustness Check for different values of n Baseline parameters values: $\sigma = 3, \lambda = 0.4, \beta = 0.04, \gamma = 1.25, L = 1$	72
3.4	Robustness Check for different values of β Baseline parameters values: $\sigma = 3, \lambda = 0.4, n = 10, \gamma = 1.25, L = 1$	73
3.5	Robustness Check for different values of γ Baseline parameters values: $\sigma = 3, \lambda = 0.4, \beta = 0.04, n = 10, L = 1$	73

List of Figures

1.1	Cost of innovation as a function of σ	17
1.2	Effects of an increase in the complexity of innovation.	19
1.3	Cost of innovation as a function of n	20
1.4	Cost of innovation in the limiting economy.	21
1.5	Expected revenues for inputs. $n=5$	24
1.6	Probability of final innovation. $n=5$	25
1.7	Increasing complexity and optimal patent strength	26
1.8	Increasing complexity and optimal patent strength for imperfect substitutes ($\varepsilon = 0.1, \bar{K} = 0.4$)	27
1.9	Increasing complexity and optimal patent strength: the case of symmetric patent policy	28
1.10	Cooperative and non-cooperative pricing.	29
2.1	Comparison of equilibria.	44
2.2	Gains from not joining the patent pool. $\varepsilon = 0.1$	45
2.3	Socially optimal innovation.	47
2.4	Comparison of expected welfare.	48
2.5	Probability of innovation and patent length.	53
2.6	Probability of innovation as a function of patent length.	54
3.1	Patent length and research effort	70
3.2	Patent length and Welfare	71
3.3	Size of the economy and optimal patent length	74
3.4	Degree of substitutability and optimal patent length	75
3.5	Uniqueness of equilibrium - Infinite patent length	78
3.6	Uniqueness of equilibrium - Infinite patent length	79

Introduction

This thesis analyzes the effect of Intellectual Property Rights (IPRs), in the form of both patents and copyrights, on innovation.

Patents are intended to enhance private investment in R&D through the monopoly power they grant to the innovator over the commercial exploitation of her invention. This vision is essentially static, with patents increasing the expected revenues of a one-stand innovator.

Sometimes innovations are sequentially linked, so that for instance the invention of the radio would have been impossible without the previous discovery of electromagnetic waves. In this case protecting the first innovator provides her with a claim over a part of second innovator's revenues. At this point the policy maker must be careful, because patents in this context play a double role: on one hand they increase the expected revenues of the innovator, but on the other hand they produce an additional cost for the innovator, as she has to pay a license fee to the previous innovator. If protection to the first innovation is too strong, it might be the case that the second innovation is not profitable. On the other hand if protection is too scant, expected revenues from the first innovation might be inadequate to provide incentives to invest in its discovery. This problem has been well understood and studied in depth by the literature on sequential innovation, pioneered by Scotchmer (1991).

The three chapters of this thesis analyze a more general version of this problem, where the number of previous innovators with a claim on the second generation inventions are more than one, something very common in high-tech industries. Going back to the case of radio, it was not only electromagnetic waves that radio was based upon, but also high-frequency alternator, high-frequency transmission arc, magnetic amplifier, the crystal detector, diode and triode valves, directional aerial, etc. In the words of Edwin Armstrong (inventor of FM radio) "it was absolutely impossible to manufacture any kind of workable apparatus without using practically all of the inventions which were then known". Patents on these previous inventions produced as many claims over the commercialization of radio. Shapiro (2001) refers to such a situation as a *patent thicket*. There are many cases like this one in hi-tech industries like software, hardware, biotechnology and electronics. For example, in the 1980s IBM accused Sun Microsystems of infringing some of its 10,000 software patents¹; development of golden rice required access to 40 patented products and processes (Graff, Cullen, Bradford, Zilberman, and Bennett 2003); and there are 39 patent families "potentially relevant in developing a malaria vaccine from MSP-1" (Commission on Intellectual Property Rights 2002).

When a patent thicket arises the innovator must pay license fees on many patented previous discoveries. Someone claims that this is reducing innovation in high-tech sectors. Heller and Eisenberg (1998) were the first to suggest that reduction in innovation activity would have stemmed from what Heller (1998) defines the *tragedy of the anticommons*. This phenomenon, specular to the tragedy of the common, arises when too many agents have rights of exclusion over a common, scarce resource, and as a consequence the common resource is under-utilized. The tragedy of the anticommons in hi-tech sectors can be interpreted as follows: too many patentees have exclusive claims on separate components of the state-of-the-art technology, placing an obstacle to future research. In this thesis I analyze the effect of patents in sequential and complex

¹<http://www.forbes.com/asap/2002/0624/044.html>

innovation, and study whether the formation of anticommons is a theoretical possibility.

In chapter 1, joint with Gastón Llanes, we set up a partial equilibrium game theoretical model where an innovator uses N patented inputs to produce a new good. We allow the elasticity of substitution between the N inputs to vary between zero (perfect complements) to infinity (perfect substitutes). We ask what happens as the number of previously patented inputs, for which the innovator has to pay a license fee, increases. We find that the answer depends on the degree of substitutability of the inputs: when inputs are complements the total cost for the innovator increases in N , eventually reducing the probability of innovation to zero. On the other hand, when inputs are substitutes, increasing N exacerbates competitive pressure, reducing the total cost of the license fees. In the basic model, the research inputs have been already invented, and are protected by strong patents. The government could reduce the patent thicket by granting less patents, or by reducing patent breadth. However, weaker patents imply a reduced incentive to discover research inputs in the first place. Therefore we extend the model to include a fixed cost of inventing the inputs and study the optimal patent policy. Patent policy affects the division of profits between the input innovators and the final innovator. We obtain the patent policy that maximizes the joint probability of inventing all the inputs and the final good. We find that this optimal policy is decreasing in complexity: when the number of inputs required in the R&D process increases, the optimal response is to reduce patent strength.

Chapter 2, also joint with Gastón Llanes, is a dynamic version of chapter 1. In this chapter we analyze a sequence of innovations $n = 1, 2, \dots$. Innovation n cannot be introduced until innovation $n - 1$ has been introduced. Each innovation has a commercial value (the profit it generates as a final good), which is random and private information of the innovator, and a deterministic cost of R&D to be developed.

Our model provides a good description of the innovation process in several industries. For example, in the software industry the first programs were written from scratch, and therefore built on little prior knowledge. As more and more programs were developed, they progressively became more dependent on technologies introduced by the first programs. According to Garfinkel, Stallman, and Kapor (1991), nowadays software programs contain thousands of mathematical algorithms and techniques, which may be patented by the innovators who developed them. Similar examples can be found in other hi-tech industries.

Formally, our model is a multi-stage game in discrete time with uncertain end. Interestingly, the probability of reaching the next period is determined endogenously. Our theoretical contribution is to present a simple dynamic model that can obtain closed form solutions for the sequence of probabilities of innovation. The equilibrium concept we use is Subgame Perfect Equilibrium with Markovian Strategies (Markov Perfect Equilibrium).

We are interested in determining equilibrium dynamics under three scenarios: patents, no patents and patent pools. The novel aspect of our model is that patents generate cumulative claims on the sequence of innovations. We find that with patents, innovation becomes harder and harder with more complex innovations. The probability of innovation goes to 0 as $n \rightarrow \infty$. The probability of innovation is higher than in the static case, but not enough to stop the tragedy of the anticommons from happening.

In the no patents case, on the other hand, the probability of innovation is constant and depends on the degree of appropriability of the commercial value of the innovation in the final goods sector. The no patents case will provide higher innovation than the patents case unless the innovator can appropriate a very small fraction of the value of the innovation.

When ideas are protected by patents, the formation of a patent pool increases the probability of innovation for all innovations. Interestingly, the probability of innovation with a pool is constant and higher than what it would be in the static case. This result strengthens the findings of Shapiro (2001), Lerner and Tirole (2004), and Llanes and Trento (2009) for static models. The comparison between the pool and the no patents case depends once again on the degree of appropriability of the value of the innovation in the latter case.

We also find that pools are dynamically unstable: the temptation to remain outside the pool increases as the sequence of innovations advances. This means that early innovators have more incentives to enter the pool than later innovators. In this model patents are perfect complements and a deviation from the pool, while privately profitable is socially undesirable. We find that the patent pool outcome can be replicated by a scheme in which each innovator buys all patent rights from the preceding innovator, instead of paying only for the permission to use the idea (innovator 1 sells all the rights over innovation 1 to innovator 2, who sells all the rights over innovations 1 and 2 to innovator 3 and so on). This means that the complete sale of patent rights will generate higher innovation than licensing. This scheme may be difficult to implement when the nature of innovations is difficult to describe ex-ante, in which case patent pools would be easier to enforce. An alternative scheme, leading to the same innovation outcome, is to allow subsequent competition between the licensee and the original licensor. This removes the monopoly power of all but the last patent, eliminating the anticommons effect.

We find the optimal innovation policy that maximizes the expected welfare of the sequence of innovations. We find that innovation is suboptimal in the three policy regimes. In the no patents regime, there is a dynamic externality: innovators do not take into account the impact of their decision on the technological possibilities of future innovators. In the two other policy regimes, the inefficiency stems from asymmetric information and market power: patent holders do not know the exact value of the innovation, but they know its probability distribution. This generates a downward sloping expected demand for the use of their ideas, and market power implies a price for old ideas above marginal cost.

Finally chapter 3 is a dynamic general equilibrium model of endogenous growth, where growth is the result of quality improvement of the final good. In order to obtain this quality improvement, firms must invest in R&D. The R&D success rate follows a stochastic poisson process, meaning that the time of innovation is uncertain. In this context patent length plays a central role: by the time R&D takes place, patents on some of the research inputs might have expired. Thus, depending on patent length, a fraction of the research inputs might be sold at a competitive price. The larger this fraction, the lower - *ceteris paribus* - the cost of R&D. But reducing patent length also reduces expected revenues from the innovation. This trade off is the central theme of the paper.

I find that the infinite patent length always provides a less than optimal research effort in the steady-state equilibrium. Then I perform numerical sim-

ulations of the model under a regime of finite patent length. These simulations display an hump-shaped level of R&D effort on patent length, suggesting the existence of a finite optimal patent length. This optimal patent length increases the equilibrium level of innovation relative to a system of infinitely lived patents. Also in this case, though, the steady state level of innovation is always below the one a benevolent planner would choose. In other words, a varying patent length is not a good policy instrument to achieve the socially optimal rate of innovation. I use numerical simulations to study the characteristics of the optimal patent length. These simulations suggest that optimal length is decreasing in the size of the economy. This could be the result of the scale effect: the larger the economy the lower the opportunity cost of investing in R&D, and therefore the higher the growth rate of the economy. Or it could be the result of the fact that the larger the economy, the larger the profits of the innovator at each moment in time, and therefore the lower the need for long patents. Or a combination of the two. In the paper I argue that the second one is the only sensible explanation. This is in line with the results of the static analysis of Boldrin and Levine (2005a) and it might be relevant in the context of the recent discussions about introducing an IPRs system in developing countries. It would suggest that, if the current patent length in advanced countries is optimal, enforcing IPRs in developing countries should be accompanied by a reduction of patent length in developed countries.

Numerical simulations also suggest that the optimal patent length is increasing in the degree of substitutability between the research inputs used in R&D, reinforcing the results from the static model of chapter 1, and calling for a weaker patent protection in sectors characterized by a higher complementarity in R&D.

Chapter 1

Anticommons and Optimal Patent Policy in a Model of Sequential Innovation

Gastón Llanes and Stefano Trento

1.1 Introduction.

Knowledge builds upon previous knowledge. This is true for most innovations nowadays, especially in hi-tech industries like molecular biology, plant biotechnology, semiconductors and software. In some cases, the innovation consists of an improvement of an older version of the same good. In other cases, the research leading to the discovery of the new good depends on the access to research tools, techniques and inputs which are previous innovations themselves.

In any case, innovation activity will in general depend on the access to previous innovations. Depending on the structure of the patent system, some of these inventions will be protected by patents. This means that patents affect not only the revenues of the innovator, but also the cost of performing an innovation.

Recent concern has arisen on the possibility that patents (or other kinds of Intellectual Property) can restrict access to research inputs, hindering innovation as a consequence. The innovator and the owners of patents on previous inventions share the revenues of the innovation. As the number of inputs needed in research increases, the innovator faces a *patent thicket* and is threatened by the possibility of *hold-up*, namely the risk that a useful innovation is not developed because of lack of agreement with the patent holders. This problem has been dubbed the *tragedy of the anticommons* (Heller 1998, Heller and Eisenberg 1998). When too many agents have exclusion rights over the use of a common resource, this resource tends to be underutilized, in clear duality with the tragedy of the commons in which too many agents hold rights of use and the resource tends to be overused.

This problem may be particularly acute in biomedical research, where there is a deep controversy over the patenting of gene fragments and research tools. Take for example the case of the MSP1 antigen (Plasmodium Falciparum Merozoite Specific Protein 1), widely recognized as the most promising candidate for an anti-malarial vaccine. A study of the Commission on Intellectual Property Rights (2002) found more than 39 patent families covering DNA fragments, methods for processing fragments, production systems, vaccine delivery systems, etc. As a consequence, a potential innovator willing to commercialize a vaccine based on MSP1 must get prior permission from the owners of these property rights.

Anticommons can arise in biotechnology as well. A good example is Golden Rice, which required payment of up to 40 licenses, depending on the country of commercialization (Graff, Cullen, Bradford, Zilberman, and Bennett 2003).

As a final example, consider the case of software patents, which cover mathematical algorithms and techniques. Software programs have become so complex that any single program may use thousands of algorithms (Garfinkel, Stallman, and Kapor 1991), possibly infringing a large number of patents. This explains the expected increase in patent litigation in this sector in the next years (think of Microsoft vs. the programmers and users of Linux), and the formation of a Patent Commons by firms involved in the Open Source community (IBM, HP, Novel, Sun, etc.).

We address these issues by constructing a model of sequential innovation in which an innovator uses n patented inputs to develop a new invention. Substitutability between the inputs goes from zero (perfect complements) to infinity (perfect substitutes) and the input sellers compete in prices but do not know the exact value of the innovation for the innovator.

We study how the probability of performing the innovation changes as technologies become more complex (n increases) and find that it decreases when the inputs are market complements and increases when they are market substitutes. Therefore, we prove that the anticommons hypothesis may hold when inputs are essential and not easy to substitute. Then, we analyze the limiting economy when $n \rightarrow \infty$, and show that the probability of innovation is always less than socially optimal unless the inputs are perfect substitutes. Moreover, the probability of innovation goes to zero when the elasticity of substitution is below a threshold level which is higher than 1.

In the basic model, the research inputs have been already invented, and are protected by strong patents. The government could reduce the patent thicket by granting less patents, or by reducing patent breadth. However, weaker patents imply a reduced incentive to discover research inputs in the first place. In Section 1.4, we turn to the analysis of optimal patent policy. Patent policy affects the division of profits between the input innovators and the final innovator. We obtain the patent policy that maximizes the joint probability of inventing all the inputs and the final good. We find that this optimal policy is decreasing in complexity: when the number of inputs required in the R&D process increases, the optimal response is to reduce patent strength.

There is an extended literature on Sequential Innovation, which is mainly concerned with the optimal division of profits between sequential innovators. The main contribution of this paper is to extend the analysis of patent policy and optimal division of profit when innovation activity requires access to several prior inventions. As such, the two key insights of the model are: (i) patents will be more harmful when research inputs are complementary, and (ii) optimal patent policy should decrease when innovations become more complex.

As we discuss in the following section, our paper is also related to the literatures of complementary monopoly and patent pools. A patent pool is a cooperative agreement among patent holders, through which they agree on the licensing terms of a subset of their patents. In Section 1.5, we analyze the effects of the formation of patent pools on innovation activity within our model, and study similarities and differences of our findings with previous papers in this literature.

1.1.1 Related literature.

There is an extended literature on Sequential Innovation (Scotchmer 1991, Green and Scotchmer 1995, Chang 1995, Scotchmer 1996, Hopenhayn, Llobet, and Mitchell 2006), which is mainly concerned with the optimal division of profits between sequential innovators. Generally, in these models, there are two innovations that have to be introduced sequentially (the second innovation cannot be introduced until the first one has been introduced), and the objective is to find the patent policy that maximizes the incentives to invest in both innovations. In this paper we generalize these models by assuming that any innovation is based on n previous innovations. We find that the effect of increasing complexity on innovation activity depends on the degree of substitutability/complementarity between these previous innovations.

Our contribution is to extend the analysis of optimal patent policy and optimal division of profits to the case in which the innovator requires access to *several* prior inventions. In doing so, we link the literature of sequential

innovation to the literatures of complementary monopoly and patent pools.

Cournot (1838) was the first to analyze the complementary monopoly problem. He modeled a competitive producer of brass who must buy zinc and copper from two separate monopolists (zinc and copper are perfect complements), and showed that (i) the price of the inputs is higher than the price that a single provider would set, (ii) the total cost of the inputs is increasing in the number of inputs, (iii) in the limit, as $n \rightarrow \infty$ the cost of the inputs is such that the demand for the final good is zero.

This analysis has been later extended in various directions. Sonnenschein (1968) showed that the Cournot's theory of complementary monopoly is the dual of its theory of duopoly with quantity competition and homogeneous goods, when the marginal cost of production is zero. Bergstrom (1978) built upon this intuition to extend the duality result to imperfect substitutes. Recently, Boldrin and Levine (2005b) applied the structure of complementary monopoly to a model of sequential innovation: they set up a model of one innovation building upon a number of (perfect complementary) previous innovations, and show that as this number increases the cost of innovating also increases, eventually hitting the upper bound for which the innovation is no longer profitable. We extend the analysis to a framework of oligopoly with any degree of substitutability. This generalization is relevant because it allows us to study the joint effect of the number of oligopolists and the degree of substitutability on the probability of innovation. In particular we find that, when the research inputs are market complements (i.e. not only when they are perfect complements), the probability of innovation decreases in the complexity of innovation.

Cournot's theory has also been used by the literature on patent pools. Shapiro (2001) was the first to suggest that patent pools may be anticompetitive when they are formed by substitute patents, and pro-competitive when formed by complementary patents. Lerner and Tirole (2004) built a model to generalize Shapiro's results for imperfect substitutability. Given the similarity between our model and Lerner and Tirole's, in Section 1.5 we present an analysis of the effect of patent pools on innovation. We find that Lerner and Tirole's results generalize to our case also. However, given that in our model input decisions are continuous, we can be more precise in our definition of complementarity, which complements previous analysis.

However, it is important to remark that, even though our model is related with Lerner and Tirole's, our focus is completely different. We are interested in analyzing the effect of increases in n on the optimal patent policy, while they are interested in the effects of price collusion on innovation and welfare for n fixed.

1.2 The model.

There are n research inputs (x_1, \dots, x_n) and a potential innovator who may use the n inputs in R&D in order to invent a new good. The n inputs have already been invented and are ready to be produced. We make this assumption in order to concentrate on the effects of the pricing of old ideas on innovation activity. In Section 1.4 we will extend the model to see what happens when the n inputs have to be invented as well.

The structure of Intellectual Property Rights is such that each input is pro-

tected by a patent, granting its owner a monopoly over it. Each patent is owned by a different patentee and thus each of the n inputs is supplied by a different producer. Given that the inputs are imperfect substitutes of each other, the factor market is a differentiated goods oligopoly. The input sellers compete in prices and the value of the innovation is private information of the innovator.

1.2.1 Technology.

The innovator can perform R&D to invent a new final good according to the following CES technology:

$$y = A \left(\sum_{i=1}^n x_i^\rho \right)^{\frac{1}{\rho}}, \quad (1.1)$$

where y is a measure of the R&D effort, A is a scale parameter, x_i is the amount of input i used, n is the number of inputs and $\rho \in (-\infty, 1]$ is a technological parameter related to the substitutability between the inputs.

We will find easier to work with the elasticity of substitution $\sigma = 1/(1 - \rho)$. As is well known, $\sigma = 0$ represents perfect complements, $0 < \sigma < \infty$ represents imperfect substitutes, and $\sigma \rightarrow \infty$ represents perfect substitutes.

The innovator faces an indivisibility problem, meaning that a minimum amount of R&D effort is required to invent a new good. When the R&D effort is below that threshold level there is no innovation. Without loss of generality we can set the threshold level at 1, so that the indicator function for the innovation is:

$$I = \begin{cases} 1 & \text{if } y \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

We set the scale parameter A in (1.1) equal to $n^{(\rho-1)/\rho}$ in order to eliminate any returns from specialization or division of labor. Usually CES production functions exhibit a property called increasing returns to specialization (or love for variety in the case of utility functions). Following an argument similar to Romer (1987), suppose that the production function is $y = (\sum_{i=1}^n x_i^\rho)^{1/\rho}$, and let X be the total quantity of inputs used in production. Because of symmetry, all inputs will be used in the same quantity in the equilibrium, so $x_i = X/n$ for all i , and output will be equal to $y = n^{(1-\rho)/\rho} X$. There are positive returns to specialization because an increase in n holding X constant causes output to increase. We eliminate this effect by introducing $A = n^{(\rho-1)/\rho}$ in the production function.

The complexity of the innovation is measured by n . More complex technologies use a larger number of components or require more research tools in order to be developed. Each input is produced with a constant marginal cost of $\varepsilon > 0$. We assume that the resources used to produce inputs are sold in a competitive market, so that the private and the social cost of producing inputs coincide. The assumption of no returns to specialization guarantees that the social cost of performing the innovation does not change as technologies become more complex. In other words, there is no technological advantage or disadvantage from increases in n .

In section 1.2.4 we will provide the intuition behind our innovation technology. Before we do it, let us continue with the description of the model.

1.2.2 Value of the innovation and structure of the information.

The social value of the innovation, v , is the total surplus generated by the new product. To focus on the factor market, we will assume that the innovator is a perfect price discriminator in the final goods market. This means that the revenue of the innovator coincides with the social value of the innovation.

The value of the innovation is private information of the innovator. This may be because the innovator has better information about the characteristics of the new product or about the valuation of the consumers. The sellers of inputs only know that v has a cumulative distribution $F(v)$. Therefore $F(v)$ is the probability that the innovation has a return less or equal to v . In Section 1.6 we show that the assumptions of perfect price discrimination and asymmetric information can be relaxed without altering the results.

The hazard function is defined as $h(v) = f(v)/(1 - F(v))$, where $f(v)$ is the density function corresponding to $F(v)$. In order to guarantee the quasi-concavity of the maximization problem of the input producers, the following assumption will hold throughout the paper:

Assumption 1 (Nondecreasing hazard function) $h(v) > 0$ and $h'(v) \geq 0$ on a support $[\underline{v}, \bar{v}]$, and $h(v) = 0$ outside of this support.

This assumption on the hazard function is very general, and holds for most continuous distribution functions. Notice that we are not restricting \underline{v} nor \bar{v} to be of finite value.

An important assumption is that the distribution of values of the innovation does not change with n . This assumption, together with the absence of returns to specialization in the R&D technology imply the following lemma:

Lemma 1 *The probability that an invention is socially optimal does not depend on its complexity.*

The probability that an innovation is socially optimal is the probability that its social value is larger or equal than its social cost. The social cost of an innovation coincides with the resources used to produce it. Therefore, the probability that an innovation is socially optimal is $Prob(v - \sum_{i=1}^n \varepsilon x_i \geq 0)$. Because of the symmetry in the innovation technology, $x_i = 1/n$, so this probability becomes $1 - F(\varepsilon)$, which depends on the distribution of social values of the innovation and the marginal cost of the inputs but *not* on the number of inputs used in R&D.

In this paper, we are interested in studying the effects of increasing technological complexity on the probability of innovation. Lemma 1 assures that a change in n affects this probability only through a change in the number of inputs used in research, but not through a change in the social value or cost of the innovation. In other words, we want to compare innovations with different n but the same net social value. In Section 1.6 we relax these assumptions by letting the value of the innovation be a function of n and allowing returns to specialization in the R&D technology. We find that the main results of the paper are not significantly affected by a change in these assumptions.

1.2.3 Market interaction.

The players of the game are the n input sellers and the innovator. A strategy for input seller i is a choice of price for her input. A strategy for the innovator is a function $g : \mathbb{R}_+^n \times v \rightarrow \mathbb{R}_+^n$, namely a demand x_i for each input, as a function of the price of all the inputs and the realization of the value of the innovation.

The timing of the game is as follows: (i) the input producers simultaneously set the price of their inputs, (ii) Nature extracts a value v from the distribution $F(v)$, and (iii) given prices, the innovator calculates the input mix that minimizes the cost of innovation and decides whether to innovate or not.

The equilibrium concept we use is Symmetric Subgame Perfect Equilibrium (SSPE). A set of strategies $\{p_i\}_{i=1}^n, g$ is a SSPE if it is a Nash equilibrium of every subgame of the original game, and $p_i = p$ for all i .

The payoff for input producer i is $x_i(p_i - \varepsilon)$ and the payoff of the innovator is $Iv - \sum_{i=1}^n p_i x_i$.

Innovator's Problem.

Given input prices $\{p_i\}_{i=1}^n$, the innovator solves the following Cost Minimization Problem (CMP):

$$\begin{aligned} c &= \min \sum_{i=1}^n p_i x_i \\ \text{s.t.} \quad & n^{-\frac{1-\rho}{\rho}} \left(\sum_{i=1}^n x_i^\rho \right)^{\frac{1}{\rho}} \geq 1. \end{aligned}$$

The solution to this problem is the set of conditional factor demands x_i and the minimum cost of innovation c . Given c , the innovator will perform the innovation ($I = 1$) if $v \geq c$.

