Patent Policy, Investment and Social Welfare.

James Bergin

Department of Economics

Queen's University

Canada

E-mail: berginj@econ.queensu.ca

April 2016

Keywords: Patents, Investment, Innovation, Welfare.

JEL Classification Numbers: D6, O3.

Corresponding author: James Bergin.

Abstract

This paper considers an environment where firms are differentiated by their technologies and where the investment response to patent policy varies across such firms. Two cases are considered. In one (the "substitutes" case), weaker firms have greater dependency on intellectual property which serves as a substitute for investment in innovation. In this case, depriving those firms of access to that technology forces them to increase investment. This has a knock-on effect of leading better firms to compete through investment, resulting in an overall increase in investment in innovation. Reversing these assumptions leads to the opposite effect — a reduction in investment. The welfare implications of such policies (as measured by consumer and producer surplus) are considered.

1 Introduction

This paper is concerned with how varying access to technologies through the patent system affects firms incentives to increase or reduce investment in R&D. For example, limiting access to technology may incentivize a firm to compensate with increased R&D investment. At the aggregate level these forces affect the competitive environment, with additional implications for overall investment. The paper studies this issue in an environment where firms have differentiated technologies and examines how varying patent restrictiveness affects those firms in terms of incentives to invest in innovation — and how this aggregates to the overall level of investment.

Access to innovation has been discussed extensively in terms of patent length and breadth. An early paper by Nordhaus [16] on innovation in a market develops optimal patent length in terms of demand elasticity, the rate at which the level of innovation responds to investment, the importance of the discovery and the cost of innovation. Gilbert and Shapiro [8] reconsider a firm's incentives when both length and breadth of the patent are considered. So, for example, if increasing breadth has substantial negative impact on welfare, it may be better