Input Seller's Problem.

When setting the price the sellers of inputs do not know the realization of v . They only know that given $\{p_i\}_{i=1}^n$ the probability that $v \geq c$ (the probability of innovation) is $1 - F(c)$. Therefore, the expected demand of input firm i is $E(x_i) = (1 - F(c)) x_i$, and its Profit Maximization Problem (PMP) is:

$$\max_{p_i} \Pi_i = (1 - F(c)) x_i(p_i - \varepsilon),$$

where both c and x_i come from the CMP of the innovator.

1.2.4 Interpretation of the innovation technology

The CES specification of the R&D technology is a simple and general way to introduce substitutability and complementarity between the inputs used in research. In our model, ideas have economic value because they are embodied in physical objects (Romer (1990), Boldrin and Levine (2002, 2005c)). The innovator uses these physical objects to innovate, not the abstract ideas. Accordingly, the input decision is not discrete (to use the idea or not), but rather continuous (the research inputs can be used in variable amounts).

Our description of the R&D process is a good description of many innovation technologies. First, we can think that the inputs are used in R&D in variable amounts, and once the innovation is performed they are no longer needed. This interpretation fits well for sectors that use a large number of research tools, like biomedical research, where the use of clones and cloning tools, laboratory equipment and machines, reagents, computer software and many other research instruments is required in R&D, and can be used in variable amounts.

The second interpretation is that inputs are actually components of the final innovation, and are used to produce each copy of it. This interpretation resembles more the case of already patented code lines used in new software, hardware components for computers, and a variety of cases in electronics, semiconductors and other similar industries. In the early radio industry, for example, according to Edwin Armstrong (inventor of FM radio) “it was absolutely impossible to manufacture any kind of workable apparatus without using practically all of the inventions which were then known”, like the high-frequency alternator, high-frequency transmission arc, magnetic amplifier, the crystal detector, diode and triode valves, directional aerial, etc.

Under this second interpretation there is a continuum of perfectly competitive innovators. The production function of output is $y = (\sum_{i=1}^n n^{\rho-1} x_i^\rho)^{1/\rho}$ and perfect competition assures that the price of the final output is equal to its marginal cost $c = \sum_{i=1}^n p_i x_i$. Then, the demand of the final good is $y = 1 - F(c)$, where y and c are the quantity and price of the final good.

The two interpretations lead to exactly the same results, except that in the second one there is a welfare loss from the anti-competitive pricing in the inputs market (the welfare loss would be approximately equal to $(c-\varepsilon)(F(c)-F(\varepsilon))/2$). For expositional purposes, we will stick to the first interpretation.

Obviously, our research technology is not a good description of all possible innovation processes. In particular, it does not fit the case in which the innovator pays only for the permission to use an idea. An elegant alternative formulation is to consider discrete input choices (1 if the input is used and 0 otherwise). This is the approach followed by Lerner and Tirole (2004). In this case, the innovator uses either embodied or abstract ideas, and pays input innovators to have access to these ideas.

In Section 1.5 we show that our model can be interpreted as a “smooth” version of Lerner and Tirole’s. Plainly, the advantage of our approach is tractability. Continuity allows us to use differential calculus, which greatly simplifies our subsequent analysis of the effect of an increase in the number of inputs. The best argument to show that our results would still hold if we were to assume discrete input choices is to notice the similarity of the conclusions of both models on the effect of patent pools when n is fixed.

1.3 Equilibrium.

In this section we solve recursively for the SSPE. Therefore, we begin by solving the Innovator’s Problem (second stage of the game). The demands are those of a typical CES production function. The conditional demand of input i and the

cost of innovation are:

$$\begin{aligned} x_i &= I n^{-\frac{1}{1-\sigma}} p_i^{-\sigma} \left(\sum_{i=1}^n p_i^{1-\sigma} \right)^{\frac{\sigma}{1-\sigma}}, \\ c &= n^{-\frac{1}{1-\sigma}} \left(\sum_{i=1}^n p_i^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \end{aligned}$$

where $\sigma = 1/(1 - \rho)$ is the elasticity of substitution between the inputs. The innovator will introduce the new good ($I=1$) if $v \geq c$.

The restrictions on ρ imply that the elasticity of substitution σ goes from 0 (perfect complements) to ∞ (perfect substitutes).

Given x_i and c , the symmetric equilibrium price p solves

$$p = \operatorname{argmax}_{p_i \geq \varepsilon} (1 - F(c)) x_i (p_i - \varepsilon),$$

where $c = n^{-\frac{1}{1-\sigma}} (p_i^{1-\sigma} + (n-1)p^{1-\sigma})^{\frac{1}{1-\sigma}}$ and $x_i = n^{-1} p_i^{-\sigma} c^\sigma$. It is useful to notice that in the symmetric equilibrium ($p_i = p$ for all i), $c = p$ and $x_i = 1/n$ for all i . Also, $p \geq \varepsilon$ in equilibrium because otherwise firms would be making negative profits and would find it profitable to deviate by setting a higher price.

Because of the nature of Nash equilibria, for any value of n , ε , and $\sigma < \infty$ there exists equilibria where p is so high that the probability of innovation is zero (i.e. profits are zero for all input sellers) but any deviation by a *single* input seller is not enough to make it positive. However, these are trivial equilibria coming from the definition of Nash equilibria without any intrinsic economic value. We are interested in the existence of equilibria with a positive probability of innovation ($p < \bar{v}$).

The following proposition characterizes the solution of the first stage of the game (the Input Seller's Problem).

Proposition 1 *A SSPE with positive probability of innovation ($p < \bar{v}$) exists and is unique. The equilibrium price solves*

$$(p - \varepsilon) h(p) = n - \sigma(n - 1)(p - \varepsilon)/p.$$

The conditional input demand is $x = 1/n$, the cost of innovation is $c = p$ and the probability of innovation is $1 - F(p)$.

The firm wants to maximize $(1 - F(c)) x_i (p_i - \varepsilon)$. The derivative with respect to price is:

$$D(p_i) = -f(c) \frac{\partial c}{\partial p_i} x_i (p_i - \varepsilon) + (1 - F(c)) \left(\frac{\partial x_i}{\partial p_i} (p_i - \varepsilon) + x_i \right).$$

By Shepard's Lemma $\partial c / \partial p_i = x_i$, and by symmetry $c = p$, $x_i = 1/n$ and $\partial x_i / \partial p_i = -(n-1)\sigma / (n^2 p)$. Therefore, the first order condition becomes:

$$D(p) = -f(p) \frac{p - \varepsilon}{n^2} + (1 - F(p)) \left(-\frac{\sigma(n-1)(p - \varepsilon)}{n^2 p} + \frac{1}{n} \right).$$

Now we prove that the solution cannot be ε nor \bar{v} for $n < \infty$. $p = \varepsilon$ cannot be the equilibrium because $D(\varepsilon) = (1 - F(\varepsilon))/n > 0$. Also, $p = \bar{v}$ cannot be the

equilibrium both if \bar{v} is finite or infinite. If $\bar{v} < \infty$, then $D(\bar{v}) = -f(\bar{v}) \left(\frac{\bar{v}-\varepsilon}{n}\right) < 0$. On the other hand, $\lim_{p \rightarrow \infty} D(p) = -\infty < 0$. Therefore, the solution must satisfy $D(p) = 0$. Multiplying $D(p)$ by $-n^2/(1-F(p))$ we get:

$$h(p)(p-\varepsilon) + \sigma(n-1)\frac{p-\varepsilon}{p} - n = 0. \quad (1.2)$$

We can be sure that equation (1.2) has exactly one solution because it is continuously increasing in p by Assumption 1, is negative when $p = \varepsilon$ and is positive when $p \rightarrow \bar{v}$ (Assumption 1 implies that $\lim_{p \rightarrow \bar{v}} h(p)p = \infty$ for finite or infinite \bar{v}). Therefore, the solution exists and is unique. Rearranging terms in equation (1.2) we get the desired result.

Example. We will find it useful to illustrate the results with the help on an example based on the uniform distribution. This example has the advantage of providing an explicit solution for the equilibrium price. Specifically, assume that the value of the innovation, v , is uniformly distributed between 0 and 1. This means that $F(v) = v$ and $h(v) = 1/(1-v)$. The equilibrium price is:

$$p = \frac{a + \sqrt{a^2 + 4\sigma\varepsilon(n-1)b}}{2b}, \quad (1.3)$$

where $a = n + \varepsilon - \sigma(n-1)(1+\varepsilon)$ and $b = 1 + n(1-\sigma) + \sigma$. The cost of innovation is equal to the price and the probability of innovation is simply $1-p$.

1.3.1 Elasticity of substitution.

The price of the inputs and the cost of innovation in equilibrium depend on the elasticity of substitution, the complexity of the innovation and the marginal cost of the inputs. In the following subsections we will analyze the comparative statics of the above equilibrium.

Proposition 2 *The cost of innovation is decreasing in σ .*

Equation (1.2) provides an implicit function of p in terms of σ . We can calculate $\partial c/\partial\sigma$ using the implicit function theorem (remember that $p=c$ in the symmetric equilibrium):

$$\frac{\partial c}{\partial\sigma} = -\frac{(n-1)(p-\varepsilon)/p}{h(p) + h'(p)(p-\varepsilon) + \sigma(n-1)\varepsilon/p^2}$$

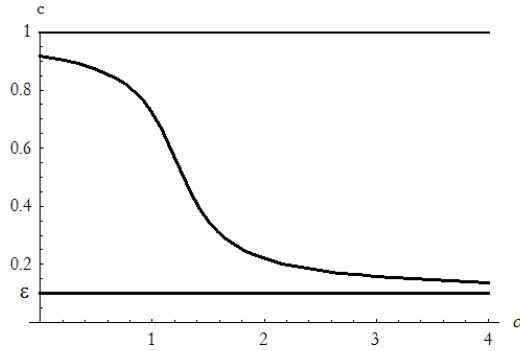
It is easy to see that this derivative is always negative (the numerator and the denominator are positive). The result follows.

Figure 3.2 depicts the cost of innovation (i.e. the price of the inputs) as a function of σ for the uniform distribution and for $n = 10$ and $\varepsilon = 0.1$.

The cost of innovation is monotonically decreasing in σ because of increased competition as the inputs become more substitutable. As $\sigma \rightarrow \infty$ price converges to marginal cost ε , which is the standard Bertrand price competition result with homogeneous goods.

1.3.2 Complements and Substitutes.

We will classify inputs in market complements and substitutes according to the sign of the cross-price derivative of expected demand which, in this setting,

Figure 1.1: Cost of innovation as a function of σ .

is equivalent to analyzing the cross-price derivative of expected profit. This classification is equivalent to the one used in game theory, where the actions of two agents are said to be complements (substitutes) when an increase in the action of one of them implies a decrease (increase) in the payoff of the other agent. In our model, the actions are prices and the payoff is expected profit. Notice that this is an equilibrium definition since it is based on the best response of the innovator.

Definition 1 (Market complements and substitutes) *Input j is a market complement (substitute) of input i if $\frac{\partial E(x_i)}{\partial p_j} < 0$ ($\frac{\partial E(x_i)}{\partial p_j} > 0$).*

An increase in the price of input j has two effects on the expected demand of input i . On one hand, the conditional demand of input i increases (substitution effect). On the other hand, the probability of innovation decreases because the inputs are more expensive to the innovator (innovation effect). The sign of the cross-price derivative depends on which of the two effects is stronger. The cross-price derivative is:

$$\frac{\partial E(x_i)}{\partial p_j} = (1 - F(c)) \frac{\partial x_i}{\partial p_j} + \frac{\partial(1 - F(c))}{\partial p_j} x_i.$$

The first effect is related to the standard substitution effect of consumer demand theory. Remember that the Cost Minimization Problem is equivalent to the Expenditure Minimization Problem and in this case there are no wealth effects of price changes (the conditional factor demands are equivalent to Hicksian demands). In principle, the derivative $\partial x_i / \partial p_j$ could be positive, negative or zero. However, the property of negative semidefiniteness of the matrix of cross-price derivatives (which implies that every input must at least have one technical substitute), together with the symmetry of the production function, implies that this derivative is non-negative. The inputs will be technical substitutes ($\partial x_i / \partial p_j > 0$) except in the case of perfect complements, where $\partial x_i / \partial p_j = 0$.

The second effect is due to the fact that the demand for inputs is downward sloping in the cost of innovation. The cost of the inputs used in research affects the profitability of innovation. Therefore, an increase in the price of any input

will lower the probability of innovation. This effect is negative, except in the case of perfect substitutes, when it is zero.

Now that our definition of complementarity and substitutability is clear, we can be precise in our exposition. In what follows when we say that inputs are complements or substitutes, we mean that they are *market* complements or substitutes. We will see that the distinction between complements and substitutes is crucial for the predictions of the model.

The following lemma shows the value of σ that makes the cross-price derivative equal to zero.

Lemma 2 *The cross-price derivative $\partial E(x_i)/\partial p_j$ is zero in the symmetric equilibrium if and only if $\sigma = \sigma^*$, where σ^* is the argument that solves $h\left(\frac{\sigma}{\sigma-1}\varepsilon\right) = \sigma - 1$.*

The cross-price derivative is:

$$\frac{\partial E(x_i)}{\partial p_j} = (1 - F(c)) \frac{\partial x_i}{\partial p_j} - f(c) x_i.$$

By Shepard's Lemma, $\partial c/\partial p_j = x_j$. Imposing symmetry, $x_i = x_j = 1/n$ and $\partial x_i/\partial p_j = \sigma/(n^2 p)$. Rearranging terms we get:

$$\frac{\partial E(x_i)}{\partial p_j} = \frac{1}{n(1 - F(p))} \left(\frac{\sigma}{p} - h(p) \right).$$

This will be zero in the symmetric equilibrium only when $h(p) = \sigma/p$. Introducing this into the first order condition (1.2) and rearranging we get $(p-\varepsilon)/p = \sigma^{-1}$ or $p = \sigma\varepsilon/(\sigma - 1)$. Plugging this value of p in $h(p) = \sigma/p$ we get the desired result.

The following proposition classifies inputs in complements and substitutes according to Definition 1. It is interesting to see that this distinction depends on the values of σ and ε , but not on the value of n .

Proposition 3 *In the symmetric equilibrium, inputs are complements when $\sigma < \sigma^*$ and substitutes when $\sigma > \sigma^*$.*

We know from Lemma 2 that the cross-price derivative is zero when $\sigma = \sigma^*$ and that its sign depends on $\sigma/p - h(p)$. The latter expression is increasing in σ because p is decreasing in σ from Proposition 2 and h is non-decreasing in p from Assumption 1. The result follows.

Interestingly, the value of σ which divides inputs in complements and substitutes has to be larger or equal than 1. To see this, suppose that $\sigma^* < 1$. This means that $h(\varepsilon\sigma^*/(\sigma^* - 1)) < 0$, which is not possible. In the case of the uniform distribution, for example, inputs are complements when $\sigma < (1 + \varepsilon)/(1 - \varepsilon)$ and substitutes when $\sigma > (1 + \varepsilon)/(1 - \varepsilon)$.

1.3.3 Increasing complexity.

Proposition 4 shows that the sign of the effect of an increase in the complexity of the innovation (n) depends on whether the inputs are complements or substitutes.

Proposition 4 *The cost of innovation increases as innovation becomes more complex if the inputs are complements ($\sigma < \sigma^*$) and decreases if the inputs are substitutes ($\sigma > \sigma^*$).*

We are looking for the effect of a unit increase in n , but it will suffice to determine the sign of $\partial c/\partial n$. Equation (1.2) provides an implicit function of c in terms of n . Therefore, we can calculate $\partial c/\partial n$ using the implicit function theorem:

$$\frac{\partial c}{\partial n} = \frac{1 - \sigma(p - \varepsilon)/p}{h'(p)(p - \varepsilon) + h(p) + \sigma(n - 1)\varepsilon/p^2}$$

We know that the denominator is always positive. Therefore, the sign of this derivative depends on the sign of the numerator.

From equation (1.2) we get the following relation in equilibrium $\sigma(p - \varepsilon)/p = (n - h(p)(p - \varepsilon))/(n - 1)$. Introducing this in the numerator and operating, it becomes $(h(p)(p - \varepsilon) - 1)/(n - 1)$. We know that $h(p)(p - \varepsilon) = 1$ when $\sigma = \sigma^*$ from the proof of Proposition 2.1. Given that $h(p)(p - \varepsilon)$ is increasing in p , it is decreasing in σ . Therefore, the numerator is positive when $\sigma < \sigma^*$ and it is negative when $\sigma > \sigma^*$. The result follows.

The probability of innovation is simply $1 - F(c)$, so it moves in an opposite direction to the cost:

$$\frac{dPr}{dn} = -f(c) \frac{dc}{dn}.$$

As before, the effect on the probability of innovation of an increase in the complexity of innovation depends on the substitutability between the inputs. If inputs are complements, then the probability decreases as n increases. If inputs are substitutes, then the probability increases as n increases.

Figure 1.2 shows what happens in the uniform distribution example as the complexity of the innovation increases from $n = 5$ to $n = 15$, for $\varepsilon = 0.1$. The cost schedules cross when $\sigma = 1.22$, which is exactly $\sigma^* = (1 + \varepsilon)/(1 - \varepsilon)$. This means that the cost of innovation increases if the inputs have low substitutability and decreases in case of high substitutability.

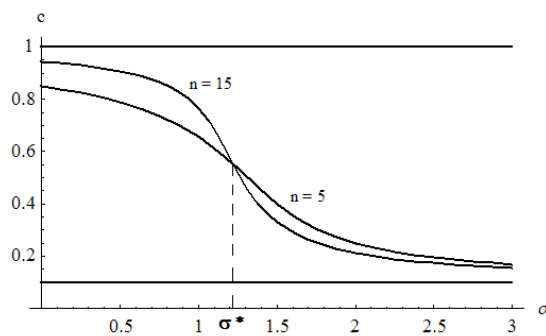


Figure 1.2: Effects of an increase in the complexity of innovation.

Proposition 4 says that patents are very harmful when innovation is sequential and the research inputs are essential or difficult to substitute, but do not pose an important problem when inputs are easily replaceable.

Figure 1.3a shows cost as a function of n for complementary inputs and $\varepsilon = 0.1$ in the case of the uniform distribution. As innovation becomes more complex, the cost of innovation increases and converges to 1 when $n \rightarrow \infty$. This means that the probability of innovation decreases with n and converges to 0. Convergence is faster when σ gets closer to zero. When the substitutability between the inputs is very low (σ close to zero), the probability of innovation is very small even for simple innovations (low n).

Figure 1.3b shows that the conclusions change when the research inputs are substitutes. In this case the cost of innovation decreases when the complexity of innovation increases (i.e. the probability of innovation increases with n).

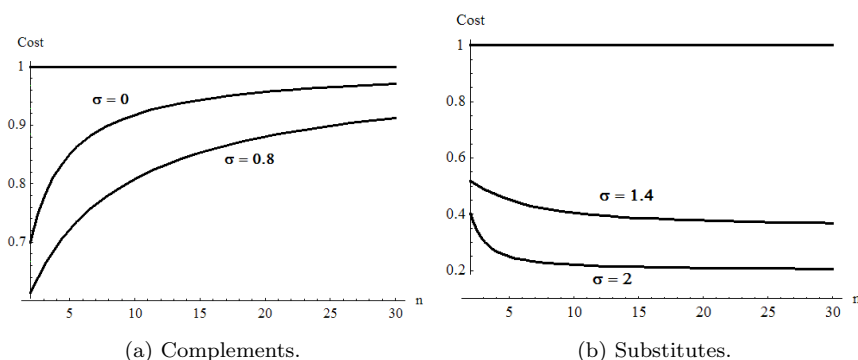


Figure 1.3: Cost of innovation as a function of n .

1.3.4 High complexity and Monopolistic Competition.

It is interesting to analyze the equilibrium of the economy when $n \rightarrow \infty$ for two reasons. First, $n \rightarrow \infty$ represents innovations that are highly complex and therefore require a *large number* of inputs to be developed. The innovator faces a patent thicket and has to gather inputs from many patentees. We know how the probability of innovation changes as n increases, but it is interesting to determine in what cases it will go to 0 or $1 - F(\varepsilon)$. Second, in this limiting economy there is an infinite number of input sellers, so the effect of a price change by a single firm has an infinitesimal impact on the cost of innovation, and the market becomes monopolistically competitive.

Proposition 5 characterizes equilibria with positive probability of innovation ($p < \bar{v}$). In this case there are values of σ for which there is no equilibrium with positive probability of innovation.

Proposition 5 *A SSPE with $p < \bar{v}$ exists only when $\sigma > \hat{\sigma}$ where $\hat{\sigma} = \frac{\bar{v}}{\bar{v} - \varepsilon}$. The equilibrium price and cost of innovation are $p = \frac{\sigma}{\sigma - 1} \varepsilon$.*

Dividing the first order condition (1.2) by n , we get:

$$h(p)(p - \varepsilon) \frac{1}{n} + \sigma \left(\frac{n-1}{n} \right) \left(\frac{p - \varepsilon}{p} \right) - 1 = 0.$$

As $n \rightarrow \infty$, the term with the hazard function goes to zero. This is because each firm becomes negligible and does not affect the probability of innovation

on its own. It is clear that the equilibrium price of the limiting economy solves:

$$\sigma \left(\frac{p - \varepsilon}{p} \right) - 1 = 0.$$

Therefore, $p = \frac{\sigma}{\sigma-1} \varepsilon$, which is between ε and \bar{v} only when $\sigma > \bar{v}/(\bar{v} - \varepsilon)$.

It is interesting to comment on three characteristics of the equilibrium. First, any $p \geq \bar{v}$ is an equilibrium for any value of σ in this limiting economy. If $p \geq \bar{v}$ the probability of innovation is zero, but if a single input seller deviates, its impact on the cost of innovation is infinitesimal, so the probability of innovation (i.e. expected profits) remains unchanged. Therefore, there are no profitable deviations when $p \geq \bar{v}$.

Second, the equilibrium quantity x_i goes to zero as $n \rightarrow \infty$. This is because the number of inputs is increasing towards infinity but the total quantity of inputs required is keeping constant, given our assumptions on the innovation technology.

Finally, it is easy to show that $1 \leq \hat{\sigma} < \sigma^*$. The first inequality follows trivially from the fact that $\hat{\sigma} = \bar{v}/(\bar{v} - \varepsilon)$. Therefore, $\hat{\sigma} = 1$ only when $\bar{v} \rightarrow \infty$ or $\varepsilon = 0$. For the second inequality, it is enough to compare the equilibrium price when $\sigma = \hat{\sigma}$ with the equilibrium price when $\sigma = \sigma^*$, since price is decreasing in σ . When $\sigma = \hat{\sigma}$, price is equal to \bar{v} . When $\sigma = \sigma^*$ we know that the equilibrium price solves $h(p)(p - \varepsilon) = 1$. If $p = \bar{v}$, then $h(p)(p - \varepsilon) \rightarrow \infty$, which is much larger than 1. For $h(p)(p - \varepsilon)$ to decrease and approach 1, p has to decrease. This means that equilibrium price is larger with $\hat{\sigma}$ and therefore $\hat{\sigma} < \sigma^*$.

Figure 1.4 shows the cost schedule as a function of σ when $\bar{v} = 1$ and $\varepsilon = 0.1$. The equilibrium of the limiting economy does not depend on the distribution of v , but it depends on the upper bound of the support of the distribution.

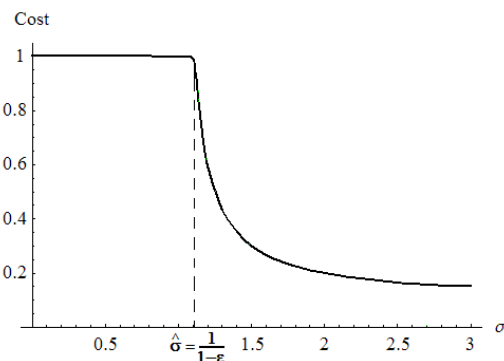


Figure 1.4: Cost of innovation in the limiting economy.

The equilibrium price is the same than Dixit and Stiglitz's (1977) monopolistic competition model. When inputs are substitutes, firms set a mark-up over marginal cost equal to $1/(\sigma - 1)$. This means that the pricing inefficiency decreases as n increases, but it does not disappear even when $n \rightarrow \infty$.

For complements, the outcome depends on whether σ is greater or less than $\hat{\sigma} = \bar{v}/(\bar{v} - \varepsilon)$. When $\sigma > \bar{v}/(\bar{v} - \varepsilon)$, firms set a mark-up just like in the substitutes case. When $\sigma \leq \bar{v}/(\bar{v} - \varepsilon)$, the only equilibria have $p \geq \bar{v}$ and so the

probability of innovation is zero. In this case, as n increases the inefficiency due to monopoly pricing increases and it is at its maximum when $n \rightarrow \infty$.

1.3.5 The Tragedy of the Anticommons Revisited.

The model presented in this paper gives a formal treatment to the tragedy of the anticommons in sequential innovation. An anticommon (Heller 1998) arises when multiple owners have the right to exclude each other from using a scarce resource, causing its inefficient under-utilization. This problem is symmetric to the tragedy of the commons, where multiple owners have the right to use a scarce resource, but nobody has exclusion rights and resources tend to be overused.

Heller and Eisenberg (1998) have raised the question whether the tragedy of the anticommons may apply to innovation in biomedical research. They point out that in this sector excessive patenting of research tools might reduce the incentive to innovate, because the innovator has to face a possibly high cost of bundling all the licenses together. This is a concern that might be shared by other hi-tech sectors where innovation follows a similar cumulative process.

The model presented above predicts that anticommons may arise in sequential innovation, but only under certain circumstances. In our model, the scarce resource is the state-of-the-art technology for innovation, over which *each* patent holder of a research input has a claim. Each patent holder decides the selling price of her input. It is interesting to notice that when $\sigma \leq 1$ all the inputs are essential to perform the innovation so all the input sellers can potentially impede the innovation by setting a high price.

Proposition 4 shows that, when the inputs are market complements, the cost of producing an innovation increases as technologies become more complex. This proposition says that when the number of patented and complementary research tools to be used in R&D increases, the probability of innovation is reduced: that is, anticommons applies to sequential innovation when research inputs are market complements. This result holds not only for perfect complementarity between the inputs, but whenever the elasticity of substitution is not sufficiently large to compensate the negative effect of price changes on the probability of innovation. The opposite is true when research tools are market substitutes: in this case there is no anticommons because, as the number of research inputs increases, competitive pressure reduces their price fast enough to reduce also the cost of innovation.

Proposition 5 reinforces the previous result: when the number of patented research tools grows large, and they are highly complementary, the anticommons is so strong that the probability of innovation goes to zero.

Finally, it is important to remark that the anticommons effect arises in the absence of any kind of transaction costs. The anticommons arise as a natural consequence of the uncoordinated market power of the input producers. As we will see in section 1.6.2, asymmetric information is not essential for our results. All we need is a downward sloping expected demand for the inputs.

1.4 Patent Policy

We have shown that strong patents and fragmentation of ownership lead to a low probability of innovation when inputs used in research are complementary. The government could reduce the patent thicket by granting less patents, or by reducing the breadth of the patents. However, weaker patents imply a reduced incentive to discover research inputs in the first place. The problem of the division of profit between sequential innovators has been studied by Scotchmer (1991), Green and Scotchmer (1995), Chang (1995), Scotchmer (1996) and Hopenhayn, Llobet, and Mitchell (2006). In this section we complement existing literature by analyzing the optimal division of profits between sequential innovators when the final innovation requires multiple inputs to be performed. Our objective is to determine the effects of higher complexity on the optimal patent policy.

The innovation technology is the same as the one in section 1.2. For tractability, we focus on the perfect complements case, so the final good can only be introduced only if all inputs are invented. Without loss of generality (for the $\sigma = 0$ case), we assume that once the inputs are invented they can be reproduced at zero marginal cost ($\varepsilon = 0$). Finally, we will concentrate on the uniform distribution case, $v \sim U[0, 1]$, which provides linear demands for the innovation.

An important difference with Section 1.2 is that now the n inputs must be invented at an earlier stage. We assume that there is a fixed (sunk) cost K/n of inventing the inputs. Each input will be introduced if expected revenues are larger than the fixed cost. As is standard in the literature of sequential innovation, the fixed cost of the input sellers is unknown to the policy maker. All the policymaker knows is that the sunk cost has a distribution $K \sim U[0, \bar{K}]$. Therefore, the patent policy cannot depend on the realization of K .

Our assumptions imply that the social cost and value of the innovation do not change as n increases. All that changes is the number of input producers with which the final innovator has to agree to perform her innovation.

We introduce a patent policy parameter $\phi \in [0, 1]$. This policy parameter can be interpreted as the patent breadth, the novelty requirement, or the strength with which Intellectual Property law is enforced in courts: there is a probability ϕ that the input innovator is granted a patent and that the patent can be successfully defended in court. We also assume that imitation is costless: Bertrand competition will ensure that inputs not protected by patents are sold at marginal cost. Consequently the innovator only has to pay a non-competitive license fees for inputs that are covered by patents, and can buy the rest at marginal cost.

We will first analyze the case in which patent policy affects only input innovators (first stage innovations). In this section therefore we abstract from the fact that in some cases the same patent policy should apply to the final (second stage) innovation. This is standard in the literature on sequential innovation where the focus is on the effect of patents on the optimal division of profits between first and second stage innovators. For completeness, in Section 1.4.2 we will analyze what happens when patents apply equally to first and second stage innovations (symmetric patent policy). We will show that the qualitative results do not change.

The timing of the game is as follows: (i) input innovators observe the sunk cost of innovation, and decide whether to invent their input or not, (ii) if all

inputs are invented, Nature decides which inputs are protected by patents, (iii) patent holders set a price for their patented inputs, (iv) Nature decides the value of the innovation, (v) the final innovator decides whether to innovate or not.

Research inputs will be invented only if the expected revenue from selling the input is higher than the sunk cost of inventing it. Expected revenue depends on whether they are granted a patent and on how many other patents are granted. Remember that each input innovator is granted a patent with probability ϕ . Suppose m patents are granted in the second stage. Then, the price of the inputs, the cost of innovation and the probability of innovation are, respectively: $p_m = \frac{n}{m+1}$, $c_m = \frac{m}{m+1}$ and $Pr_m = \frac{1}{m+1}$.

Consider an input innovator who is granted a patent. In the third stage, her revenues depend on how many other patents have been granted. Let $k = m - 1$ denote the number of patents granted, in addition to the patent of the input innovator we are considering. Actual revenues (after uncertainty is resolved) are $\Pi_k = 1/(k + 2)^2$. Expected revenues are:

$$E(\Pi) = \phi \sum_{k=0}^{n-1} \left(\frac{1}{k+2} \right)^2 \frac{(n-1)!}{(n-1-k)!k!} \phi^k (1-\phi)^{n-1-k} \quad (1.4)$$

There are two effects of increases in ϕ on $E(\Pi)$. First, a higher ϕ increases the probability of being granted a patent, which increases $E(\Pi)$. Second, the increase on ϕ increases the probability that more patents are granted in addition to mine, which decreases $E(\Pi)$ because of the anticommons effect. Our simulations indicate that the first effect always dominates the second effect for n small, so that $E(\Pi)$ are increasing in ϕ . For larger n , though, the second effect dominates the first for large ϕ , so $E(\Pi)$ first increases and then decreases with ϕ . Figure 1.5 shows the expected revenue of input producers as a function of patent strength for $n = 5$.

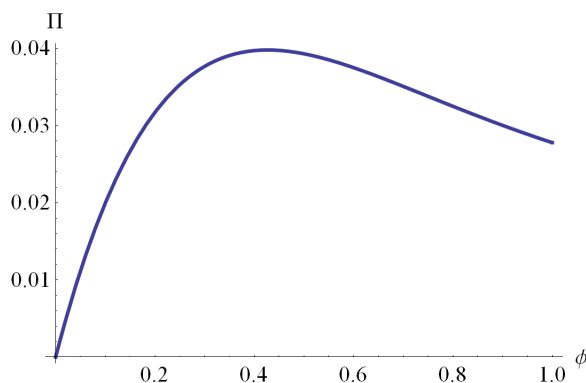


Figure 1.5: Expected revenues for inputs. $n=5$

Let us now focus on how patent policy affects the probability of innovation. In order for the final innovation to be introduced, two things must happen: (i) expected revenues have to be larger than the fixed cost for the input sellers, and (ii) the value of the innovation for the final innovator has to be larger than the cost of paying the inputs protected by patents. The probability that (i) happens

is $Pr(E(\Pi) > K/n) = nE(\Pi)/\bar{K}$. The expected probability that (ii) happens is:

$$E(Pr(v > c)) = \left(\sum_{m=0}^n \frac{1}{m+1} \frac{n!}{(n-m)!m!} \phi^m (1-\phi)^{n-m} \right).$$

It is easy to see that $E(Pr(v > c))$ is decreasing in ϕ . This is because, assuming the inputs are invented, a higher ϕ increases the probability that more patents are granted, which implies a lower probability of introducing the final innovation due to the anticommons problem.

The probability of introducing the final innovation is simply the product of the previous two probabilities: $Pr = Pr(E(\Pi) > K/n) E(Pr(v > c))$.

$$Pr = \frac{nE(\Pi)}{\bar{K}} \left(\sum_{k=0}^n \frac{1}{k+1} \frac{n!}{(n-k)!k!} \phi^k (1-\phi)^{n-k} \right) \quad (1.5)$$

Figure 1.6 shows that the probability increases with respect to ϕ , reaches a maximum (for the optimal policy ϕ^*), and then decreases, for $n = 5$. There are two effects pulling in opposite directions: on one hand increasing ϕ increases the probability of inventing the inputs, on the other hand it increases the cost of the final innovator.

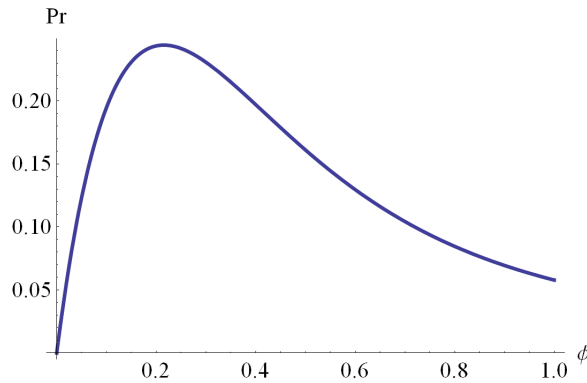


Figure 1.6: Probability of final innovation. $n=5$

Interestingly, the policy that maximizes Pr is smaller than the policy that maximizes $E(\Pi)$. This is because $E(\Pi)$ is always decreasing in ϕ . This has an implication for rent seeking activities: at the optimal policy, input innovators would push for stronger patents on research inputs, and final innovators would push for weaker patents.

The following proposition presents the most interesting result of this section.

Proposition 6 *The optimal patent strength is decreasing in the complexity of the innovation.*

Rather than going through an analytical proof, it suffices to analyze Figure 1.7. This figure shows the policy that maximizes the probability of innovation, as a function of n . Given that the optimal policy does not depend on \bar{K} or any

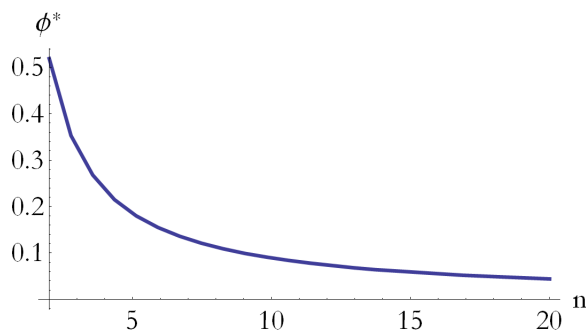


Figure 1.7: Increasing complexity and optimal patent strength

other parameter, this actually shows the effect of increases on n on ϕ^* , which is clearly decreasing.

In models of sequential innovation, the degree of patent protection determines the division of profits between sequential innovators. Stronger patents increase the protection for early innovators, as they grant them a claim over following innovations. When innovation builds on several prior inventions, the uncoordinated market power of earlier patentees generates an anticommons effect. Proposition 6 shows that as the number of claims increase, the anticommons gets worse, and the optimal response is to reduce the degree of patent protection.

It is important to remark that in our model a minimum amount of protection is always needed, otherwise input producers would have no incentives to invent the inputs in the first place. The important result, however, is that the degree of patent protection should decrease as n increases. In section 1.4.2 we will show that the same result holds in a more sophisticated model where patent policy also affect the revenues of the final innovator.

1.4.1 Patent Policy with Imperfect Substitutability

In this section we relax any constraint on the substitutability between research inputs, and allow it to vary between zero and infinity. Any degree of substitutability higher than one requires a positive marginal cost ε for the inputs, otherwise the innovator would only use unpatented inputs and would be able to innovate at zero cost. As before, inputs require a fixed cost of K/n to be invented, with $K \sim U[0, \bar{K}]$, and $v \sim U[0, 1]$.

This setting represents a generalization of the analysis of Section 1.4. However, the problem becomes analytically untractable, and we have to resort to numerical simulations.

The results in this section are in line with those of Section 1.4: increasing complexity reduces the optimal patent strength. Figure 1.8 shows the optimal patent breadth ϕ^* as a function of the elasticity of substitution σ , with $\varepsilon = 0.1$ and $\bar{K} = 0.4$, for $n = 5$ and $n = 10$. Increasing n reduces ϕ^* for any σ , and the result holds for any $n \geq 2$.

Another interesting result is that the optimal patent strength ϕ^* is increasing in σ . This is because substitutability increases competition among input

producers, thus reducing their ability to set a price above the marginal cost. This is equivalent to redistribute revenues from input producers to the final innovator. In order to compensate for this redistribution, and return to its optimal level, patents must be strengthened. Notice however that, for this same reason, patents in this case are less harmful as they provide a more limited market power. Finally, it is interesting to note that there is some value of σ , larger than 1, for which the optimal patent policy is 1, i.e. inputs have to be protected with strong patents when the substitutability is very large.

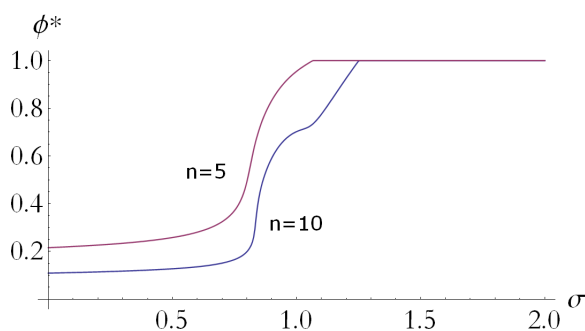


Figure 1.8: Increasing complexity and optimal patent strength for imperfect substitutes ($\varepsilon = 0.1, \bar{K} = 0.4$)

1.4.2 Symmetric Patent Policy.

Here we complete the analysis of section 1.4 by assuming that the patent policy also applies to the final good innovator: there is a probability $\phi \in [0, 1]$ that, if the final innovation is introduced, it will be protected by a patent. The expected revenue of the final innovator is ϕv , where $v \sim U[0, 1]$ is the gross social value of the innovation, just as before. This change introduces an additional factor in favor of strong patents: the expected private value of the innovation ϕv , to be shared between the final innovator and the input producers, it is now increasing in patent strength.

The rest of the model is the same as in Section 1.4, so we focus on the perfect complements case with linear demand for the final innovation. Research inputs need to be invented in a previous stage, with a fixed cost $K \sim U[0, \bar{K}]$, and the necessary condition for this to happen is that expected profits from selling the input $E(\Pi_i) = (1 - F(c/\phi)) x_i p_i$ are higher than this sunk cost of innovation. The timing of the game is also left unchanged.

Interestingly enough, introducing this extension does not change our results. Input producers completely internalize the effect of ϕ on the expected profits of the final innovator, and leave the probability of innovation unaffected. When m inputs receive patent protection, inputs price and cost of innovation become respectively $p = \frac{\phi n}{m+1}$ and $c = \frac{\phi m}{m+1}$. As has been said, the probability of innovation is the same: $Pr(\phi v > c) = \frac{1}{m+1}$.

On the other hand, the revenue of a patent holder when $k = m - 1$ additional patents are granted is now ϕ times lower at $\frac{\phi}{(1+k)^2}$. This obviously has a negative effect on the overall probability of getting the final innovation $Pr = Pr(E(\Pi) >$

$K/n) E(Pr(v > c))$. It also has a positive effect on the *level* of the optimal patent strength (the one that maximizes Pr). Still it does not affect the main result: optimal patent strength is decreasing in the complexity of the innovation, as shown in Figure 1.9

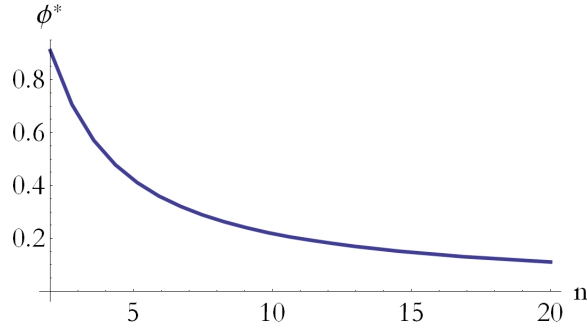


Figure 1.9: Increasing complexity and optimal patent strength: the case of symmetric patent policy

This result reinforces the one in section 1.4, and gives an idea how strong the anticommons effect is as technological complexity increases. The result is even more remarkable when taking into account that, in this model, the final innovator is a perfect price discriminator. This implies that reducing patent strength reduces the expected producer revenue (ϕv) on a one-to-one basis, which we see as a realistic upper bound on the impact of patents on the incentive to innovate.

1.5 Patent Pools

Even though our objective is to analyze the optimal patent policy, it is interesting to analyze what would happen if research inputs were priced cooperatively, either by a collective institution such as a patent pool or by a single patent holder (monopolist) that owns all the patents. This analysis is interesting because the USPTO (US Patents and Trademarks Office) itself has recommended the creation of patent pools to ease the access to biotechnology research tools (Clark, Piccolo, Stanton, and Tyson 2000).

For this analysis, we come back to the basic model of Section 1.2 (the innovator requires n inputs, which are already invented and protected by patents of infinite breadth). Proposition 7 shows the equilibrium price in this case. The difference with the previous case is that now the patent holder maximizes joint-profits and therefore takes into account the cross-price effects between expected demands.

Proposition 7 (Patent Pool) *The equilibrium price when all the inputs are priced cooperatively, p^* , is the argument that solves $h(p)(p - \varepsilon) = 1$.*

Given the symmetric input demands, the pool wants to sell a symmetric bundle. Therefore $x_i = 1/n$ and $p_i = p$ for all i and the pool wants to maximize

total profits $n(1 - F(p))(p - \varepsilon)$. The first order condition is $n(-f(p)(p - \varepsilon) + 1 - F(p)) = 0$. Rearranging terms we get the desired result.

Notice that p^* depends only on the functional form of h and the value of ε , but not on the values of σ or n . The following proposition compares the cost of innovation when the inputs are priced individually, c , with that of a patent pool, p^* .

Proposition 8 *The cost of innovation when the inputs are priced non-cooperatively, c , is equal to that of a patent pool, p^* , when the cross-price derivative is zero ($\sigma = \sigma^*$), it is larger when the inputs are complements ($\sigma < \sigma^*$) and it is smaller when the inputs are substitutes ($\sigma > \sigma^*$).*

We know from the proof of Lemma 2 that when $\sigma = \sigma^*$, the cross-price derivative is zero and $\sigma = p h(p)$. Replacing this in (1.2) and rearranging we get $h(p)(p - \varepsilon) = 1$, which is the cooperative result. Given that p is decreasing in σ , whereas p^* is independent of σ , $p > p^*$ when $\sigma < \sigma^*$ and $p < p^*$ when $\sigma > \sigma^*$.

The difference between cooperative and non-cooperative pricing is that in the first case the firms take into account the effect of an increase in the price of one input on the demand for the rest. When $\sigma = \sigma^*$ this effect is zero so the price of the pool coincides with that of the non-cooperative equilibrium. When $\sigma < \sigma^*$ the effect is negative, so the pool knows that an increase in price will decrease the demand for the rest and will set a price smaller than the uncoordinated input sellers. The opposite happens when $\sigma > \sigma^*$.

In the case of the uniform distribution, the pool price is $p^* = (1 + \varepsilon)/2$. Figure 1.10 compares this price with the non-cooperative price for $\varepsilon = 0.1$ and $n = 5$.

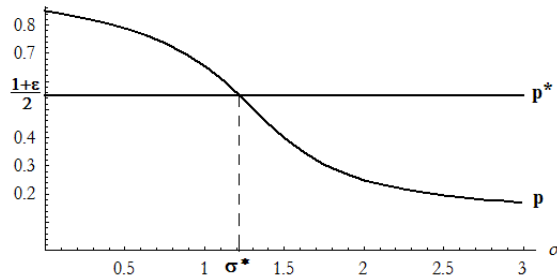


Figure 1.10: Cooperative and non-cooperative pricing.

Our results are similar to those found in Lerner and Tirole (2004). As we discussed in Section 1.2.4, the difference between the two papers is that we have assumed that inputs are used in research in a continuous fashion, while Lerner and Tirole (2004) consider discrete input choices (1 if the input is used and 0 otherwise).

Under the latter approach, the equilibrium will depend on whether the competition margin or demand margin bind. When the competition margin binds, if the input seller raises her price, her input would be evicted from the bundle of patents bought by the final innovator. When the demand margin binds, the

input seller can raise price without excluding its input from the basket, but the overall demand for the bundle would decrease.

These competition and demand margins are related to our substitution and innovation effects. The substitution effect says that an increase in the price of one of the inputs lowers the demand for that input and increases the demand for the rest (holding overall demand constant). The innovation effect says that an increase in the price of an input lowers the overall demand for the basket of inputs (holding the relative demand of the inputs constant).

Therefore, our model can be interpreted as a "smooth" version of Lerner and Tirole's. In Lerner and Tirole, as inputs become more complementary, it is more likely that the demand margin will bind. In our model, both margins always bind (except when $\sigma = 0$ or $\sigma \rightarrow \infty$), but as inputs become more complementary, the innovation effect becomes more important relative to the substitution effect.

In this sense, the contribution of our model is twofold: (i) we show that Lerner and Tirole's results extend to the continuous innovation technology case, and (ii) we explain the effects of patent pools on innovation using the traditional definition of complementarity based on cross-price derivatives (precisely because the effects of price changes are smooth), while Lerner and Tirole base their definition of complementarity on the shape of the revenue function of the innovator.

1.6 Extensions.

In this section we analyze the consequences of relaxing some of the basic assumptions of the model.

1.6.1 Social value and cost depend on complexity.

Until now, we have assumed that the distribution of values of the innovation and the social cost of the inputs do not depend on n , and that there are no returns from specialization. Under these assumptions, a change in n only changes the number of producers from whom the innovator has to buy the research inputs in order to innovate, but does not change the probability that the innovation is socially valuable.

However, it could be argued that the revenues of the innovator or the cost of the inputs are increasing or decreasing in n , or that a higher number of inputs has a positive impact in the R&D technology due to a higher division of labor. All these changes have equivalent effects on the probability of innovation so we will concentrate on changes in the distribution of returns of the innovation.

Let the return of the innovation be $a(n)v$, with $a'(n) \geq 0$ or $a'(n) \leq 0$ and $\lim_{n \rightarrow \infty} a(n) = a_\infty > 0$. v has a cumulative distribution $F(v)$ as before. Notice that we are not setting an upper bound on a_∞ . All we require is that if a is non-increasing it does not go to zero as $n \rightarrow \infty$. This is because if $a_\infty = 0$ then the distribution of values of the innovation will collapse to zero and the innovation will never be profitable when n is very large by assumption.

The probability of innovation is $1 - F(c/a)$, and in the symmetric equilibrium $c = p$ and $x = 1/n$. The equilibrium price of the inputs (i.e. the cost of innovation) solves:

$$(p - \varepsilon) h(p/a)/a = n - \sigma(n - 1)(p - \varepsilon)/p$$

but we are more interested in the ratio $k = c/a$, which enters in the probability of innovation. Replacing in the previous equation we have:

$$(k - \varepsilon/a) h(k) = n - \sigma(n - 1)(k - \varepsilon/a)/k. \quad (1.6)$$

This equilibrium is equivalent to the one in Proposition 1, thinking of $k = c/a$ as the cost of innovation and ε/a as the social cost of the inputs. We can prove the same theorems as before with respect to the difference between complements and substitutes, the welfare effects of patent pools and $\partial k/\partial \sigma$. However, $\partial k/\partial n$ will be different because now ε/a is a function of n .

Using the implicit function theorem on the equilibrium relation (1.6) we get:

$$\frac{\partial k}{\partial n} = \frac{\frac{h(k)(k-\varepsilon/a)-1}{n-1} - \frac{na'}{a} \frac{\varepsilon/a}{k-\varepsilon/a}}{h'(k)(k-\varepsilon/a) + h(k) + \sigma(n-1)(\varepsilon/a)/k^2}$$

As before, the sign of this derivative depends only on the sign of the numerator, but now there is an additional term which shifts the threshold value of σ that divides positive and negative changes in k . This threshold value will be to the left of σ^* when $a'(n) > 0$ and to the right of σ^* when $a'(n) < 0$.

Two important remarks are in order. First, if $a'(n)$ is large then the last term in the numerator will determine the sign of the derivative. In this case, the effect of changes in n on revenues completely overcomes the effect on the pricing of inputs, and $\partial k/\partial n$ has the opposite sign of $a'(n)$ irrespective of the value of σ . Second, even for small $a'(n)$, when $a'(n) > 0$ and $\sigma \rightarrow \infty$ the derivative is always positive. Therefore when $a'(n)$ is small and positive, there are two regions where $\partial k/\partial n$ is positive: one with low values of σ and another with large values of σ .

According to the previous analysis, assuming that the return of the innovation depends on n has an effect on the derivative of the probability of innovation with respect to n . Next, we will show that this assumption has no significant effect on the analysis of the equilibrium as $n \rightarrow \infty$.

The equilibrium price solves:

$$h(p)(p - \varepsilon) \frac{1}{an} + \sigma \left(\frac{n-1}{n} \right) \left(\frac{p-\varepsilon}{p} \right) - 1 = 0.$$

When $n \rightarrow \infty$, the first term will go to zero because $a_\infty > 0$. Therefore, the equilibrium price is the same as before, $p = \frac{\sigma}{\sigma-1} \varepsilon$, which is less than the maximum possible revenue ($a_\infty \bar{v}$) only if $\sigma > a_\infty \bar{v}/(a_\infty \bar{v} - \varepsilon)$. When $\sigma \leq a_\infty \bar{v}/(a_\infty \bar{v} - \varepsilon)$, on the other hand, there is no equilibrium price such that the probability of innovation is positive.

The probability of innovation is $1 - F(p/a_\infty)$. There are two possible cases. If $a_\infty < \infty$ then the probability of innovation is less than optimal, just as in the basic model. If $a_\infty = \infty$ then the probability of innovation will go to 1 for $\sigma > 1$ and 0 for $\sigma \leq 1$, which is the same as assuming $\varepsilon = 0$ in the basic model.

1.6.2 No Asymmetric Information.

Another assumption of the basic model is that there is asymmetric information on the value of the innovation (? , read)for good discussions of why this assumption makes sense]Gallini.Wright.1990,Bessen.2004. However, our results do not

depend on the existence of asymmetric information. All that is needed for the results is a downward sloping demand for innovations.

An alternative interpretation could be that there is a continuum of innovators with decreasing returns to their innovations. Suppose that the innovators are indexed by the return to their innovations, which ranges between \underline{v} and \bar{v} . Now, $F(v)$ is the measure of innovations with a return less or equal to v . Also, assume that the innovations do not compete against each other in the final goods market and that the input sellers cannot price discriminate between the innovators. It is easy to see that all the previous results translate directly into this setting. All that changes is that now $1 - F(c)$ is not the probability of innovation but the measure of innovations performed.

1.6.3 No price discrimination.

We can also relax the assumption that the innovator is a perfect price discriminator. Dropping this assumption introduces a wedge between the social and private values of the innovation. This means that the distribution of values of innovation changes, and that now there is also an inefficiency in the final goods sector. Assume that the social value of the innovation is still distributed according to $F(v)$, with probability density function $f(v)$. The private value of the innovation is now v_p , which is less than the social value of the innovation. With a linear demand for the final good, for example, the private return of the innovation would be $v_p = v/2$, which has a probability density function given by $2f(2v_p)$. The qualitative results are the same as before. All that changes is that now the probability of innovation decreases for each value of σ , and so the values of σ^* , $\hat{\sigma}$ and $\bar{\sigma}$ increase. Also, the optimal patent protection is lower for each value of σ and n than in the case of perfect price discrimination.

1.6.4 Uncertain return of the innovation.

We have also assumed that the innovator is the only one that knows the value of the innovation. In this section we ask what happens if v is also unknown to the innovator. Formally, we do this by changing the timing of the game: (i) the input producers simultaneously set the price of their inputs, (ii) given prices, the innovator calculates the input mix that minimizes the cost of innovation and decides whether to innovate or not, and (iii) Nature extracts a value v for the innovation from the distribution $F(v)$.

We begin by solving the second stage of the game. The innovator decides what would be the optimal combination of inputs to perform the innovation in case he decides to perform it. This leads to the same cost of innovation and conditional demands as before. Then, the innovator decides whether to perform the innovation or not, in order to maximize expected profits $E(v) - c$. The innovation will be performed if $E(v) \geq c$ and will not be performed otherwise. If $E(v) < \varepsilon$, then the innovation will never be performed, so we assume that $E(v) \geq \varepsilon$. We also assume that the innovator will perform the innovation if $E(v) = c$.

The uncertainty has now passed from the input sellers to the input producer. The problem of the input sellers is deterministic, they know $E(v)$ and they know that if the price is higher than $E(v)$ the innovation will not be performed (i.e. their demands will be zero). Now, the inputs are always market substitutes

unless $\sigma = 0$. It is easy to show that the innovation will always be performed, and that the elasticity of substitution only affects the distribution of payoffs between the input sellers and the innovator.

Lemma 3 shows that input demands are discontinuous at a certain price, and Proposition 9 proves that in the symmetric equilibrium $c \leq E(v)$ so the innovation is always performed.

Lemma 3 *Input demands are discontinuous at*

$$p_i = \left(nE(v)^{1-\sigma} - \sum_{j \neq i} p_j^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

The demand for inputs is positive if the cost of innovation is not larger than the expected value of the innovation, that is:

$$n^{-\frac{1}{1-\sigma}} \left(\sum_{i=1}^n p_i^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \leq E(v).$$

Rearranging terms, we get the condition on the price of the input:

$$p_i \leq \left(nE(v)^{1-\sigma} - \sum_{j \neq i} p_j^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

If p_i is larger than this value, then the innovation is not performed and the demand for all inputs is zero.

The input sellers want to maximize profits $x_i(p_i - \varepsilon)$. Proposition 9 states the solution of the game.

Proposition 9 *The equilibrium price when the return of the innovation is uncertain for the innovator is:*

$$p = \begin{cases} \frac{\sigma(n-1)}{\sigma(n-1)-n} \varepsilon & \text{if } \sigma > \frac{n}{n-1} \frac{E(v)}{E(v)-\varepsilon}, \\ E(v) & \text{otherwise.} \end{cases}$$

After imposing symmetry, the derivative of $x_i(p_i - \varepsilon)$ with respect to p_i becomes:

$$D(p) = \frac{1}{n} \left(-\frac{\sigma(n-1)}{n} \frac{p-\varepsilon}{p} + 1 \right).$$

Lemma 3 implies that if the derivative with respect to price is positive at $p = E(v)$, this is a symmetric equilibrium, as firms are making positive profit, do not want to lower price ($D \geq 0$), and would have a zero profit if they would rise price. This happens when $\sigma \leq \frac{n}{n-1} \frac{E(v)}{E(v)-\varepsilon}$.

When $\sigma > \frac{n}{n-1} \frac{E(v)}{E(v)-\varepsilon}$, on the other hand, the equilibrium price solves the unrestricted first order condition $D(p) = 0$.

1.7 Conclusions.

In this paper we extend the literature of sequential innovation in two directions. First, we study how the probability that an innovation is privately profitable changes as technologies become more complex, and the inputs used in research are patented. We find that the results depend on the substitutability between these research inputs.

When the inputs are complements, the profitability of the innovation is decreasing in the technological complexity. In the limit (when $n \rightarrow \infty$), when the degree of substitutability is below a threshold level, which is higher than 1, the innovation is never profitable. This paper therefore gives a formal treatment of the tragedy of the anticommons.

On the other hand, when the inputs are substitutes, the profitability of the innovation is increasing in technological complexity. Even in this case, when $n \rightarrow \infty$, the cost of gathering all the inputs for the innovation is always too high from a social point of view and thus the probability of innovation is suboptimal.

Second, we study the optimal response of patent policy to increasing complexity of innovation. We find that, because of the anticommons effect, the optimal degree of patent protection is decreasing in the complexity of the innovation. The degree of patent protection determines the division of profits between sequential innovators. Stronger patents distribute more revenues from the last innovator to the previous innovators. This result says that increasing complexity of innovation reduces the optimal amount of protection granted to earlier innovators. This result is very robust: numerical simulations suggest that it holds for any degree of substitutability between previous innovations, and also holds when patent protection affects the revenues of the final innovator (i.e patents also affect the amount of profits to be distributed between successive innovators).

These results are at odds with respect to what we observe in the real world: the complexity of technology is increasing but patents are becoming stronger. Not only they have been recently extended to sectors previously lacking protection (sexually reproduced plants, software, business methods, products and processes of biotechnology, including plants and animals). Also patent length has been increasing over the years, and patent systems are being created in countries where they did not previously exist. We think this is a contradiction worth being studied further.

Finally, we also study what happens when inputs are priced cooperatively, either by a collective organization as a patent pool or by a single owner of all the inputs. We find that the cost of the innovation decreases with respect to the non-cooperative pricing, when inputs are market complements, while it increases when inputs are market substitutes. This result is in line with the intuition of Shapiro (2001) and the model of Lerner and Tirole (2004). Still, we believe it complements these previous papers by using a standard definition of complementarity/substitutability, allowing us to exactly identify the economic forces driving the result.

Chapter 2

Patent Policy, Patent Pools, and the Accumulation of Claims in Sequential Innovation

Gastón Llanes and Stefano Trento

[letter,12pt]amsart graphicx harvard, dsfont, subfig
 Patent policy, patent pools, and the accumulation of
 claims in sequential innovatio

2.1 Introduction

Patents are intended to enhance private investment in R&D through the monopoly power they grant to the innovator over the commercial exploitation of her invention. Generally, innovations are sequentially linked. For instance, the invention of the radio would have been impossible without the previous discovery of electromagnetic waves. This sequential nature of innovation introduces the issue of how to divide the revenues of the sequence of innovations among the different innovators. Suppose that two innovations may be introduced sequentially (the second innovation cannot be introduced until the first one has been introduced). If a patent is granted to the first innovator, she may obtain a claim over part of the revenues of the second innovator. The policy maker is confronted with an important trade-off: if the patent covering the first innovation is strong, it may imply that the second innovation becomes unprofitable. On the other hand, if the patent is weak, it may provide low incentives to introduce the first innovation.

This problem has been studied in depth by the literature on sequential innovation, pioneered by Scotchmer (1991). Usually, this literature analyzed the optimal division of profits between two sequential innovators. But what happens when each innovation builds on several prior inventions? Going back to the case of radio, it was not only electromagnetic waves that radio was based upon, but also high-frequency alternator, high-frequency transmission arc, magnetic amplifier, the crystal detector, diode and triode valves, directional aerial, etc. In the words of Edwin Armstrong (inventor of FM radio) “it was absolutely impossible to manufacture any kind of workable apparatus without using practically all of the inventions which were then known”.

In this paper, we extend the literature on sequential innovation by analyzing the case in which patents generate cumulative claims on subsequent innovations. If the number of claims is large, innovators may face a patent thicket and may be threatened by the possibility of hold up, namely the possibility that a socially desirable innovation fails to be performed due to the lack of agreement with previous inventors Shapiro (2001).

Patent thickets are pervasive in hi-tech industries, like software, hardware, biotechnology and electronics. For example, in the 1980s IBM accused Sun Microsystems of infringing some of its 10,000 software patents ¹; development of golden rice required access to 40 patented products and processes (Graff, Cullen, Bradford, Zilberman, and Bennett 2003); and there are 39 patent families “potentially relevant in developing a malaria vaccine from MSP-1” (Commission on Intellectual Property Rights 2002).

When a patent thicket arises the innovator must pay license fees on many patented previous discoveries, which may lead to low innovation. Heller and Eisenberg (1998) were the first to suggest that a reduction in innovation activity would have stemmed from what Heller (1998) defines the *tragedy of the anticommons*. This phenomenon arises when too many agents have rights of

¹<http://www.forbes.com/asap/2002/0624/044.html>

exclusion over a common, scarce resource, and as a consequence the common resource is under-utilized, in clear duality with the tragedy of the commons. In our case, the anticommons could arise if too many patentees have exclusive claims on separate components of the state-of-the-art technology, placing an obstacle to future research. In this paper we build a model to analyze whether anticommons in sequential innovation is a theoretical possibility.

Abstracting from transaction costs, or the possibility that one or more patentees refuse to license their idea therefore blocking innovation, anticommons are similar to complementary monopoly: many monopolistic input providers selling their inputs to a final good producer. The problem of complementary monopoly was first analyzed by Cournot (1838). He modeled a competitive producer of brass who has to use copper and zinc as perfect complement inputs in production. He showed that, when the inputs are sold by two different monopolists, the total cost of producing brass is higher than when the two inputs are sold by the same monopolist. Sonnenschein (1968) showed that complementary monopoly is equivalent to duopoly in quantity with homogeneous goods, and Bergstrom (1978) generalized this result to a general number n of inputs and any degree of complementarity among them.

Recently Shapiro (2001) and Lerner and Tirole (2004) applied the instruments of complementary monopoly to the analysis of patent pools. Their results reinforced the results on complementary monopoly: patent pools (or equivalently a single monopolist owning all the patented inputs) reduces the cost of innovation when patents are complements, and it increases it when they are substitutes. Boldrin and Levine (2005b) and Llanes and Trento (2009) also made use of complementary monopoly to show that, as the number of complementary patents increases, the probability that a future innovation is profitable goes to zero.

All these papers, while making important contributions, present static models. In other words, the first innovation has been already invented, so patents and patent pools only affect the profitability of introducing a second innovation. This introduces an important asymmetry between previous and future innovations. We believe that adding a dynamic dimension is an important step towards a better understanding of the mechanism of anticommons in sequential innovation.

Developing a dynamic model is important for several reasons. First, it will eliminate any bias stemming from the asymmetric treatment of old and new ideas. Second, patent policy will affect not only current but also future innovative activity. Third, it will allow us to analyze the problem of assigning resources to promote innovations with low commercial value (basic research). Fourth, some of the previous findings in the literature will be affected by the introduction of the temporal dimension, and new issues will arise precisely because of this modification.

We present a dynamic model to study the division of profit when each innovation builds on several prior inventions. There is a sequence of innovations $n = 1, 2, \dots$. Innovation n cannot be introduced until innovation $n - 1$ has been introduced. Each innovation has a commercial value (the profit it generates as a final good), which is random and private information of the innovator, and a deterministic cost of R&D to be developed.

Our model provides a good description of the innovation process in several industries. For example, in the software industry the first programs were written

from scratch, and therefore built on little prior knowledge. As more and more programs were developed, they progressively became more dependent on technologies introduced by the first programs. According to Garfinkel, Stallman, and Kapor (1991), nowadays software programs contain thousands of mathematical algorithms and techniques, which may be patented by the innovators who developed them. Similar examples can be found in other hi-tech industries.

Formally, our model is a multi-stage game in discrete time with uncertain end. Interestingly, the probability of reaching the next period is determined endogenously. Our theoretical contribution is to present a simple dynamic model that can obtain closed form solutions for the sequence of probabilities of innovation. The equilibrium concept we use is Subgame Perfect Equilibrium with Markovian Strategies (Markov Perfect Equilibrium).

We are interested in determining equilibrium dynamics under three scenarios: patents, no patents and patent pools. The novel aspect of our model is that patents generate cumulative claims on the sequence of innovations. Patents affect the innovator in two ways: on one hand, the innovator has to pay licensing fees to all previous inventors, but on the other hand, she will collect licensing revenues from all subsequent innovators, in case they decide to innovate. Therefore, it is not clear what is the net effect of patents on innovation as the sequence of innovations progresses.

We find that with patents, innovation becomes harder and harder with more complex innovations. The probability of innovation goes to 0 as $n \rightarrow \infty$. The probability of innovation is higher than in the static case, but not enough to stop the tragedy of the anticommons from happening.

In the no patents case, on the other hand, the probability of innovation is constant and depends on the degree of appropriability of the commercial value of the innovation in the final goods sector. The no patents case will provide higher innovation than the patents case unless the innovator can appropriate a very small fraction of the value of the innovation.

When ideas are protected by patents, the formation of a patent pool increases the probability of innovation for all innovations. Interestingly, the probability of innovation with a pool is constant and higher than what it would be in the static case. This result strengthens the findings of Shapiro (2001), Lerner and Tirole (2004), and Llanes and Trento (2009) for static models. The comparison between the pool and the no patents case depends once again on the degree of appropriability of the value of the innovation in the latter case.

We find that pools are dynamically unstable: the temptation to remain outside the pool increases as the sequence of innovations advances. This means that early innovators have more incentives to enter the pool than later innovators. In this model patents are perfect complements and a deviation from the pool, while privately profitable is socially undesirable. The design of a mechanism to solve the pool instability problem, along the line of Brenner (2009), is beyond the scope of this paper. However we find that the patent pool outcome can be replicated by a scheme in which each innovator buys all patent rights from the preceding innovator, instead of paying only for the permission to use the idea (innovator 1 sells all the rights over innovation 1 to innovator 2, who sells all the rights over innovations 1 and 2 to innovator 3 and so on). This means that the complete sale of patent rights will generate higher innovation than licensing. This scheme may be difficult to implement when the nature of innovations is difficult to describe ex-ante, in which case patent pools would be easier to

enforce. An alternative scheme, leading to the same innovation outcome, is to allow subsequent competition between the licensee and the original licensor. This removes the monopoly power of all but the last patent, eliminating the anticommons effect.

We find the optimal innovation policy that maximizes the expected welfare of the sequence of innovations. We find that innovation is suboptimal in the three policy regimes. In the no patents regime, there is a dynamic externality: innovators do not take into account the impact of their decision on the technological possibilities of future innovators. In the two other policy regimes, the inefficiency stems from asymmetric information and market power: patent holders do not know the exact value of the innovation, but they know its probability distribution. This generates a downward sloping expected demand for the use of their ideas, and market power implies a price for old ideas above marginal cost.

It is interesting to analyze the second best innovation policy, implemented through transfers between innovators when the patent office does not know the value of the innovations. In this case, we find that patents are larger than these transfers, and therefore the patent regime cannot even attain the second best. Finally, we extend the basic model to analyze what happens when the sequence of innovations does not stop after an innovation fails, and to analyze the optimal patent policy when patents have finite length.

Our paper is related to Hopenhayn, Llobet, and Mitchell (2006), which also presents a model of cumulative innovation with asymmetric information. However, the focus of that paper is different. In their paper innovations are substitutes: the introduction of a new product automatically implies the disappearance of old versions in the market. Patents block the introduction of subsequent innovations for the duration of the patent. The question they study is how to allocate monopoly power in the final goods market to successive innovations. A trade-off arises because the promise of property rights to the first innovator limits what can be offered to the second innovator. In our paper, innovations are complementary: all prior inventions are necessary to introduce a new idea. We study what is the effect of intellectual property rights on the pricing of old ideas. The problem is that granting too many rights on sequential innovations implies an increase in licensing fees, potentially hindering innovation as a consequence. Therefore, the two papers offer complementary analysis of the process of sequential innovation when the value of innovations is private information.

2.2 Innovation with patents

There is a sequence of innovations $n = 1, 2, \dots$. Innovation n cannot be introduced until innovation $n - 1$ has been introduced. Formally, the model is a multi-stage game with uncertain end. At each stage, an innovator may introduce an innovation. If the innovation is performed, the game continues and further innovations may be introduced. If the innovation fails to be performed, the game ends and no other innovation can be introduced (we will relax this assumption in Section 2.10). We will see that the probability that the game continues is determined endogenously.

The innovation process is deterministic. At stage n , the innovator may introduce the new idea by incurring in an R&D cost of ε . The new idea is

based on $n - 1$ previous ideas. These previous ideas are protected by patents, which means that the innovator has to pay licensing fees to the $n - 1$ previous innovators (patent holders), in case she wants to introduce the innovation. The cost of innovation is the sum of the cost of R&D and the licensing fees paid to patent holders.

The innovation process reflects the fact that usually earlier innovations do not have a solid background to build upon, while as the market comes to maturity further innovations are more and more indebted to previous ones.

Each idea has a commercial value v_n , which represents the revenues obtained by selling the new product in the final goods market. In order to concentrate on the effects of patents on innovation activity, we will assume that the innovator is a perfect price discriminator in the final goods market, which means that the commercial value of the innovation is equal to the social surplus generated by the new product.

The value of the innovation is private information of the innovator. Patentees only know that v_n is drawn from a uniform distribution between 0 and 1, with cumulative distribution function $F(v_n) = v_n$.

The new idea will be protected by a patent of infinite length, which means that the innovator can request licensing fees from all subsequent innovators (we will relax this assumption in Section 2.11). The total revenues of the innovation equal the commercial value of the innovation plus the future licensing revenues.

The timing of the game within each stage is the following: (i) the $n - 1$ patent holders set licensing fees p_n^i , (ii) Nature extracts a value for v_n from distribution $F(v_n)$, (iii) the innovator decides whether to innovate ($I_n = 1$) or not ($I_n = 0$).

At each stage, patent holders only care about maximizing the expected future licensing revenues. Let J_n^i denote the expected future licensing revenues of innovator i at stage n , given that stage n has been reached. Then,

$$\begin{aligned} J_n^i &= Pr_n p_n^i + Pr_n Pr_{n+1} p_{n+1}^i + \dots \\ &= \sum_{m=n}^{\infty} p_m^i \prod_{k=n}^m Pr_k, \end{aligned} \quad (2.1)$$

where Pr_n is the probability that the n^{th} innovation is introduced, given that all prior innovations have been introduced. Notice that the probabilities Pr_n work as intertemporal discount factors, which arise endogenously from the specification of the model.

J_n^i can also be expressed in a recursive way:

$$J_n^i = Pr_n (p_n^i + J_{n+1}^i), \quad (2.2)$$

This means that with probability Pr_n the innovation is performed, and the patent holder gets the licensing fee from the innovator plus the continuation value of her expected licensing revenues.

The innovator's payoff is $I_n(v_n + J_{n+1}^n - c_n - \varepsilon)$, where $c_n = \sum_{i=1}^{n-1} p_n^i$ is the sum of licensing fees paid to previous innovators.

We will focus on Markov strategies. A strategy for player i specifies an action conditioned on the state, where actions are prices and the state is simply the number of previous innovations introduced. The equilibrium concept is Markov Perfect Equilibrium.

Perfectness implies that future prices will be determined in following subgames, as the result of a Nash equilibrium. Thus, players understand that no action taken today can influence future prices and probabilities. Current actions only affect the probability that the following stage is reached, through the effect of current prices on the probability of innovation. We have just proved the following lemma:

Lemma 4 J_m^i for $m > n$ does not depend on any action taken at stage n .

The game is solved recursively. The solution to the innovator's problem is straightforward. Given v_n and c_n , the innovator forecasts J_{n+1}^n , and decides to innovate ($I_n = 1$) if the revenues from the innovation exceed the cost of innovation:

$$I_n = \begin{cases} 1 & \text{if } v_n + J_{n+1}^n \geq c_n + \varepsilon, \\ 0 & \text{otherwise,} \end{cases} \quad (2.3)$$

which implies that the probability of innovation is $Pr_n = 1 + J_{n+1}^n - c_n - \varepsilon$.

At stage n , patent holders want to maximize their expected licensing revenues from stage n onwards. They know their decisions do not affect J_{n+1}^i (they can only affect the probability that stage $n + 1$ is reached), and decide a licensing fee p_n^i , taking the decision of the other patent holders as given. The patent holder's problem is:

$$\max_{p_n^i} J_n^i = Pr_n (p_n^i + J_{n+1}^i), \quad (2.4)$$

which leads to a price equal to $p_n^i = (1 - \varepsilon)/n$.

Imposing symmetry, $p_n^i = p_n$ and $J_n^i = J_n$ for all i . Replacing prices and probability in J_n , we get $J_n = \left(\frac{1-\varepsilon}{n} + J_{n+1}\right)^2$. Rearranging this equation, $J_{n+1} = \sqrt{J_n} - \frac{1-\varepsilon}{n}$, which is a decreasing sequence, converging to 0 as $n \rightarrow \infty$.

The sequence of probabilities of innovation is:

$$Pr_{n+1} = \left(Pr_n - \frac{1-\varepsilon}{n}\right)^{1/2}, \quad (2.5)$$

which is also a decreasing sequence converging to 0 as $n \rightarrow \infty$. This means that innovation gets harder and harder with more complex innovations (those that are based on more previous innovations).

The probability of innovation decreases with complexity because patent holders do not take into account cross-price effects: patent holder i set the price of her patent by equating the marginal revenue and the marginal cost of increasing her license fee. The marginal revenue is simply the additional revenue in case the new innovation is performed. The marginal cost is the reduction in expected demand, and depends on the fact that - since all patents are essential for the new innovation - increasing the price of patent i decreases the probability of innovation. As a matter of fact, increasing the price of patent i reduces the expected demand for all other inputs, but this effect does not enter the marginal cost for patent holder i . This generates the anticommons effect that closely resembles the tragedy of the commons: patent holders ignore this cross-price effects and set a price that is higher than the price they would set if they coordinated (see section 2.4).

In practice the anticommons can be interpreted as a combination of coordination failure and market power. Each patent produces a claim over part of the

revenues generated by subsequent innovations. Since each patent is essential, and patent holders do not take into account cross-price effects, they set a license fee that is too high. As the number of claims increases, the coordination problem gets worse, until eventually the new innovator is left with negative expected profits with probability one.

2.3 Innovation without patents

Suppose that a policy reform completely removes patents. This change has two effects on innovation. First, the revenues of the innovator in the final goods sector will decrease as a result of imitation. Specifically, assume that the innovator can only appropriate a fraction $\phi \in [0, 1]$ of the consumer surplus generated by the innovation. Second, innovators will not pay licensing fees to previous innovators, nor will they charge for the use of their ideas in subsequent innovations. Therefore, $c_n = 0$ and $J_n = 0$ in the previous model.

At each stage: (i) nature extracts a value of the innovation v_n , and (ii) the innovator decides to innovate or not. The innovator will innovate if $\phi v_n \geq \varepsilon$ and will not innovate otherwise. Thus, the probability of innovation is constant and equal to $1 - \varepsilon/\phi$ if $\phi > \varepsilon$. If $\phi \leq \varepsilon$, then the probability of innovation is zero.

2.4 Patent pools

In this section we analyze what happens when licensing fees are set cooperatively by a collective institution like a patent pool. At each stage, the pool maximizes the future expected revenues of *current* patent holders. The pool will set a symmetric price for all current patent holders. Once an innovation is performed, the innovator becomes a member of the pool in all subsequent stages. In the first stage there is no pool because no innovation has been introduced (the pool plays from stage 2 onwards).

The probability of innovation is $Pr_n = 1 + J_{n+1} - (n-1)p_n - \varepsilon$, and the pool's problem is

$$\max_{p_n} J_n = Pr_n (p_n + J_{n+1}). \quad (2.6)$$

The difference with respect to the non-cooperative case is that the pool recognizes cross-price effects, and therefore is encouraged to set lower prices than in the no-pool case.

A higher J_{n+1} fosters innovation in two ways. First, it increases the future revenues of the innovator. Second, it encourages the pool to set a lower price, because it increases the loss of current patent holders if the sequence of innovations is stopped.

The equilibrium price is $p_n = \frac{1-\varepsilon}{2(n-1)} - \frac{n-2}{2(n-1)}J_{n+1}$, which is equal to the price a pool would set in a static model (see section 2.8.1) *minus* an additional term arising from the pool's concern of keeping future revenues.

The probability of innovation becomes $Pr_n = \frac{1-\varepsilon}{2} + \frac{n}{2}J_{n+1}$. Introducing price and probability in J_n , we get

$$J_n = \frac{1}{n-1} \left(\frac{1-\varepsilon}{2} + \frac{n}{2}J_{n+1} \right)^2. \quad (2.7)$$

This is a first order non-linear difference equation. J_n is decreasing in n and converges to 0 as $n \rightarrow \infty$.

The sequence in terms of probabilities is

$$Pr_n = \frac{1 - \varepsilon}{2} + \frac{1}{2}Pr_{n+1}^2. \quad (2.8)$$

which is a constant sequence such that $Pr_n = 1 - \sqrt{\varepsilon}$ for $n \geq 2$. To determine Pr_1 we need J_2 , which is equal to $(1 - \sqrt{\varepsilon})^2$. Then, $Pr_1 = \min\{1, 2(1 - \sqrt{\varepsilon})\}$.

2.5 Comparison

Figure 2.1 shows the evolution of the probability of innovation in the three cases studied above: infinitely lived patents, no patents and patent pool. The cost of R&D is $\varepsilon = 0.2$ and we consider $\phi = 1$ (full appropriation) and $\phi = 0.3$ (the innovator appropriates 30% of the social surplus generated by the new product) for the no patents case.

Comparing the patent and no-patent cases, we can see that patents increase the probability of the first innovations but decrease the probability of further innovations. The number of innovations for which patents increase the probability depends on ϕ . For example, when $\phi = 1$, patents only increase the probability of the first innovation. Nevertheless, even when $\phi = 0.3$ the probability increases only for the first two innovations. For patents to increase the probability of several innovations, it is necessary that ϕ is very small and close to ε (i.e. when there is very little appropriability without patents).

When ideas are protected by patents, the formation of a patent pool increases the probability of innovation. Figure 2.1 shows that the probability of innovation with patent pools is always larger than the patents case. Moreover, with a pool the probability of innovation does not go to zero as $n \rightarrow \infty$. The comparison with the no patents case depends on ε and ϕ . If ϕ is low, a patent pool increases the probability of all innovations. When ϕ is high, the pool increases the probability of the first innovation, and decreases the probability of all posterior innovations.

2.6 Complete sale of patent rights

The tragedy of the anticommons stems from fragmented ownership of complementary patents. In this case, the probability of innovation decreases as more innovations are introduced, converging to 0 as $n \rightarrow \infty$. The formation of a patent pool would alleviate this problem by concentrating all pricing decisions on one entity. In this section, we discuss a possible alternative solution, which is to enforce the sale of complete patent rights, instead of allowing the sale of individual access rights through licensing fees. These patent rights, in turn, can be resold to other innovators. In this case, innovator 1 would sell the complete patent rights over innovation 1 to innovator 2 for a price φ_1 . Innovator 2 then would sell the patent rights on innovations 1 and 2 to innovator 3 for a price φ_2 , and so on. We will show that this mechanism eliminates the coordination failure at the basis of the anticommons, and that it replicates the innovation outcome under a patent pool.

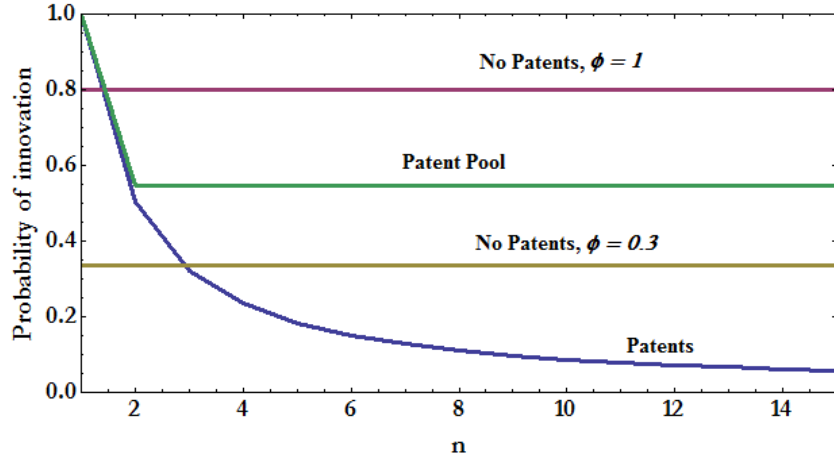


Figure 2.1: Comparison of equilibria.

The cost of innovation n becomes $\varepsilon + \wp_{n-1}$, its expected revenues $v_n + Pr_{n+1}\wp_n$, and the probability that innovation n is performed $Pr_n = 1 - \varepsilon - \wp_{n-1} + Pr_{n+1}\wp_n$. At stage n , innovator $n - 1$ solves the following maximization problem:

$$\max_{\wp_{n-1}} J_n^s = Pr_n \wp_{n-1} \quad (2.9)$$

where J_n^s represent expected revenues of selling the $n - 1$ patent rights to innovator n .

Solving the maximization problem yields a price for patent rights $\wp_{n-1} = \frac{1-\varepsilon}{2} + \frac{1}{2}J_{n+1}^s$, and a sequence of probabilities of innovation $Pr_n = \frac{1-\varepsilon}{2} + \frac{1}{2}Pr_{n+1}^2$ exactly as in the patent pool case. As before, the sequence is $Pr_1 = \min\{1, 2(1 - \sqrt{\varepsilon})\}$, and $Pr_n = 1 - \sqrt{\varepsilon}$ for $n \geq 2$.

This scheme may be difficult to implement when the nature of innovations is difficult to describe ex-ante. For example, when selling the rights over innovation n to innovator $n + 1$, it is difficult to describe what innovation $n + 2$ may be. In this case, complete contracts may be difficult to write, making patent pools easier to enforce.

An alternative policy arrangement leading to the same result would be the following: restoring the possibility of licensing access rights, but at the same time allowing subsequent competition between the licensee and the original licensor. In this case, if innovator n licenses the use of innovation n to innovator $n + 1$, then innovator $n + 2$ can license the use of innovation n from innovators n and $n + 1$. Under this policy arrangement, innovator n will only get positive revenues from the licensing of her innovation to innovator $n + 1$, because at stage $n + 1$ she is a monopolist. After that, she will face competition from other innovators, and Bertrand competition will imply a licensing fee equal to zero.

2.7 Endogenous patent pool formation

In section 2.4 we have assumed that all innovators, after innovating, automatically join the patent pool. Let us now endogeneize this choice, by analyzing the incentives for innovator n to join the pool. In particular let us compare the expected revenues from joining the pool (J_n from section 2.4) with the expected revenues from setting the price of her patent non-cooperatively (J_n^o).

We start with the non-cooperative choice. For expositional clarity let us refer to the patent pool members as *insiders* and to the non-cooperative member as the *outsider*. The pool maximizes the expected revenues of the insiders:

$$\max_{p_n^i} J_n^i = Pr_n (p_n^i + J_{n+1}^i) \quad (2.10)$$

where p_n^i is the cooperative price of any insider's patent, and $Pr_n = 1 + J_{n+1}^i - (n-2)p_n^i - p_n^o - \varepsilon$, with p_n^o denoting the price of the outsider's patent.

On the other hand, the outsider maximizes:

$$\max_{p_n^o} J_n^o = Pr_n (p_n^o + J_{n+1}^o). \quad (2.11)$$

From first order conditions we know that $J_n^o = (n-2)J_n^i$, meaning that if there is one outsider it is convenient to be her. Now let us compare the expected revenues from not joining the pool (J_n^o) to the expected revenues of joining the pool given that everybody else is in the pool (J_n from section 2.4). In equilibrium, deviating from the pool produces an expected revenue of $J_n^o = (3 - \sqrt{1 + 8\varepsilon})/4$, which is constant and only depends on the R&D cost ε . If, on the other hand, innovator n decides to join the pool together with the $n-1$ previous innovators, her expected revenue is $J_n = (1 - \varepsilon)^2/(n-1)$ which is decreasing in n . This is because the patent pool maximizes the joint profits, thus keeping total cost of innovating constant. This constant amount must be divided among an increasing number of insiders, therefore the expected revenue of an insider is decreasing in n . Figure 2.2 shows the gains from deviating from the pool, that is the difference between J_n^o and J_n , as a function of n , for $\varepsilon = 0.1$. For any $\varepsilon \in [0, 1]$ there is a $n^{un}(\varepsilon)$ such that for $n > n^{un}(\varepsilon)$ it is convenient for the innovator to set her price non-cooperatively².

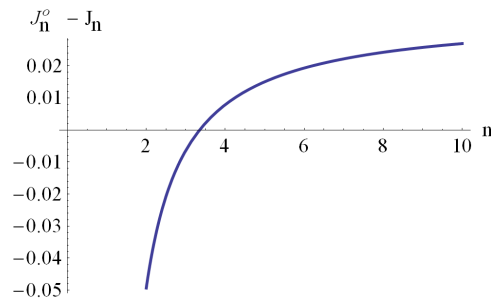


Figure 2.2: Gains from not joining the patent pool. $\varepsilon = 0.1$

²Note that $n^{un}(0) = 5$ and $\frac{\partial n^{un}(\varepsilon)}{\partial \varepsilon} < 0$ for $\varepsilon \in [0, 1]$

Patent pools can improve innovation activity, but are dynamically unstable. Early innovators have more incentives to enter the pool than subsequent innovators. Brenner (2009) finds an elegant mechanism to solve the instability problem of socially desirable patent pools in a static model. We leave the design of an equivalent mechanism in the context of a dynamic model for future research. Without such a mechanism, however, patent pools are likely to be unstable. This might explain why governments in some cases have to enforce the creation of patent pools, as the US government did in the radio and aircraft cases for example.

2.8 Socially optimal innovation

The relevant measure of welfare is the expected social value generated by the sequence of innovations. The social value of an innovation is equal to the increase in consumer surplus minus the cost of the resources spent in R&D. Therefore, at stage n , the social value generated is $v_n - \varepsilon$ if an innovation is performed, and 0 otherwise. Let W be the expected social value. Then,

$$\begin{aligned}
 W &= \sum_{n=1}^{\infty} E(v_n - \varepsilon / I_n = 1) Pr(I_n = 1) + E(0 / I_n = 0) Pr(I_n = 0) \\
 &= \sum_{n=1}^{\infty} E(v_n - \varepsilon / I_n = 1) \prod_{m=1}^n Pr_m \\
 &= \sum_{n=1}^{\infty} \int_{w_n}^1 \frac{v_n - \varepsilon}{1 - w_n} dv_n \prod_{m=1}^n (1 - w_m) \\
 &= \sum_{n=1}^{\infty} \left(\frac{1 + w_n - 2\varepsilon}{2} \right) \prod_{m=1}^n (1 - w_m), \tag{2.12}
 \end{aligned}$$

where w_n is the smallest v_n such that the innovation is performed. In the cases studied above, $w_n = \varepsilon/\phi$ when there are no patents and $w_n = c_n + \varepsilon - J_{n+1}$ with patents or patent pools.

Suppose now that the decision of whether to innovate or not is taken by a centralized agency or social planner. The social planner has to determine $\{w_n\}_{n=1}^{\infty}$, namely what is the minimum value of v_n she would require to perform the innovation at stage n . The planner may decide to perform an innovation even when the realization of v_n is less than ε , if the expected gain from future innovations exceeds the current loss in terms of welfare.

Proposition 10 (Socially optimal innovation) *In order to maximize expected social welfare, the innovation should be performed at stage n if and only if $v_n \geq w_n^*$, where*

$$w_n^* = \begin{cases} 0 & \text{if } \varepsilon \leq E(v_n) = 1/2, \\ \sqrt{2\varepsilon - 1} & \text{if } \varepsilon > E(v_n) = 1/2. \end{cases} \tag{2.13}$$

Because previous decisions are irrelevant once a stage is reached, the social planner's problem at stage n is exactly the same as the problem at stage $n + 1$,

which means that $w_n = w$ for all n . The social planner wants to maximize

$$\begin{aligned} W &= \sum_{n=1}^{\infty} \left(\frac{1+w-2\varepsilon}{2} \right) (1-w)^n \\ &= \left(\frac{1+w-2\varepsilon}{2} \right) \frac{1-w}{w} \end{aligned} \quad (2.14)$$

The first order condition is $-(1+w^2-2\varepsilon)/(2w^2) \leq 0$, with equality if $w \geq 0$. The value that equates the first order condition, $w^* = \sqrt{2\varepsilon-1}$, makes sense only when $\varepsilon \geq 1/2$. On the other hand, when $w \rightarrow 0$, the first order condition converges to $\text{sign}(2\varepsilon-1)\infty$, which means that $w^* = 0$ only if $\varepsilon < 1/2$.

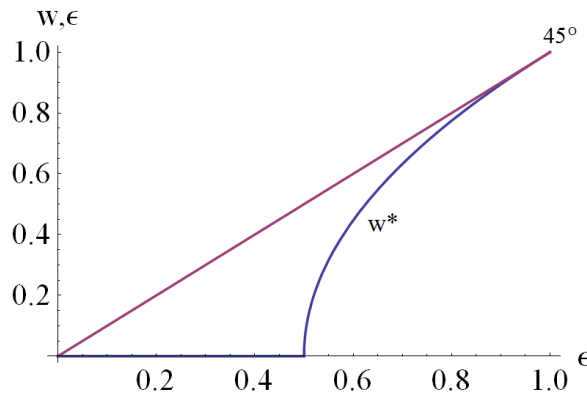


Figure 2.3: Socially optimal innovation.

Proposition 10 implies that innovation will be suboptimal in the three cases studied above, unless $\varepsilon = 0$. There are three reasons why this is so: dynamic externalities, market power and asymmetric information.

The dynamic externality is best described by analyzing the no patents case. Without patents, the innovator will perform the innovation when $v_n \geq \varepsilon/\phi$. Given that $w_n^* \leq \varepsilon$, the innovator may decide not to perform an innovation when it is socially optimal to do so, even if $\phi = 1$. There is a dynamic externality: the innovator ignores the effect of her decision on the technological possibilities of future innovators. This effect is well known in the literature of sequential innovation (Scotchmer 1991, Hopenhayn, Llobet, and Mitchell 2006), and is similar to the one found in the literature of moral hazard in teams (see for example Holmstrom, (1982)), where each agent internalizes only his reward from the effort exerted.

With respect to the patents and patent pool cases, the inefficiency arises from a different source: market power and asymmetric information. Because patentees care about the stream of future licensing revenues, they internalize the effect of today's decision on future innovation. However, asymmetric information implies a downward sloping expected demand for old innovations, and market power implies inefficient pricing of patents, which leads to suboptimal innovation. As the number of holders of rights on innovation increases, the inefficiency due to market power increases.

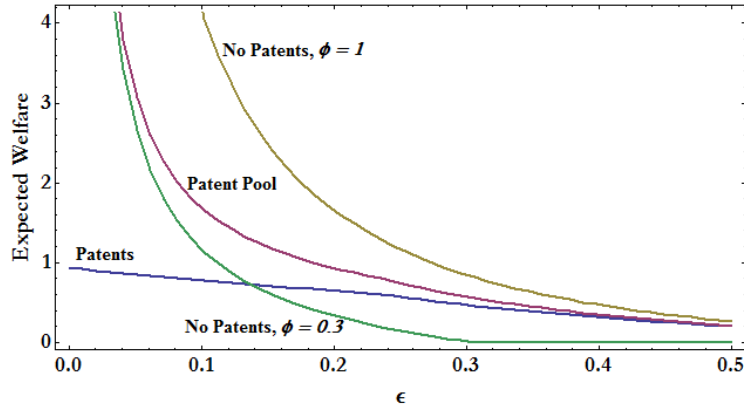


Figure 2.4: Comparison of expected welfare.

2.8.1 Static versus dynamic incentives

Previous models of complementary monopoly, sequential innovation and patent pools were static (Shapiro 2001, Lerner and Tirole 2004, Boldrin and Levine 2005b, Llanes and Trento 2009). It is interesting to ask what changes when we add the dynamic dimension.

To see what happens in the static case, assume only one innovation is being considered. The innovation uses $n - 1$ old ideas, which have already been invented. If the innovation is performed, the innovator obtains a value v from a uniform distribution between 0 and 1, and incurs in a cost ε of R&D. The probability of innovation is $Pr = 1 - \varepsilon - c_n$, with patents or patent pool and $Pr = 1 - \varepsilon/\phi$ without patents.

With patents, the patent holder's problem is to maximize $Pr p^i$, the equilibrium price is $\frac{1-\varepsilon}{n}$ and the probability of innovation is $\frac{1-\varepsilon}{n}$. We have shown that in the dynamic model, the probability is $\frac{1-\varepsilon}{n} + J_{n+1}$, with $J_{n+1} > 0$. This extra term arises because the innovator gets licensing revenues from future innovators. Dynamic incentives imply a higher probability of innovation, but the increase is not enough to prevent the probability from converging to 0 as $n \rightarrow \infty$.

A patent pool would consider cross-price effects, which would lead to a price of $\frac{1-\varepsilon}{2(n-1)}$ and a probability of innovation of $\frac{1-\varepsilon}{2}$. The probability of the corresponding dynamic model is $\frac{1-\varepsilon}{2} + (n-1)J_n$, with $J_{n+1} > 0$. In this case, the extra term arises not only due to the future licensing revenues of the innovator, but also because the pool is concerned with keeping the future licensing revenues of current patent holders.

With respect to the no patents case, the profit-maximizing decision is the same as in the dynamic case. This means that innovators will perform the innovation if $\phi v_n \geq \varepsilon$, which leads to a probability of $Pr = 1 - \varepsilon/\phi$. However, in the dynamic case innovation is suboptimal even when $\phi = 1$, which contrast with the static case, where innovation is socially optimal because there is no intertemporal link between innovations and therefore there is no externality.

2.9 Dynamic externalities and optimal transfers

In the previous section we have shown that sequential innovation is suboptimal because of the presence of dynamic externalities and asymmetric information. Without patents, current innovators do not take into account the effect of their decisions on the innovation possibilities of future innovators. The solution to this problem would require intertemporal transfers between innovators. Patents provide a way to transfer resources from future innovators to current innovators, but we have shown that with patents, market power leads to high licensing fees for old innovations, and therefore to low innovation. In this section we ask how close can the government get to the social optimum when it does not have information on the value of innovations (second best analysis).

To do this, we will use a simplified 2-period version of the general model. In the first period, innovator 1 has the option of introducing an innovation with value v_1 and cost ε . If innovator 1 decides to perform the innovation, in the second period, innovator 2 can introduce an innovation with value v_2 and cost ε . Innovator 1 does not know v_2 .

To determine the social optimum, we have to assume the social planner knows v_1 at stage 1, and v_2 at stage 2. It is likely to think that the government would have reduced information on v_n , but assuming the social planner does not know v_n would imply that innovation decisions without patents give higher welfare than the social optimum, which does not make sense. Later we will analyze government policy, and we will assume that the government does not know v_n .

2.9.1 First best

Let us begin by finding the optimal innovation policy in this 2-stage model. At stage 2 the value and cost of the first innovation are sunk. Therefore, the second innovation should be performed if $v_2 \geq \varepsilon$, and should not be performed otherwise. Consider now the first innovation. The social planner will introduce this innovation if

$$\begin{aligned} v_1 + Pr(v_2 \geq \varepsilon) E(v_2 - \varepsilon | v_2 \geq \varepsilon) &\geq \varepsilon, \\ v_1 + (1 - \varepsilon) \left(\frac{1 - \varepsilon}{2} \right) &\geq \varepsilon. \end{aligned} \quad (2.15)$$

which leads to a probability of innovation $Pr_1^* = \min\{1, \frac{(1-\varepsilon)(3-\varepsilon)}{2}\}$.

Without patents, the probability of introducing the second innovation is $Pr_2 = 1 - \varepsilon$, which is optimal. However, the probability of introducing the first innovation is also $Pr_1 = 1 - \varepsilon$, which is less than optimal. The reason is the same as in Section 2.8: the first innovator does not take into account the effect of her decision on the innovation possibilities of the second innovator.

With patents, innovator 1 sets a licensing fee p_1 to try to extract part of the surplus of innovator 2 (in this 2-period model, the patent and patent pool cases are the same). The probability of innovation of innovator 2 is $Pr_2 = 1 - \varepsilon - p_1$. Innovator 1 maximizes:

$$\max_{p_1} v_1 - \varepsilon + (1 - \varepsilon - p_1) p_1, \quad (2.16)$$

which leads to a price $p_1 = (1 - \varepsilon)/2$. The probabilities of innovation are $Pr_1 = \min\{1, \frac{(1-\varepsilon)(5-\varepsilon)}{4}\}$ and $Pr_2 = (1 - \varepsilon)/2$, so $Pr_1 \leq Pr_1^*$ and $Pr_2 < Pr_2^*$. This is due to the combined effects of asymmetric information and market power.

Therefore, the 2-period model presents a simplified version of the general model but still allows to capture the externality and asymmetric information problems.

2.9.2 Second best: optimal transfers

One way to correct the dynamic externality would be to allow for transfers from future innovators to current innovators. We have seen that patents fail to convey appropriate incentives because of asymmetric information. In this subsection we analyze what is the optimal transfer a government should set to maximize expected welfare when it does not have information on the value of innovations, and we compare it with that of the patents case.

Assume that the government does not know v_1 nor v_2 . In this case, the government cannot make the transfer depend on the realization of v_2 , and it will be impossible to reach the optimum. The best the government can do is to set a transfer equal to t if innovator 2 innovates, and 0 otherwise.

In this case, innovator 1 will innovate if $v_1 + Pr_2 t \geq \varepsilon$, and innovator 2 will innovate if $v_2 \geq \varepsilon + t$. The government wants to maximize expected welfare:

$$\begin{aligned} W &= Pr_1 \left(E(v_1 - \varepsilon / v_1 + Pr_2 t \geq \varepsilon) + Pr_2 E(v_2 - \varepsilon / v_2 \geq \varepsilon + t) \right) \\ &= \frac{1}{2} (1 - \varepsilon) (2 - \varepsilon - t) (1 - \varepsilon + t(1 - t - \varepsilon)). \end{aligned} \quad (2.17)$$

Solving this problem we get that the second best transfer with asymmetric information is:

$$t^* = \frac{3 - 2\varepsilon - \sqrt{6 - 6\varepsilon + \varepsilon^2}}{3}, \quad (2.18)$$

where $t^* < p_1$. Therefore, even if the government does not know v_2 , it will set a lower transfer than the licensing fee of innovator 1 with patents. This is due to the combined effects of asymmetric information and market power with patents.

2.10 Ongoing innovation

In this section we analyze what happens if the sequence of innovations does not stop after one innovation fails to be performed. There is a sequence of innovations $n = 1, 2, \dots$, just as before, but now there can be many trials until an innovation is successful.

Innovator n, j is the j^{th} innovator trying to introduce innovation n ($j - 1$ previous innovators tried to introduce innovation n without success). The innovator has to pay licensing fees to $n - 1$ patentees (the $n - 1$ previous successful innovators), and obtains a draw v_{nj} from the same distribution as before. If the revenues from the innovation are higher than the cost, innovator n, j will introduce the innovation, and in the next stage, innovator $n + 1, 1$ will try to introduce innovation $n + 1$. If revenues are lower than cost, innovator n, j fails to introduce innovation n , which will then be tried by innovator $n, j + 1$ in the

following stage. This innovator will face the same $n - 1$ patent holders and will have a new draw of the value of innovation $v_{n,j+1}$.

For this model, we need to be more specific about the time dimension. Specifically, assume that stages correspond to time periods. At each period only one trial for one innovation is performed. The discount factor of innovator and patent holders is β . At stage $n + j$ the game is summarized by a state $\{n, j\}$.

Let J_{nj}^i be the expected future licensing revenues of patentee i at trial j of innovation n , given that stage $n + j$ has been reached under state $\{n, j\}$. Expressed in a recursive way:

$$J_{nj}^i = Pr_{nj} (p_{nj}^i + \beta J_{n+1,1}^i) + (1 - Pr_{nj}) \beta J_{n,j+1}^i, \quad (2.19)$$

where Pr_{nj} is the probability that innovation n is introduced in trial j . With probability Pr_{nj} , the patent holder gets the price p_{nj}^i plus the continuation value of J of the first trial of the next innovation, $J_{n+1,1}^i$, appropriately discounted by β . With probability $1 - Pr_{nj}$, the innovation will not be introduced and the patent holder gets the continuation value of J corresponding to the next trial of the current innovation.

The profit of the innovator is $I_{nj} (v_{nj} + \beta J_{n+1,1} - c_{nj} - \varepsilon)$.

Just as before, subgame perfection implies that the patent holders take $J_{n+1,1}^i$ and $J_{n,j+1}^i$ as given when deciding p_{nj}^i . The profit maximizing price is $p_{nj}^i = (1 - \varepsilon + \beta J_{n,j+1}^i) / n$. In a symmetric equilibrium, $p_{nj}^i = p_n$ and $J_{nj}^i = J_n$ for all i, j .

Replacing in the probability of innovation, we get $Pr_n = \frac{1-\varepsilon}{n} + \beta J_{n+1} - \frac{n-1}{n} \beta J_n$, and introducing in the expression for J_{nj}^i :

$$J_n = \frac{1}{1-\beta} \left(\frac{1-\varepsilon}{n} + \beta J_{n+1} - \frac{n-1}{n} \beta J_n \right)^2. \quad (2.20)$$

Rearranging this expression:

$$J_{n+1} = \frac{1}{\beta} \left(\sqrt{(1-\beta)J_n} - \frac{1-\varepsilon}{n} \right) + \frac{n-1}{n} J_n, \quad (2.21)$$

which is a decreasing sequence converging to 0 as $n \rightarrow \infty$.

The sequence in terms of probabilities is:

$$Pr_{n+1} = \frac{1-\beta}{\beta} \left(Pr_n + \frac{n-1}{n} \frac{\beta}{1-\beta} Pr_n^2 - \frac{1-\varepsilon}{n} \right)^{1/2}, \quad (2.22)$$

which is also a decreasing sequence converging to 0 as $n \rightarrow \infty$. Therefore, the main conclusions of the basic model still hold under when innovation does not stop when a single innovation fails.

2.11 Finite patents

In this section we analyze what happens if patents have finite length. Each stage corresponds to one period and only one innovation is attempted at each period. If the innovator decides to introduce the innovation, she obtains a patent for L periods. This means that the innovator has to pay patents for L previous innovations, but also charges licenses to L future innovators.

The main difficulty of the present analysis is that now the identity of the patent holders matters. The price and future expected licensing revenues will be different for different patent holders, depending on how long will it take for her patent to expire.

The innovator will introduce the innovation if the revenues from innovation are larger than the cost:

$$v_n + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k \geq \sum_{i=n-L}^{n-1} p_n^i + \varepsilon, \quad (2.23)$$

which means that the probability of innovation is

$$Pr_n = 1 + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k - \sum_{i=n-L}^{n-1} p_n^i - \varepsilon. \quad (2.24)$$

The L current patent holders differ in their objective functions. Let J_n^i be the future expected revenues of patent holder i at stage n , given that stage n has been reached. Then,

$$J_n^i = Pr_n(p_n^i + J_{n+1}^i). \quad (2.25)$$

The patent holder charging a license for the last time is patent holder $n-L$, so $J_{n+1}^{n-L} = 0$. The patent of $n-L+1$, on the other hand, will last for one more period, so $J_{n+1}^{n-L+1} = Pr_{n+1} p_{n+1}^{n-L+1}$. In this way, we can construct the future expected revenues of the L patent holders.

The profit maximization problem is

$$\max_{p_n^i} J_n^i = Pr_n(p_n^i + J_{n+1}^i). \quad (2.26)$$

The first order condition is $-p_n^i - J_{n+1}^i + Pr_n = 0$, so $p_n^i + J_{n+1}^i = Pr_n$ and $J_n^i = Pr_n^2$ for all i . This also implies that $p_n^{n-L} = Pr_n$.

We are interested in stationary equilibria, which means that $Pr_n = Pr$ for all n . This, together with the first order condition, implies that $p_n^i = Pr(1-Pr)$ for $i \geq n-L$. Replacing in the probability of innovation, we get:

$$\begin{aligned} Pr &= 1 + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k - \sum_{i=n-L}^{n-1} p_n^i - \varepsilon \\ &= 1 + \sum_{m=1}^{L-1} Pr(1-Pr)Pr^m + PrPr^L - (L-1)Pr(1-Pr) - Pr - \varepsilon \end{aligned} \quad (2.27)$$

Solving for Pr , we get:

$$Pr = \frac{L+1 - \sqrt{(L-1)^2 + 4L\varepsilon}}{2L} \quad (2.28)$$

which is the stationary equilibrium probability of innovation.

Figure 2.5 shows the probability of innovation as a function of the patent length for $\varepsilon = 0.2$. We can see that the probability of innovation decreases with L , which means that patents hurt more than benefit the innovator. This

is because the innovator has to pay licenses that are certain to the patent holders, but the future licensing revenues are uncertain, as they depend on future innovations being performed.

It is also interesting to see that $Pr \rightarrow 0$ when $L \rightarrow \infty$ and $Pr \rightarrow 1 - \varepsilon$ when $L \rightarrow 0$, which correspond to the previously analyzed patents and no patents cases (with $\phi = 1$).

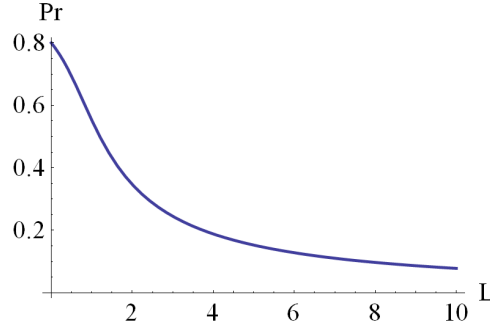


Figure 2.5: Probability of innovation and patent length.

2.11.1 Revenues depend on patent length

We have assumed that the revenues from selling the new product in the final goods market are independent of patent length. In this subsection we analyze what happens when we relax this assumption. Assume the revenues of the innovator are $\phi(L)v_n$, with $\phi'(L) \geq 0$, $\phi''(L) \leq 0$, $\lim_{L \rightarrow 0} \phi(L) = \underline{\phi}$ and $\lim_{L \rightarrow \infty} \phi(L) = 1$. Here, $\underline{\phi}$ is the fraction of social surplus the innovator would appropriate without any patent protection, due to trade secrets or first mover advantages.

In this case, the innovator will innovate if

$$\phi(L)v_n + \sum_{m=n+1}^{n+L} p_m^n \prod_{k=n+1}^m Pr_k \geq \varepsilon + \sum_{i=n-L}^{n-1} p_n^i. \quad (2.29)$$

Applying a similar procedure as that of the previous case, we obtain the probability of innovation in the stationary equilibrium:

$$Pr = \frac{L + 1 - \sqrt{(L-1)^2 + 4L\varepsilon/\phi(L)}}{2L}. \quad (2.30)$$

The effect of patent length on the probability of innovation depends on the functional form of $\phi(L)$. Let $\phi(L) = 1 - \frac{1-\underline{\phi}}{(L+1)^\gamma}$, where γ measures the speed at which revenues grow when L increases. Figure 2.6a shows that when ϕ is more concave ($\gamma = 1$), the probability of innovation first increases and then decreases with patent length. The optimal length is positive and finite (in this case $L = 1$). Figure 2.6b shows that for a lower degree of concavity of $\phi(L)$ it is optimal to completely remove patents. Therefore, the results do not change significantly when the revenues in the final goods sector depend on patent length.

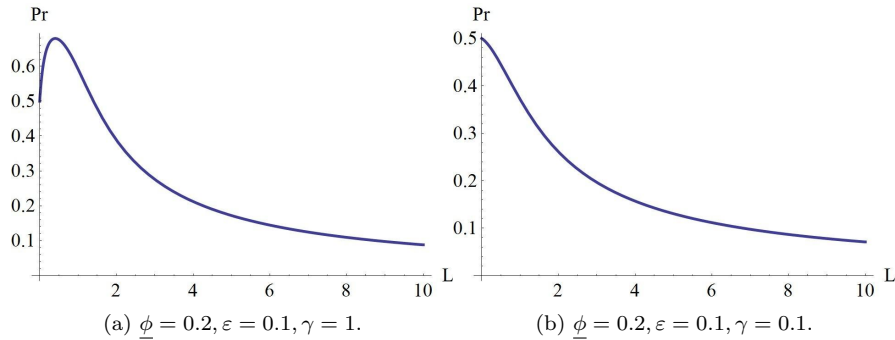


Figure 2.6: Probability of innovation as a function of patent length.

2.12 Conclusion

In this paper we build a dynamic model where accumulation of patents generates an increasing number of claims on cumulative innovation. The model is intended to reproduce the central feature of innovation activity in hi-tech industries: new products are more complex than old products, because they build on a larger stock of previously accumulated knowledge.

We study the policy that maximizes expected social welfare and compare it with the outcome of three patent policy regimes: patents, patent pools and no patents. We find that, even abstracting from the monopolistic inefficiencies of patents, none of these policies attains the optimum.

With patents, the innovator has to pay an increasing number of license fees to previous innovator. Asymmetric information on the value of the innovation and uncoordinated market power of licensor create an anticommons effect that reduces the incentives to innovate as innovation becomes more complex. The anticommons effect is weaker than in the static case, but it is still strong enough to drive the probability of innovation to zero as the number of licenses grows large. Enforcing a patent pool solves the lack of coordination but not the asymmetric information problem. As a result the outcome of patent pools is more desirable but still it does not achieve the first best. Eliminating patent protection solves the two problems but introduce a non-internalized externality: previous innovations set the foundations for future innovations. Therefore it might be the case that the social cost of one innovation is higher than its instantaneous social value (the social value the innovation creates *per se*), and yet the innovation is socially desirable because it allows the development of further innovations. This is the typical problem faced by basic research.

Then we study alternative solutions to the anticommons: (i) the complete sale of patent rights of each innovator to the next one, and (ii) the possibility that the licensee compete with the original licensor. Both alternatives exactly replicate the sequence of innovations under the patent pool regime. Another interesting result of the paper is that patent pools are dynamically unstable, as the incentives to remain outside the pool increase as the sequence of innovations progresses.

We also find that the outcome of these three policy arrangements does not even attain the second best. The second best is achievable by means of govern-

ment transfers, assuming that the government does not know the value of the innovation.

This paper shows that patent protection may be the wrong way to provide incentives to innovation in complex industries like electronics, software and hardware. Enforcing patent pools or eliminating patent protection would improve welfare, but still would not reach the social optimum. We hope this paper will contribute to future research on the design of an optimal innovation policy.

Chapter 3

Intellectual Property Rights in a Model of Economic Growth

Stefano Trento

3.1 Introduction

Technological progress, especially in hi-tech sectors, is widely recognized as one of the main forces behind economic growth. In these sectors innovation is sequential, meaning that new products build upon existing ones.

This is evident when the new good is an improvement over existing products. It is also true when the discovery of the new good requires the use in R&D of research inputs which are previous discoveries themselves. In hi-tech sectors the number of these research inputs often grows large. Patent pools constitutes considerable evidence of this phenomenon. A patent pool is an agreement between two or more patent holders to license one or more of their patents together as a package. The first patent pool was formed in 1856 around intellectual property conflicts in the sewing machine industry. Other important patent pools are those on movie projectors in 1908, on aircrafts in 1917, on radio in 1924, and many more until the most recent one on Blue-ray Disc promoted by MPEG LA in 2007. Some of these patent pools are composed by an incredibly high number of patents: the MPEG2 pool contains more than 600 patents by 25 patent holders. In the absence of a patent pool, the innovator has to purchase a high number of research inputs from different patent holders. This was the case of Golden Rice, whose discovery required the use of around forty patented products and processes (Graff, Cullen, Bradford, Zilberman, and Bennett 2003). Or the development of a malaria vaccine based on the MSP1 protein, that would infringe upon 39 patent families (Commission on Intellectual Property Rights 2002).

These examples show that Intellectual Property Rights (IPRs) affect both the cost and the revenues of the innovator. Therefore the widely believed view that stronger IPRs necessarily increase the incentives to innovate no longer holds. This intuition is at the basis of the literature on sequential innovation initiated by Scotchmer (1991). This literature, with the contributions of Green and Scotchmer (1995), Chang (1995), Scotchmer (1996), and many others, uses two stage innovation models where the first innovation builds the foundations for the second one. It then studies how the optimal patent policy should be designed in order to redistribute profits between the two innovators, so that all the socially valuable innovations are performed. Through static partial equilibrium models, this literature analyzes carefully the effects of IPRs on the profitability of innovation.

The literature on endogenous growth, on the other hand, focuses on the dynamic effects of R&D and innovation on growth. The contribution of this literature has been to make technological progress the result of agents' rational economic decisions. The papers of Romer (1990), Aghion and Howitt (1992), Grossman and Helpman (1991), just to cite a few, have brought fundamental insights on the relationship between R&D and economic growth. This literature also recognizes the cumulative nature of innovation, with each innovation building upon the previous one. In these models the sequentiality of innovation is treated as a big externality from the previous to the next innovator. Innovators do not have to pay for the technologies that have been previously invented and patented, and can freely improve upon them, so that the IPRs policy does not affect the cost of innovation. It does however affect the revenues of the innovator. In this context the incentives to innovate are monotonically increasing in patent length, hence the adoption of an infinite patent life policy.

Therefore there are two separate streams of literature: the one on sequential innovation studies in depth the effects of IPRs on innovation by means of static partial equilibrium models. The one on endogenous growth uses dynamic general equilibrium models to study the effects of innovation on economic growth, but it is not concerned with the response of innovation to different IPRs policy. In order to fully understand how patent policy can stimulate economic growth, there is a need to build a bridge between these two literatures.

Surprisingly, no attempt has been done yet in this sense. There is a growing literature on optimal patent length, but this literature is mostly concerned with a different issue. These models assume that the incentive to innovate is strictly increasing in patent length, and then study the trade-off between the growth enhancing effect and the static inefficiency of the monopoly provided by the patent. Judd (1985) finds that, under certain conditions, infinite patents achieve the efficient rate of growth. This is because when all goods are patented and priced the same, the marginal rate of substitution still equals the marginal rate of transformation and the inefficiency disappears. The same conclusion does not hold true for finite patent length: in this case goods whose patent is expired are priced at marginal cost, while patent protected goods are sold at a mark-up over the marginal cost. ? reverse this result and find that the optimal patent length is finite. This is because they add labor in the production function of the final good. Therefore, even if all intermediate goods are priced the same in the infinite patent case, there is an inefficiency in the choice between labor and intermediate goods. Acemoglu and Akcigit (2007) expand the analysis to include state-dependant IPRs. They find that patent policy should depend on the technological gap between the leading firm and the followers. O'Donoghue and Zweimuller (2004) assume infinite patent duration and focus on policy instruments other than patent length. In particular they examine in depth the patentability requirement and the leading breadth. All these papers, while clarifying many aspects of the role of patent policy in growth, do not take into account its effect on the cost of the innovation.

To the best of my knowledge this is the first paper studying sequential innovation in a dynamic general equilibrium model where IPRs affect both the cost and the revenues of the innovation. In this paper I develop a model of endogenous growth with cumulative innovation, where the innovator has to use a generic number n of research inputs in R&D in order to come up with a new product. The n research inputs are previous innovations, so that the model is a dynamic general equilibrium version of Llanes and Trento (2009). The R&D success rate follows a stochastic poisson process, meaning that the time of innovation is uncertain. In this context patent length plays a central role: by the time R&D takes place, patents on some of the research inputs might have expired. Thus, depending on patent length, a fraction of the research inputs might be sold at a competitive price. The larger this fraction, the lower - *ceteris paribus* - the cost of R&D. But reducing patent length also reduces expected revenues from the innovation. This trade off is the central theme of the paper.

I find that the infinite patent length always provides a less than optimal research effort in the steady-state equilibrium. This result is new with respect to the existing literature on quality based growth. This literature finds that infinite patents might lead to an equilibrium rate of innovation that can be slower, equal or faster than the optimal one. The new result of this paper stems from the fact that the model takes into account the effect of IPRs on the cost

of the innovation.

Then I perform numerical simulations of the model under a regime of finite patent length. These simulations display an hump-shaped level of R&D effort on patent length, suggesting the existence of a finite optimal patent length. This optimal patent length increases the equilibrium level of innovation relative to a system of infinitely lived patents. Also in this case, though, the steady state level of innovation is always below the one a benevolent planner would choose. In other words, a varying patent length is not a good policy instrument to achieve the socially optimal rate of innovation. I use numerical simulations to study the characteristics of the optimal patent length. These simulations suggest that optimal length is decreasing in the size of the economy. This could be the result of the scale effect: the larger the economy the lower the opportunity cost of investing in R&D, and therefore the higher the growth rate of the economy. Or it could be the result of the fact that the larger the economy, the larger the profits of the innovator at each moment in time, and therefore the lower the need for long patents. Or a combination of the two. In the paper I argue that the second one is the only sensible explanation. This is in line with the results of the static analysis of Boldrin and Levine (2005a) and it might be relevant in the context of the recent discussions about introducing an IPRs system in developing countries.

Numerical simulations also suggest that the optimal patent length is increasing in the degree of substitutability between the research inputs used in R&D, reinforcing the results from the static model of ?, and calling for a weaker patent protection in sectors characterized by a higher complementarity in R&D.

The rest of the paper has the following structure: in section 3.2 I set up the model to be used throughout the paper. Section 3.3 studies the efficient R&D effort as the solution to the social planner problem. Section 3.4 analyzes the decentralized equilibrium with infinite patent length, and compares it to the efficient one. In section 3.5 I study the decentralized equilibrium with finite patent life, and I analyze the optimal patent length and its characteristics. Then in section 3.6 I draw some conclusion and policy implications of this model.

3.2 The Model

The structure of the model is a classical endogenous growth model, in line with Grossman and Helpman (1991). In this framework I introduce a technology for R&D that mimics the innovation process of hi-tech sectors, where many R&D inputs are previous innovations themselves. R&D effort depends on the access to these previous discoveries, which in turns depends on the IPRs policy. This structure allows me to study the optimal patent length in a decentralized equilibrium.

3.2.1 Environment.

Time is continuous and denoted by τ . There is one final good supplied in many different qualities. The number of qualities available to the consumers expands with innovation. I denote the quality of the good by the subscript j , where j is an integer number. Each innovation increases the quality of the good by a factor $\gamma > 1$. I denote quality j by q_j , therefore $q_j = \gamma^s q_{j-s}$.

There is a mass L of identical households, each provided with one unit of labor that he supplies inelastically in exchange for a wage w . Households only derive utility from the consumption of the final good.

3.2.2 Preferences.

Households' utility is increasing in both the quality and the quantity of the good consumed. Let d_j be total consumption of quality j good, then lifetime utility of the representative household is:

$$U = \int_0^{\infty} e^{-\beta\tau} \log \left[\sum_j \gamma^j d_j(\tau) \right] d\tau \quad (3.1)$$

where β is the subjective discount rate.

3.2.3 Technology.

Labor is the only input in the production of goods of any quality. The production function is linear, with one unit of good requiring one unit of labor to be produced, independently of its quality:

$$x_j = l_j \quad (3.2)$$

3.2.4 R&D

Quality improvement, as the source of economic growth, is determined endogenously by R&D. I build the R&D production function taking into account two intrinsic features of hi-tech sectors: (i) innovation is sequential, meaning that new goods are built upon existing goods; (ii) a large number of recently discovered techniques and products are used in research. In order to capture these features I assume that innovation requires the use in R&D of n lower quality goods. In particular I assume that the discovery of quality $j + 1$ requires the use in R&D of qualities $j - 1, \dots, j - n$. The R&D effort for the $j + 1$ innovation is measured as follows:

$$y_{j+1} = A \left(\sum_{i=j-n}^{j-1} x_i^\rho \right)^{\frac{1}{\rho}} \quad (3.3)$$

where A is a scale parameter, x_i is the amount of quality i good used as input in R&D, and $\rho \in [0, 1]$ is a parameter related to the substitutability between inputs. In what follow I set $A = n^{(\rho-1)/\rho}$ in order to eliminate increasing returns from specialization. Therefore y only depends on the total amount of inputs used¹.

There is free access to the R&D technology (3.3). Innovation is characterized by uncertainty: the time of the discovery of a new product follows a Poisson process. The larger the research effort, the higher the instantaneous probability of discovering a new good, λy . Because of this, the expected time between the

¹In most models of monopolistic competition $A = 1$. This implies that, for a fixed amount of total input $X = nx$, output is increasing in n , see (Romer 1987)

discovery of quality $j - 1$ and quality j is $\frac{1}{\lambda y_j}$. Notice that because (3.3) is a constant returns to scale technology, and because the sum of N Poisson processes with parameter λ is a poisson process with parameter $N \lambda$, the number of firms engaging in R&D is irrelevant. If there are H firms indexed by $h = 1, 2, \dots, H$ engaging in R&D, then the instantaneous probability of innovation is $\lambda y_{j,h}$ for firm h and $\lambda \sum_{h=1}^H y_{j,h}$ for the economy. Since the number of R&D firms is irrelevant we can assume that in equilibrium there will be only one, in order to save notation.

3.3 Social Planner

The social planner has to allocate resources efficiently between the production of goods for consumption and the production goods to be used in R&D. The total amount of resources (labor) is L .

The planner maximizes (3.1), subject to the production technology (3.2), the R&D technology (3.3), and the resource constraint:

$$L = \sum_{i=0}^j l_i \quad (3.4)$$

Since all goods share the same technology, the planner will choose to produce only the highest quality good (j) for consumption. Also, because R&D is a concave and symmetric technology, he will choose to use the same quantity of each input: $x_{j-1} = \dots = x_{j-n} = x$. Substituting this choice of inputs in (3.3), the research effort becomes: $y_{j+1} = n x$. Also, since qualities $j - n - 1, \dots, 0$ are not used in consumption nor in R&D, it is efficient not to produce them at all, meaning that goods eventually become obsolete.

The resource constraint becomes $L = l_j + y_{j+1}$, where l_j and y_{j+1} are the amount of resources used in the production of the consumption good and in R&D respectively. It follows that total quantity produced of the consumption good is $d_j = L - y_{j+1}$. Therefore the optimization problem for the planner is:

$$\begin{aligned} \max_{y_{j+1}} E(U) &= E \int_0^\infty e^{-\beta\tau} \log [q_j(\tau) (L - y_{j+1}(\tau))] d\tau & (3.5) \\ \text{s.t.} & & (3.2), (3.3), (3.4) \end{aligned}$$

After some calculations, the expected lifetime utility becomes (See Appendix):

$$E(U) = \frac{1}{\beta} \left(\frac{\lambda y \log(\gamma)}{\beta} + \log(L - y) \right) \quad (3.6)$$

The planner faces a trade-off between more consumption today and a more valuable consumption tomorrow. Maximizing (3.6) with respect to the constant research effort y , gives the following first order condition:

$$\frac{\lambda \log(\gamma)}{\beta} = \frac{1}{L - y} \quad (3.7)$$

This condition states that the marginal benefit of employing resources in R&D (the *lhs*) must be equal to its marginal cost (the *rhs*). The marginal benefit

is increasing in the Poisson parameter λ , which sets the pace of innovation for any given research effort, and in γ , the quality step. Not surprisingly it is decreasing in the subjective discount rate. The marginal cost is the opportunity cost of employing resources in R&D rather than in the production of the good for consumption: the foregone marginal utility of consumption.

Solving (3.7) for y , the following proposition holds:

Proposition 11 (Planner Allocation) *The optimal amount of research effort in the steady-state equilibrium is $y^* = L - \frac{\beta}{\lambda \log(\gamma)}$. This is increasing in total labor resources (L), the size of quality step (γ) and the effectiveness of R&D (λ). It is decreasing in the subjective discount rate (β).*

These relationships are all very intuitive. The research effort y is an increasing function of λ , the effectiveness of R&D, and γ , the increase in marginal utility after an innovation arrives. It is decreasing in β because β decreases the present value of future utility. And it is increasing in the total amount of resources available in the economy L . This is because increasing total resources reduces the opportunity cost of investing in R&D. This relationship creates what is known as "scale effect", which is an unwelcome feature of many growth models (Romer 1990, Aghion and Howitt 1992, Grossman and Helpman 1991). The scale effect, in its strongest form, means that there is a positive correlation between the size of the economy and its growth rate. Recently there has been a successful attempt to correct for the scale effect (see Jones (1999) for a survey). This paper, like other endogenous growth papers interested in the effect of patents (Judd 1985, O'Donoghue and Zweimuller 2004, ?), abstracts from this scale effect by just assuming that population is constant.

3.4 Decentralized Steady-State Equilibrium With Infinitely Lived Patents

This section analyzes the decentralized steady-state equilibrium level of the R&D effort. This level is determined by the interaction of households, R&D firms and producers.

3.4.1 Representative Household.

There is a continuum of households of measure L . Household ℓ maximizes total lifetime utility subject to an intertemporal budget constraint. Utility only depends on consumption, while labor is inelastically supplied for any wage $w > 0$. The household values both quantity and quality of consumption. Households own the firms, but we will see that firms have zero expected profits: the Law of Large Numbers then implies that the value of households' share is zero, this is why it does not appear in the budget constraint. The maximization problem is the following:

$$\begin{aligned} \max_{d_j^\ell} \quad & U^\ell = \int_0^\infty e^{-\beta\tau} \log \left[\sum_j q_j^\ell(\tau) d_j^\ell(\tau) \right] d\tau \\ \text{s.t} \quad & \int_0^\infty e^{-r\tau} E^\ell(\tau) d\tau \leq A^\ell(0) + \int_0^\infty e^{-r\tau} w d\tau \end{aligned} \quad (3.8)$$

where $E^\ell = \sum_j p_j d_j^\ell$ represents total expenditures of household ℓ , $A^\ell(0)$ is household's initial wealth, and r is the instantaneous interest rate. The representative household solves the maximization problem in two steps: first she solves a static problem by deciding the composition of her expenditure at each point in time. Second she decides the optimal expenditure plan. Given preferences, at each point in time the household only consumes the good with the lowest p_j/q_j ratio, where p_j is the price of quality j . I assume that, when faced with goods of the same price/quality ratio, the household prefers the good with highest quality. Therefore at each point in time household ℓ demands $d_j^\ell = E^\ell/p_j$ of the good with lowest price/quality ratio, and zero of the rest of the goods.

Initial wealth $A^\ell(0)$ is identical across all household. Household's wealth increases according to $\dot{A}^\ell(\tau) = rA^\ell(\tau) + w - E^\ell(\tau)$. I focus on the steady state with $\dot{A}^\ell(\tau) = 0$, which leads to a constant expenditure plan:

$$E^\ell(\tau) = rA^\ell(0) + w \quad (3.9)$$

Therefore the interest rate equals the discount rate, and consumption expenditures are constant in time. Aggregate demand is just $E(\tau) = \int_0^L E^\ell(\tau) d\ell$. Because all households are identical $E(\tau) = L E^\ell(\tau)$.

3.4.2 R&D firms.

At each point in time R&D firms use technology (3.3) in a constant effort to innovate. If firm h exerts effort $y_{j+1,h}$, its instantaneous probability of innovation is $\lambda y_{j+1,h}$, and its expected arrival rate of innovation $j+1$ is $1/(\lambda y_{j+1,h})$. As noted above, there is free access to the R&D technology, and we assume that there are no externalities in research, meaning that individual R&D firms' probability to innovate are independent. It follows that the number of firms engaging in R&D is indeterminate, so we can get rid of subscript h .

When one R&D firm discovers quality $j+1$, it gets a patent, becomes a producer and starts selling quality $j+1$ to consumers as the highest quality good in the market. And after quality $j+2$ is discovered, quality $j+1$ is also sold as a research input for the discovery of the next n qualities. I will be more precise about this timing later on. Also, as soon as quality $j+1$ is invented by one firm, all the other R&D firms close their R&D race for the invention of quality $j+1$ with losses, and a new R&D process begins for the discovery of quality $j+2$.

Since the number of firms in the R&D sector is indeterminate, I will use a representative firm. The R&D firm chooses the research effort y to maximize expected profits. When engaging in the discovery of quality $j+1$, it solves:

$$\max_{y_{j+1}} \Pi = \lambda y_{j+1} V_{j+1} - c_{j+1}(y_{j+1}) \quad (3.10)$$

where c_{j+1} is the cost of performing y_{j+1} level of R&D, V_{j+1} denotes total profits from selling quality $j+1$ in case the R&D process is successful, and again λy_{j+1} is the instantaneous probability of innovation. In section 3.4.4 I will make V_{j+1} explicit, as a function of production technology, patent policy, and R&D technology. For now it is only important to know that V_{j+1} is the sum of (i) the profits from being the quality leader in the final good sector, plus

(ii) the profits from being the supplier of one of the n inputs used in the R&D process.

The R&D firm maximizes (3.10) in two steps: first, given prices of quality i , for $i = j-1, \dots, j-n$, it derives conditional factor demands for the corresponding inputs x_i in order to minimize c_{j+1} :

$$c_{j+1} = \min_{x_i} \sum_{i=j-n}^{j-1} p_i x_i \quad (3.11)$$

$$s.t. \quad n^{\frac{\rho-1}{\rho}} \left(\sum_{i=j-n}^{j-1} x_i^\rho \right)^{\frac{1}{\rho}} \geq y_{j+1} \quad (3.12)$$

Solving (3.11) we get the conditional factor demands, and the cost function of R&D effort:

$$x_i = n^{-\frac{1}{1-\sigma}} p_i^{-\sigma} \left(\sum_{i=j-n}^{j-1} p_i^{1-\sigma} \right)^{\frac{\sigma}{1-\sigma}} y_{j+1} \quad (3.13)$$

$$c_{j+1} = n^{-\frac{1}{1-\sigma}} \left(\sum_{i=j-n}^{j-1} p_i^{1-\sigma} \right)^{\frac{1}{1-\sigma}} y_{j+1} \quad (3.14)$$

where $\sigma = (1 - \rho)^{-1}$ is the elasticity of substitution between inputs.

After deriving conditional factor demands, the R&D firm maximizes (3.10). The first order condition is:

$$\frac{\partial c_{j+1}}{\partial y_{j+1}} = \lambda V_{j+1} \quad (3.15)$$

Because of constant returns to scale, Poisson arrival rate of innovation which makes expected revenues linear in y , and free access to the R&D technology, expected profits from R&D will be zero. We will state the zero-profit condition in section 3.4.4.

3.4.3 Producers.

At each point in time there are: an indeterminate number of R&D firms and $n + 1$ producers, each one of them producing one of the last $n + 1$ qualities in the market. Nothing prevents qualities to be sold both to consumers and to R&D firms, but we will see that in equilibrium only the last quality is sold to consumers, and the other n qualities are used in R&D. For simplicity we assume that the producer of any quality is the R&D firm that invented it. Also, in order to avoid strategic pricing, we assume that the goods are sold anonymously to the R&D firms. In the next subsection we will see that the optimal pricing depends on whether the quality is used in R&D or it is sold to consumers. I will refer to qualities used in R&D as inputs.

One last remark: for the sake of simplicity I maintain throughout the assumption that the highest quality in the market does not enter the R&D technology. This assumption is crucial because it means that in equilibrium no quality is sold both to consumers and to R&D firms. While certainly not very realistic, this assumption allows me to solve the model while keeping the two main characteristics of innovation in hi-tech sectors: (i) the quality of new goods is higher than that of older goods; (ii) older innovations are inputs in the R&D process leading to the discovery of new products.

3.4.4 Balanced Growth Path.

Input producers set the usual monopolistic competition price $p_i = \frac{\sigma w}{\sigma-1}$, and since the R&D production function is symmetric they all set the same price $p_i = p$. Since all inputs are priced the same, and because of symmetry and concavity of (3.3), the innovator will demand the same quantity of each input. In fact, substituting p into the conditional demand (3.13) we get $x_i = y_{j+1}/n$ for $i = j - n, \dots, j - 1$. Also, substituting p into the cost function and plugging it into (3.15), we get the zero profit conditions:

$$V = \frac{\sigma w}{(\sigma - 1)\lambda} \quad (3.16)$$

which we will use later to find the equilibrium value of the R&D effort in the balanced growth path. Since V depends only on constant parameters and on the numeraire w , I removed the subscript $j + 1$.

Now that we have the no profit condition we can be more specific about the value of quality $j + 1$. We said that this value depends on the profits the innovator will make: (i) as a final good producer until quality $j + 2$ is invented, and (ii) as an input producers by selling its good to R&D firms for the discovery of better qualities. Let us start from the last one. In equilibrium, given the mark-up pricing and (3.13), instantaneous profits of the inputs producers are:

$$\Pi_{ip} = (p_i - w) x_i = \frac{\sigma w y_{j+1}}{n(\sigma - 1)} \quad (3.17)$$

Let us now turn to the optimal pricing of the qualities sold to consumers. The successful innovator is the quality leader in the market for the consumption good. His quality q_{j+1} is γ times higher than the previous innovation's quality q_j . Since consumers only consume the quality with the lowest p_i/q_i ratio, the innovator maximize profits by setting a price that is γ times higher than the price of his competitors. The constant marginal cost of one unit of any good is w . Therefore the optimal price for the last innovator is² γw . At each point in time she makes profits:

$$\Pi_L = (\gamma w - w) \frac{E}{\gamma w} = \left(1 - \frac{1}{\gamma}\right) E \quad (3.18)$$

After the following innovation takes place, quality $j + 1$ is replaced by quality $j + 2$ in the final good market. At this point it is sold as an input in the R&D market for the next n innovations.

Therefore total expected revenues of engaging in R&D for the discovery of quality $j + 1$ are:

$$V_{j+1} = \int_0^\infty e^{-r\tau} \left[e^{-(\lambda y_{j+2})\tau} \Pi_L + \sum_{i=2}^{n+1} e^{-(\lambda y_{j+i})\tau} \frac{((\lambda y_{j+i})\tau)^{i-1}}{(i-1)!} \Pi_{ip} \right] d\tau \quad (3.19)$$

This expression represents the expected value of the innovation $j + 1$. It is equal to the discounted expected profits from selling quality $j + 1$ as a quality

²The innovator does not set a lower price because he is facing a unitary elastic demand (remember $d_j = E/p_j$ for the quality with the lowest price/quality ratio). With a unitary elastic demand the innovator would want to set a price equal to infinity, but he is constrained by competition to set $p = \gamma w$, otherwise the demand for his good will be zero

leader in the consumption good market first, and then as one of the n inputs producers for the next n innovations. The first term inside the square bracket is the probability that nobody has discovered quality $j + 2$ at time τ , times the profits as quality leader in the market for consumption. The second term is the probability that, at time τ , quality $j + 1$ good is still in the R&D process, times the profits as input producer. For quality $j + 1$ to be part of the R&D process we need that: (i) quality $j + 2$ has already been invented; (ii) no more than n innovations have been made after the invention of quality $j + 1$.

Substituting (3.17), (3.18) and the resource constraint³ into (3.19) and solving for V in the balanced growth path, we obtain:

$$V^* = \frac{(\gamma - 1)w(L - y)}{r + \lambda y} + \frac{\lambda y^2 w \left(1 - \left(\frac{\lambda y}{r + \lambda y}\right)^n\right)}{n r (r + \lambda y)(\sigma - 1)} \quad (3.20)$$

The following result holds:

Proposition 12 (Decentralized Equilibrium) *A decentralized steady state equilibrium with infinite patents exists and it is unique. The research effort (y_∞) and total Welfare (W_∞) are less than optimal. y_∞ is decreasing in n and r , and it is increasing in γ , λ and σ .*

in the Appendix

The result that, in the presence of infinite patents, the decentralized equilibrium delivers an insufficient level of R&D is new. The previous literature on quality-ladder based growth, by ignoring the cost-effect of patents studied here, concluded that, depending on different constellations of parameters, R&D effort could be higher, equal or lower than socially desirable. This paper shows that incorporating a more realistic R&D production function, where IPRs affect both cost and revenues of the innovator, leads to a different conclusion. Patents of infinite length always fail short of delivering the correct incentives to innovate as the monopoly power they attribute to past innovators more than compensate for the extra profits they bestow upon the current one. The next section analyzes whether a finite patent length may achieve the socially desirable innovation effort and, in that case, the properties of the optimal patent length.

3.5 Decentralized Steady-State Equilibrium With Finite Patent Length

This section analyzes the decentralized equilibrium when patents have a finite life. Let the patent policy parameter $\phi > 0$ represent patent length. I assume that the innovator does not have any advantage over imitators so that, after the patent expires, a costless imitation process immediately drives the price down to the marginal cost w .

A finite patent length reduces both expected cost and expected revenue of the R&D firms. We have seen in the previous section that when patents are

³Notice that $\frac{E}{\gamma w}$ in (3.18) is the quantity sold in the market for consumption good. If the highest quality is j , then $d_j = \frac{E}{\gamma w} = L - y_{j+1}$

infinitely lived the research effort is lower than optimal. Therefore if a finite patent length has to bring an efficiency gain, it must increase the incentives to innovate by reducing the cost of R&D more than it reduces the expected revenue.

With finite patent life the representative household analysis stays the same: she will only consume the good with the highest quality/price ratio. Because of the effects on cost and revenues from R&D, the innovator's problem changes. Equation (3.10) becomes:

$$\max_{y_{\phi,j+1}} \Pi_{\phi} = \lambda y_{\phi,j+1} V_{\phi,j+1} - c_{\phi,j+1}(y_{\phi,j+1}) \quad (3.21)$$

where all variables are defined as in (3.10), and the subscript ϕ indicates that we are now operating with finite patent life.

The R&D firms solve the same cost minimization problem, therefore (3.11) and (3.13) still hold. Inputs still protected by a patent are sold at the monopolistic competition price $p_i = \frac{\sigma w}{\sigma-1}$, but goods whose patent has expired are sold at marginal cost w . The policy maker takes into account the effect of patent length on expected cost of R&D. For a given patent length ϕ , the expected cost becomes (see Appendix):

$$E(c_{\phi}) = n^{-\frac{1}{1-\sigma}} \left(np^{1-\sigma} - (p^{1-\sigma} - w^{1-\sigma}) e^{-\lambda y_{\phi}} \left(n + \sum_{i=1}^n i \frac{(\lambda y_{\phi})^{n-i+1}}{(n-i+1)!} \right) \right)^{\frac{1}{1-\sigma}} y_{\phi} \quad (3.22)$$

The first order condition of (3.21) is:

$$\frac{\partial c_{\phi,j+1}}{\partial y_{\phi,j+1}} = \lambda V_{\phi,j+1} \quad (3.23)$$

and again, constant returns to scale, Poisson arrival rate of innovation and free access to the R&D technology, make expected profits from R&D zero. In section 3.5.1 I will make the zero-profit condition explicit.

3.5.1 Balanced Growth Path.

In order to find the steady-state equilibrium I follow the same steps as in the previous section: I first find the equilibrium prices and demands, and then use them to find the zero profits conditions. Finally I use the zero profit condition together with the the resource constraint to find the equilibrium research effort in the balanced growth path.

Because of the symmetry of the R&D production function, all inputs covered by a patent have the same price, which is the usual mark up over the marginal cost $p = \frac{\sigma w}{\sigma-1}$. Inputs whose patent is expired are sold at marginal cost w . Notice that, because inputs are priced differently depending on being patent protected or not, the demands will no longer be symmetric. Also, substituting the optimal p in the cost function and solving the first order condition of the R&D firm (3.23) in the balanced growth path, we get the zero profit condition:

$$V_{\phi} = \frac{E(c_{\phi})}{\lambda y_{\phi}} \quad (3.24)$$

where $E(c_{\phi})$ is as in (3.22).

3.5. DECENTRALIZED STEADY-STATE EQUILIBRIUM WITH FINITE PATENT LENGTH69

Let us now be more precise about the value of quality $j + 1$, once it has been invented. We said that the value depends on the profits made by selling quality $j + 1$ as a final good until quality $j + 2$ is discovered, and as a research input to R&D firms. Let us start from the last one. Instantaneous profits for the input producer in the balanced growth path are:

$$\Pi_{\phi,ip} = (p_i - w)x_i^\phi \quad (3.25)$$

where expected demand in the Balanced Growth Path for input i is (see Appendix):

$$x_i^\phi = n^{-\frac{1}{1-\sigma}} p_i^{-\sigma} \left(np^{1-\sigma} - (p^{1-\sigma} - w^{1-\sigma}) e^{-\lambda y \phi} \left(n + \sum_{i=1}^n i \frac{(\lambda y \phi)^{n-i+1}}{(n-i+1)!} \right) \right)^{\frac{\sigma}{1-\sigma}} y_\phi \quad (3.26)$$

and $p_i = \frac{\sigma w}{\sigma-1}$ if input i is still protected by a patent, while $p_i = w$ if the patent on quality i has expired. As in the infinite patents case, the last innovator finds it profitable to set the price γw . Instantaneous profits for the innovator as the quality leader in the consumption good market is:

$$\Pi_{\phi,L} = (\gamma w - w) \frac{E}{\gamma w} = \left(1 - \frac{1}{\gamma} \right) E \quad (3.27)$$

Total expected profits for the successful innovator are:

$$V_{\phi,j+1} = \int_0^\phi e^{-r\tau} \left[e^{-(\lambda y_{\phi,j+2})\tau} \Pi_{\phi,L} + \sum_{i=2}^{n+1} e^{-(\lambda y_{\phi,j+i})\tau} \frac{(\lambda y_{\phi,j+i}\tau)^{i-1}}{(i-1)!} \Pi_{\phi,ip} \right] d\tau \quad (3.28)$$

The following proposition holds:

Proposition 13 *In the decentralized economy with finite patent life, a steady-state equilibrium with positive innovative effort exists and is unique*

in the Appendix

In what follows I present the results of numerical simulations that analyze the characteristics of the optimal patent policy. Table (3.1) shows the benchmark values of the parameters (the values used in the graphs) and the range of values upon which robustness checks have been ran.

	Benchmark	Range
σ	3	[1.1, 50]
β	0.04	[0.03, 0.1]
L	1	[0.9, 1.1]
n	10	[2, 50]
λ	0.4	[0.1, 1.37]
γ	1.25	[1.05, 2]

Table 3.1: Parameters values

The parameters have been chosen as follows: for the annual discount rate, equal to the interest rate, I take the standard value $r = 0.04$, but then I run robustness checks for $r \in [0.03, 0, 1]$.

For γ and λ I used a reasoning similar to Stokey (1995): the average annual rate of economic growth due to technological progress is between 0.5 and 1 per cent. In the model the expected rate of economic growth is the probability of innovation times the quality improvement $\lambda y(\gamma - 1)$. Since y is endogenous to the model it is complicated to calibrate the model to exactly match the data on economic growth. Also I have one degree of freedom in the choice of γ and λ that match the data. Therefore I let both parameters vary within a reasonable range. The benchmark values of γ and λ imply a growth rate due to innovation ($\lambda y(\gamma - 1)$) of about one percent. Higher values of γ and lower values of λ fit better the data in sectors where the innovation activity is rare yet radical. On the contrary, lower values of γ and higher values of λ better illustrate the case of constant but little innovation. The wage rate does not affect the equilibrium, since it is just the numeraire in the model. In the benchmark case I take the elasticity of substitution between inputs $\sigma = 3$, but then I run robustness check for $\sigma \in [1.1, 50]$. The number of inputs entering the R&D process in the baseline case is 10, but I allow it to decrease and increase between 2 and 50. I take total labor $L = 1$ and then I let it vary to study how the size of the economy affects the optimal patent policy.

Although the results of the simulation obviously change quantitatively for different values of the parameters, the qualitative results are very robust. The simulation suggests that the steady-state level of R&D effort (y_ϕ) and total Welfare (W_ϕ) with finite patents are hump shaped in ϕ . No finite patent length reaches the social optimum. Figure (3.1) depicts the research effort under infinite patent policy y_∞ and under finite patent policy y_ϕ , as a function of patent length. The efficient level of R&D y^* , is much higher, at 0.55 and of course it does not depend on patent length.

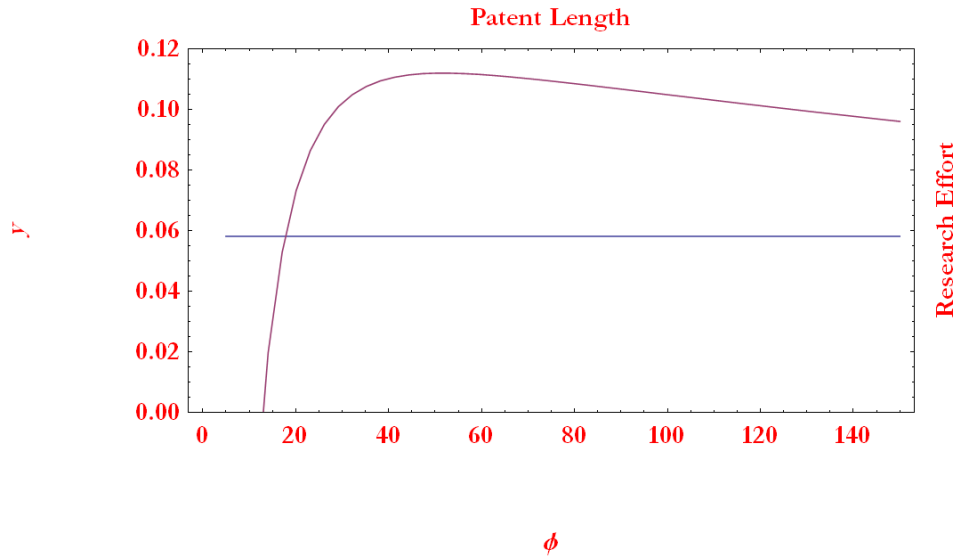


Figure 3.1: Patent length and research effort

Figure (3.2) shows the behavior of the total Welfare of the economy as a function of patent length. It is evident that total Welfare behaves similarly

to research effort, so that the patent policy that maximizes the research effort also maximizes total Welfare of the economy. Patents affects Welfare in two ways: through its effect on the rate of growth of the economy and through the inefficient allocation stemming from the monopoly it provides. In this model the first effect is the crucial one. There exist two inefficiencies of the monopoly: the monopoly on the good for consumption (the latest innovation), and the monopoly on the research inputs (the older innovations). The optimal patent length is always longer than the arrival rate of one innovation, therefore the first monopoly is almost irrelevant. In words, reducing or eliminating this inefficiency would require a patent length that is shorter than the arrival rate of one innovation, but this patent length is never optimal, as it is too short to provide incentives to innovate in this model of costless and timeless imitation. The second monopolistic inefficiency, the one on research inputs, turns out to be marginal. This is because the inefficiency here is the misallocation of resources between patented and non-patented inputs, which stems from the competitive versus non-competitive pricing of equally productive inputs. Any patent length longer than the arrival rate of one innovation implies such a misallocation. It turns out that the difference in the magnitude of this inefficiency is always dominated by the effect on economic growth. This is why, in this model, a patent life that maximizes economic growth also maximizes Welfare.

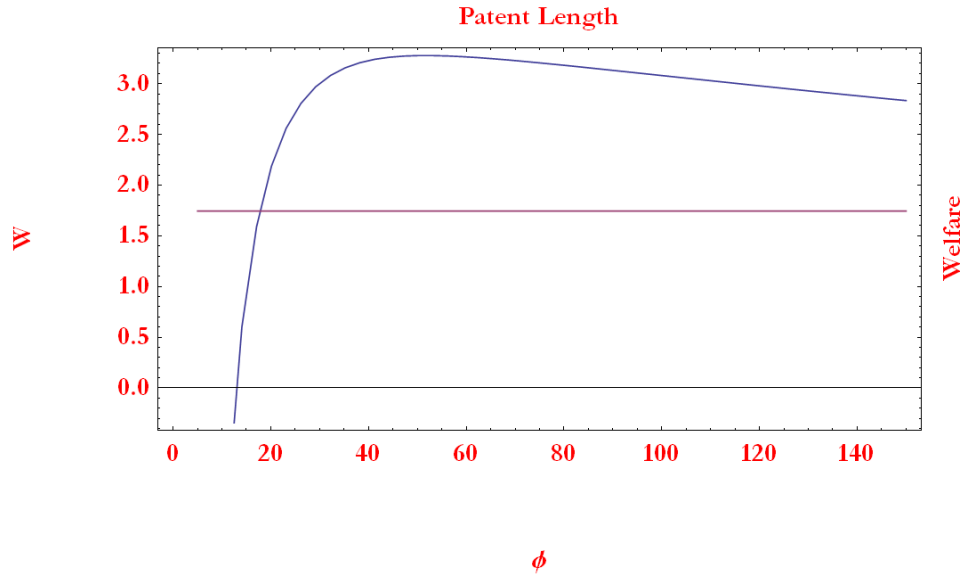


Figure 3.2: Patent length and Welfare

Tables (3.2),(3.3),(3.4),(3.5) report the values of the research effort, the level of welfare and the growth rate of the economy under the optimal finite patent policy (y_ϕ, W_ϕ, g_ϕ) , and under infinite patents $(y_\infty, W_\infty, g_\infty)$ for varying values of the parameters σ, n, β and γ .

The tables suggest that under a finite patent duration regime the economy can perform better than in the infinite patent length regime, both in terms of economic growth and welfare. Nevertheless, in this numerical simulation, no

	σ				
	1.1	3	5	25	50
y_{ϕ^*}	0.091	0.112	0.115	0.118	0.119
y_{∞}	0.000	0.058	0.084	0.113	0.117
W_{ϕ^*}	2.697	3.278	3.358	3.452	3.463
W_{∞}	0.000	1.745	2.507	3.315	3.408
g_{ϕ^*}	0.912%	1.120%	1.149%	1.183%	1.187%
g_{∞}	0.000%	0.581%	0.845%	1.113%	1.117%

Table 3.2: Robustness Check for different values of σ Baseline parameters values: $\lambda = 0.4$, $\beta = 0.04$, $n = 10$, $\gamma = 1.25$, $L = 1$

	n				
	2	5	10	20	50
y_{ϕ^*}	0.098	0.107	0.112	0.115	0.116
y_{∞}	0.062	0.059	0.058	0.058	0.057
W_{ϕ^*}	2.897	3.145	3.278	3.366	3.387
W_{∞}	1.853	1.775	1.745	1.731	1.728
g_{ϕ^*}	0.983%	1.072%	1.120%	1.152%	1.159%
g_{∞}	0.618%	0.591%	0.581%	0.576%	0.573%

Table 3.3: Robustness Check for different values of n Baseline parameters values: $\sigma = 3$, $\lambda = 0.4$, $\beta = 0.04$, $\gamma = 1.25$, $L = 1$

finite patent length reaches the optimal level of innovation and welfare, suggesting that varying patent length is not an optimal policy instrument. A simple way to reach the efficient level of R&D in this model would be to eliminate patent protection and give a fixed subsidy equal to wy^* (the cost of the socially optimal R&D effort) to the innovating firms.

A second interesting result of the simulation is that the optimal patent length is decreasing in the size of the economy, L . There are potentially two forces pushing in this direction. One is that this model incorporates a *scale effect*: the larger the size of the economy, the faster *ceteris paribus* the growth rate of the economy. The mechanism works through the marginal cost of employing resources in R&D, which is decreasing in L . Therefore, as it is clear from Proposition (11), increasing L increases the R&D effort more than proportionally. The second force depends on the fact that L also represents the size of market for the new innovation. The larger the market, the larger the profit of the innovator at each point in time. Meaning that the innovator will be able to recoup the cost of R&D sooner, and patent length can be reduced, with the beneficial consequence of reducing the R&D cost of following innovations.

I do not believe that the first one is the leading effect on the optimal patent length. In fact I do not think it has an effect at all. If that was the case, the argument would go this way: since, by increasing L I increase R&D effort more than proportionally, I want to offset this excessive reaction of y by reducing patent length. But this argument must be wrong because, in the numerical simulations, R&D effort under finite patent length is always lower than the socially optimal one ($y_{\phi} < y^*$). Then there is no reason why we would want to

3.5. DECENTRALIZED STEADY-STATE EQUILIBRIUM WITH FINITE PATENT LENGTH73

	β				
	0.03	0.04	0.05	0.06	0.10
y_{ϕ^*}	0.131	0.112	0.093	0.075	0.000
y_{∞}	0.081	0.058	0.036	0.014	0.000
W_{ϕ^*}	8.306	3.278	1.372	0.559	0.000
W_{∞}	5.222	1.745	0.552	0.115	0.000
g_{ϕ^*}	1.309%	1.120%	0.933%	0.748%	0.000%
g_{∞}	0.811%	0.581%	0.360%	0.143%	0.000%

Table 3.4: Robustness Check for different values of β
 Baseline parameters values: $\sigma = 3$, $\lambda = 0.4$, $n = 10$, $\gamma = 1.25$, $L = 1$

	γ				
	1.2	1.3	1.4	1.5	1.6
y_{ϕ^*}	0.078	0.144	0.202	0.253	0.297
y_{∞}	0.030	0.085	0.136	0.183	0.226
W_{ϕ^*}	1.517	5.561	11.348	18.329	26.116
W_{∞}	0.599	3.368	7.811	13.518	20.163
g_{ϕ^*}	0.620%	1.729%	3.231%	5.053%	7.136%
g_{∞}	0.237%	1.024%	2.183%	3.666%	5.428%

Table 3.5: Robustness Check for different values of γ
 Baseline parameters values: $\sigma = 3$, $\lambda = 0.4$, $\beta = 0.04$, $n = 10$, $L = 1$

reduce patent length to reduce the R&D effort y_{ϕ} .

I therefore conjecture that the second force is the most important if not only force behind the effect of L on optimal patent length.

The result of the numerical simulation on the effect of L on optimal patent protection is in line with the results of Boldrin and Levine (2005a). Using a static framework, they show that an increase in the market size for innovation should be accompanied by a decrease of patent strength. Figure (3.3) shows the negative relationship between L and ϕ^* .

On the other hand numerical simulation suggests that optimal patent length is increasing in the degree of substitutability between research inputs. Figure (3.4) shows this positive relationship between optimal patent length and σ .

This numerical finding is quite intuitive because inputs producers compete in price, and the degree of substitutability sets a bound on their market power. When inputs are perfect substitutes for instance, no matter if they are patent protected they will set a price equal to the marginal cost. In this extreme case patents do not affect the cost of the innovation. Nevertheless they still affect expected revenues from innovating. Consequently an infinite patent length would maximize economic growth.

The other side of the medal is that the higher the complementarity between the research inputs, the weaker the optimal patent protection. This would point in the direction of allowing for different patent policies in differently structured sectors: hi-tech sectors, often characterized by a high complementarity in R&D, are likely to suffer from an excessive patent protection. This is in line with the concern about the surge of anticommons in these sectors, namely that strong

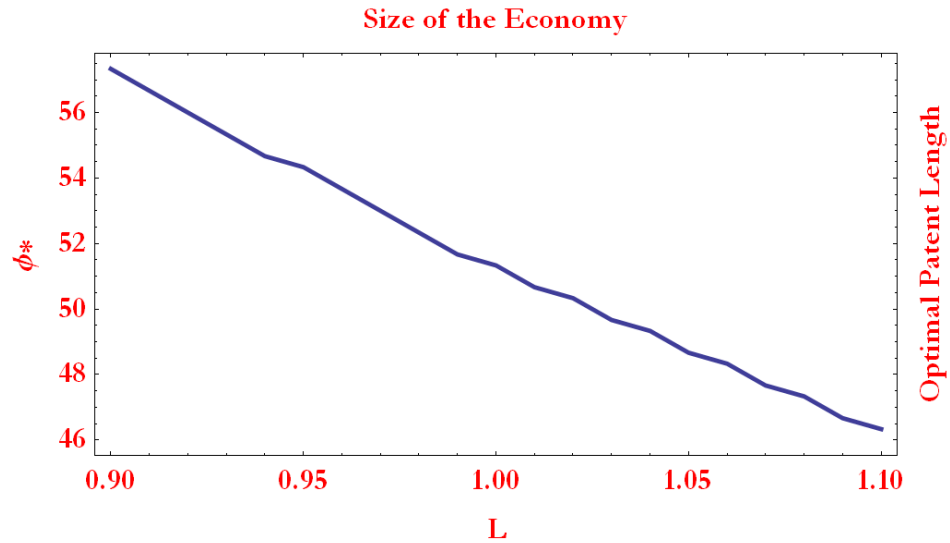


Figure 3.3: Size of the economy and optimal patent length

and fragmented IPRs make it so difficult for the innovator to gather all the needed licenses that they reduce rather than spur innovation.

3.6 Conclusions and Policy Implications

In this paper I investigate the effects of intellectual property rights on economic growth when the latter is due to sequential innovations feeding on earlier ones. Previous models of endogenous innovation abstracts from the fact that earlier innovations are inputs in the R&D process for new ones, thereby assuming an increasing relationship between the length of the patent and the incentives to innovate. The literature on static sequential innovation has made clear that, when innovation is sequential, patents affect both the revenue from and the cost of innovations. This is especially true in high tech sectors, where many of the products and processes used in R&D are previous innovation themselves, and therefore likely to be patent protected. In this context, stronger patent protection does not necessarily increase the incentives to innovate and it may well reduce it.

This is the first time in which this dual effect of IPRs on innovation is considered in a dynamic general equilibrium setting, and the results are novel. In particular, I find that, in the steady-state equilibrium, an IPRs system with infinitely lived patents does not maximize either social welfare or the rate of economic growth: it provides a level of R&D which is too low.

I also analyze the steady-state balanced growth path with finite and optimal patent life. Numerical simulations suggest that R&D effort is initially increasing and then decreasing in patent length. This inverse-U shape is easily explained by two effects. Because of the, extreme, assumption of costless and timeless imitation, zero patent protection results in no innovation. As patent protection

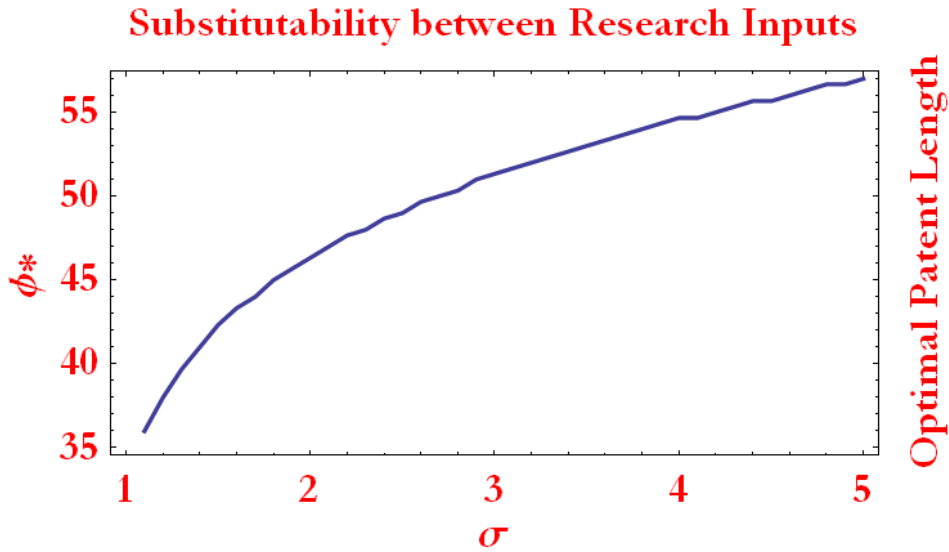


Figure 3.4: Degree of substitutability and optimal patent length

increases, the level of R&D steadily increases, until it reaches its maximum at a level which is higher than its corresponding level with infinite patents. After that it slowly decreases until it reaches its infinite patent length equilibrium value. It would be interesting to study what happens to the optimal patent length if imitation is neither costless nor timeless

I also run numerical simulation of the model which suggest that, under the maintained assumptions, there exists a unique optimal finite patent length. The numerical exercise exhibits a negative relationship between this optimal length and the size of the economy, which coincides with the market for the innovation. In the paper I argue that this relationship does not depend on the scale effect, but rather on the fact that large markets imply larger monopolistic rents at each moment in time, reducing the necessity for long patent life. Reducing patent length then has the beneficial effect of reducing the R&D cost of further innovation, by reducing the amount of license fees the innovator has to pay. I believe that this point should be further investigated, especially in relation to the recent debate about enforcing IPRs in developing countries. If this happens then the size of the market for the innovator obviously increases, with some repercussion on the optimal patent length

Another interesting numerical result relates the optimal patent life to the degree of substitutability of the research inputs. There seems to be a positive relationship between the two: higher substitutability calls for a longer patent protection, meaning that sectors characterized by a higher complementarity in R&D should receive weaker patent protection.

The model can be extended in a least two ways: by introducing specialized labor into the R&D production function and by relaxing the assumption of costless and timeless imitation. In the first extension one could assume that new goods can be obtained either by means of patented inputs or by using special-

ized labor to discover around existing patents, or a combination of both. This extension would introduce an additional inefficiency: because of the monopoly distortion created by the presence of patents, the demand for specialized labor is higher than it would be at the efficient allocation, with some implication on the wage of specialized labor. The second extension would be to provide the innovator with a first mover advantage, yielding competitive rents that in turn affect the optimal patent policy. I leave these extensions for future research.

3.7 Appendix

3.7.1 Derivation of expected utility

The Bernoulli instantaneous utility function is separable in quality $q_j(\tau)$ and quantity $(L - y_{j+1}(\tau))$. It is easier to work with this two term separately. Expected utility is:

$$E(U) = E \int_0^\infty e^{-\beta\tau} \log(q_j(\tau)) d\tau + E \int_0^\infty e^{-\beta\tau} \log(L - y_{j+1}(\tau)) d\tau$$

Let us label $E(U_q)$ the first term on the *r.h.s.*, and $E(U_d)$ the second term on the *r.h.s.* $E(U_q)$ depends on quality, which is a random variable, while $E(U_d)$ depends on quantity, which is a deterministic variable.

Given that innovation follows a Poisson process with arrival rate⁴ λy ,

$$\begin{aligned} E \int_0^\infty e^{-\beta\tau} \log[q_j(\tau)] d\tau &= \int_0^\infty e^{-\beta\tau} \sum_{j=0}^\infty \log[q_j] e^{-\lambda y_{j+1}\tau} \frac{(\lambda y_{j+1}\tau)^j}{j!} d\tau \\ &= \frac{\log(q_0)}{\beta + \lambda y} + \frac{\lambda y \log(\gamma q_0)}{(\beta + \lambda y)^2} + \frac{(\lambda y)^2 \log(\gamma^2 q_0)}{(\beta + \lambda y)^3} + \dots \\ &= \frac{\log(q_0)}{\beta + \lambda y} \sum_{i=0}^\infty \left(\frac{\lambda y}{\beta + \lambda y} \right)^i + \frac{\log(\gamma)}{\beta + \lambda y} \sum_{i=0}^\infty i \left(\frac{\lambda y}{\beta + \lambda y} \right)^i \\ &= \frac{\log(q_0)}{\beta} + \frac{\lambda y \log(\gamma)}{\beta^2} \end{aligned}$$

$E(U_d) = E \int_0^\infty e^{-\beta\tau} \log[L - y_{j+1}(\tau)] d\tau = \frac{\log(L-y)}{\beta}$. Setting $q_0 = 1$ we have that total expected utility $E(U) = E(U_q) + E(U_d)$ is:

$$E(U) = \frac{1}{\beta} \left(\frac{\lambda y \log(\gamma)}{\beta} + \log(L - y) \right)$$

3.7.2 Proof of Proposition (12)

I find the equilibrium by equating the zero profit condition (3.16) and the resource constraint (3.20). Because of constant returns to scale, the zero profit condition (3.16) does not depend on y . The same is not true for equation (3.20).

⁴The research effort y_j is the same for any j , since the dynamic problem is actually a sequence of identical static problems

In particular:

$$\frac{\partial V^*}{\partial y} = \frac{w \left(\lambda L \left(2r + \lambda y - \left(\frac{\lambda y}{r + \lambda y} \right)^n ((2+n)r + \lambda y) \right) - nr(\sigma - 1)(\gamma - 1)(r + \lambda L) \right)}{nr(r + \lambda y)^2(\sigma - 1)} \quad (3.29)$$

This expression is negative if H is negative, where H is:

$$H = -n\theta^n + (1 - \theta^n) \left(2 + \frac{\theta}{1 - \theta} \right) \quad (3.30)$$

with $\theta = \left(\frac{\lambda y}{r + \lambda y} \right) < 1$. H is positive if and only if $n < \left(\frac{1}{\theta^n} - 1 \right) \left(2 + \frac{\theta}{1 - \theta} \right)$. We have $\frac{\partial \left[\left(\frac{1}{\theta^n} - 1 \right) \left(2 + \frac{\theta}{1 - \theta} \right) \right]}{\partial \theta} < 0$ and, applying de L'Hospital, $\lim_{\theta \rightarrow 1} \left[\left(\frac{1}{\theta^n} - 1 \right) \left(2 + \frac{\theta}{1 - \theta} \right) \right] = n$. Therefore H is negative and so is (3.29).

Also, for $y = 0$, V^* is the discounted value of the infinite stream of profits as leader in the consumption good: $V^* |_{\{y=0\}} = \frac{(\gamma - 1)wL}{r}$. And for $y = L$, $V^* |_{\{y=L\}} = \frac{\lambda L^2 w \left(1 - \left(\frac{\lambda L}{r + \lambda L} \right)^n \right)}{nr(r + \lambda L)(\sigma - 1)} > 0$

Now we know that: (i) (3.16) is positive and constant in y ; (ii) (3.20) is positive for all $y \in [0, L]$ and it is decreasing in y . Therefore there exist an equilibrium if and only if $V^* |_{\{y=0\}} < V < V^* |_{\{y=L\}}$. In this case the equilibrium is unique. Notice that, in the end, the existence of the equilibrium depends on the value of L . Therefore an opportune choice of L assures the existence and uniqueness of the equilibrium.

Also substituting the optimal value of R&D, y^* (the one solving the social planner problem) we get (3.16) > (3.20):

$$\begin{aligned} V - V^* &= \frac{\sigma}{(\sigma - 1)\lambda} + \frac{r(\gamma - 1)}{(r + \lambda L)\log(\gamma) - r} \\ &+ \frac{(r - \lambda L \log(\gamma))^2 \left(1 - \left(\frac{\lambda L \log(\gamma) - r}{(r + \lambda L)\log(\gamma) - r} \right)^n \right)}{nr(\sigma - 1)\log(\gamma)((r + \lambda L)\log(\gamma) - r)} \end{aligned}$$

where the three terms are positive, since $L > \frac{r}{\lambda \log(\gamma)}$. This proves that the decentralized level of research effort is suboptimal and it is lower than y^* . Figure (3.5) depicts the uniqueness and the sub-optimality of the equilibrium.

The effects of r , γ , λ and σ on y_∞ follows from straightforward applications of the implicit function theorem.

3.7.3 Derivation of equation (3.22)

We start with the cost function of equation (3.13). When patent length is finite there is the possibility that some inputs are no longer protected by patent. Let $Prob(N(\tau) - N(\tau - \phi) = k)$ be the probability that in the period from $\tau - \phi$ to τ exactly k innovations are discovered. Given the time of innovation is a Poisson process, this probability is $e^{-\lambda y \phi} \frac{(\lambda y \phi)^k}{k!}$. Then the expected cost of innovation becomes:

$$\begin{aligned} E(c) &= n^{-\frac{1}{1-\sigma}} \left[nw^{1-\sigma} e^{-\lambda y \phi} \right. \\ &+ nw^{1-\sigma} e^{-\lambda y \phi} \lambda y \phi \end{aligned}$$

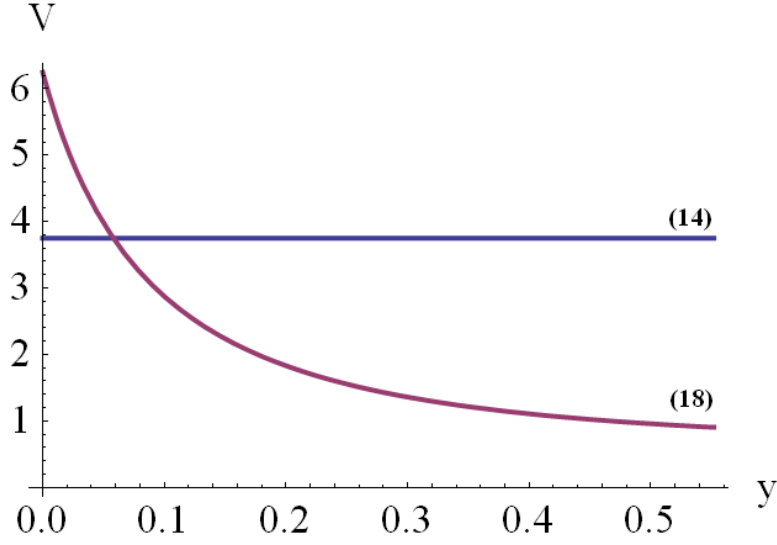


Figure 3.5: Uniqueness of equilibrium - Infinite patent length

$$\begin{aligned}
& + ((n-1)w^{1-\sigma} + p^{1-\sigma}) e^{-\lambda y \phi} \frac{(\lambda y \phi)^2}{2!} \\
& + ((n-2)w^{1-\sigma} + 2p^{1-\sigma}) e^{-\lambda y \phi} \frac{(\lambda y \phi)^3}{3!} \\
& \dots \\
& + (w^{1-\sigma} + (n-1)p^{1-\sigma}) e^{-\lambda y \phi} \frac{(\lambda y \phi)^n}{n!} \\
& + np^{1-\sigma} \left(1 - e^{-\lambda y \phi} - e^{-\lambda y \phi} (\lambda y \phi) - e^{-\lambda y \phi} \frac{(\lambda y \phi)^2}{2!} - \dots - e^{-\lambda y \phi} \frac{(\lambda y \phi)^n}{n!} \right) \Big]^{-\frac{1}{1-\sigma}} y \phi
\end{aligned}$$

Reorganizing we get equation(3.22) :

$$E(c_\phi) = n^{-\frac{1}{1-\sigma}} \left(np^{1-\sigma} - (p^{1-\sigma} - w^{1-\sigma}) e^{-\lambda y \phi} \left(n + \sum_{i=1}^n i \frac{(\lambda y \phi)^{n-i+1}}{(n-i+1)!} \right) \right)^{-\frac{1}{1-\sigma}} y \phi$$

3.7.4 Derivation of equation (3.26)

Starting from (3.13), this is analogous to derivation of equation (3.22).

3.7.5 Proof of Proposition (13)

The solution to (3.28) along the balanced growth path, given the resource constraint, is:

$$V_\phi^* = \frac{(1 - e^{-(\lambda y + r)\phi}) (\gamma - 1) w (L - y)}{r + \lambda y} + B \Pi_{\phi, ip} \quad (3.31)$$

where $\Pi_{\phi,ip}$ is as in (3.25) and

$$B = \sum_{i=0}^n \left[\left(\frac{\lambda y}{(\lambda y + r)} \right)^{i+1} \left(1 - e^{-(\lambda y + r)\phi} \sum_{j=0}^{n-i} \frac{(\lambda y \phi)^j}{j!} \right) \right] - \frac{e^{-(\lambda y + r)\phi}}{\lambda y + r} \left(\frac{(\lambda y \phi)^n}{n!} + \frac{(\lambda y)^2 \phi}{\lambda y + r} \sum_{i=2}^n \frac{(\lambda y \phi)^i}{i!} \right)$$

Equations (3.24) and (3.31) determine the steady-state equilibrium level of R&D y_ϕ . (3.24) is strictly increasing in y , while (3.31) is decreasing in y . As in the case of the infinite patent length, a large enough value of L assures that the two lines cross. Figure (3.6) shows the steady state equilibrium where the two curves cross, for benchmark value of the parameters.

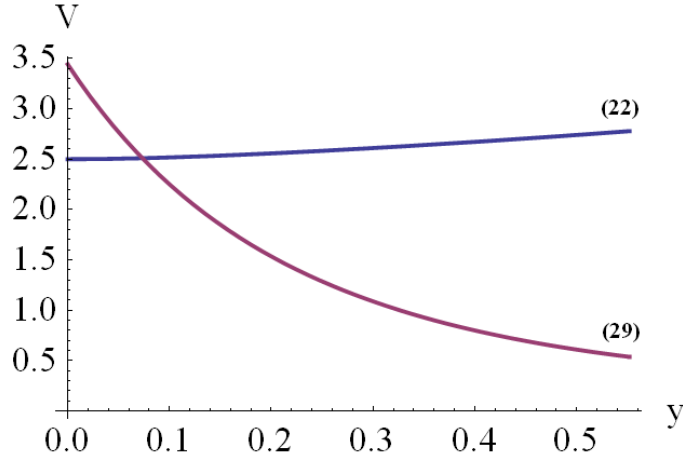


Figure 3.6: Uniqueness of equilibrium - Infinite patent length

Bibliography

- ACEMOGLU, D., AND U. AKCIGIT (2007): “State-Dependent Intellectual Property Rights Policy,” *NBER Working Paper*.
- AGHION, P., AND P. HOWITT (1992): “A Model of Growth Through Creative Destruction,” *Econometrica*, 60(2), 323–351.
- BERGSTROM, T. C. (1978): “Cournot Equilibrium in Factor Markets,” Discussion paper, University of Michigan working paper.
- BOLDRIN, M., AND D. LEVINE (2005a): “IP and Market Size,” *Mimeo*.
- BOLDRIN, M., AND D. K. LEVINE (2002): “The Case against Intellectual Property,” *American Economic Review (Papers and Proceedings)*, 92(2), 209–212.
- (2005b): “The economics of ideas and intellectual property,” *Proceedings of the National Academy of Sciences*, 102(4), 1252–1256.
- (2005c): “Intellectual Property and the Efficient Allocation of Surplus from Creation,” *Review of Economic Research on Copyright Issues*, 2(1), 45–66.
- BRENNER, S. (2009): “Optimal formation rules for patent pools,” *Economic Theory*, 40(3), 373–388.
- CHANG, H. F. (1995): “Patent Scope, Antitrust Policy, and Cumulative Innovation,” *RAND Journal of Economics*, 26(1), 34–57.
- CLARK, J., J. PICCOLO, B. STANTON, AND K. TYSON (2000): “Patent pools: a solution to the problem of access in biotechnology patents?,” Discussion paper, United States Patent and Trademark Office.
- COMMISSION ON INTELLECTUAL PROPERTY RIGHTS (2002): “Integrating Intellectual Property Rights and Development Policy: Report of the Commission on Intellectual Property Rights,” London.
- COURNOT, A. (1838): *Researches Into the Mathematical Principles of the Theory of Wealth*. Irwin (1963).
- DIXIT, A. K., AND J. E. STIGLITZ (1977): “Monopolistic Competition and Optimum Product Diversity,” *American Economic Review*, 67(3), 297–308.
- GARFINKEL, S., R. STALLMAN, AND M. KAPOR (1991): “Why patents are bad for software.,” *Issues in Science and Technology*, 8(1), 50–55.

- GRAFF, G. D., S. E. CULLEN, K. J. BRADFORD, D. ZILBERMAN, AND A. B. BENNETT (2003): "The public-private structure of intellectual property ownership in agricultural biotechnology," *Nature Biotechnology*, 21(9), 989–995.
- GREEN, J. R., AND S. SCOTCHMER (1995): "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*, 26(1), 20–33.
- GROSSMAN, G. M., AND E. HELPMAN (1991): "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 58(1), 43–61.
- HELLER, M. A. (1998): "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets," *Harvard Law Review*, 111(3), 621–688.
- HELLER, M. A., AND R. S. EISENBERG (1998): "Can Patents Deter Innovation? The Anticommons in Biomedical Research," *Science*, 280(5364), 698–701.
- HOLMSTROM, B. (1982): "Moral Hazard in Teams," *The Bell Journal of Economics*, 13(2), 324–340.
- HOPENHAYN, H., G. LLOBET, AND M. MITCHELL (2006): "Rewarding Sequential Innovators: Prizes, Patents, and Buyouts," *Journal of Political Economy*, 114(6), 1041–1068.
- JONES, C. (1999): "Growth: With or without scale effects?," *American Economic Review*, pp. 139–144.
- JUDD, K. (1985): "On the Performance of Patents," *Econometrica*, 53(3), 567–586.
- LERNER, J., AND J. TIROLE (2004): "Efficient Patent Pools," *American Economic Review*, 94(3), 691–711.
- LLANES, G., AND S. TRENTO (2009): "Anticommons and optimal patent policy in a model of sequential innovation," Working Paper 09-148, Harvard Business School.
- O'DONOGHUE, T., AND J. ZWEIMULLER (2004): "Patents in a Model of Endogenous Growth," *Journal of Economic Growth*, 9(1), 81–123.
- ROMER, P. M. (1987): "Growth Based on Increasing Returns Due to Specialization," *American Economic Review*, 77(2), 56–62.
- (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), 71–102.
- SCOTCHMER, S. (1991): "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives*, 5(1), 29–41.
- (1996): "Protecting Early Innovators: Should Second-Generation Products Be Patentable?," *RAND Journal of Economics*, 27(2), 322–331.
- SHAPIRO, C. (2001): "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting," in *Innovation policy and the economy*, ed. by A. B. Jaffe, J. Lerner, and S. Stern, vol. 1, pp. 119–150. MIT Press for the NBER.

SONNENSCHN, H. (1968): "The dual of duopoly is complementary monopoly: or, two of Cournot's theories are one," *The Journal of Political Economy*, pp. 316–318.

STOKEY, N. (1995): "R&D and Economic Growth," *The Review of Economic Studies*, 62(3), 469–489.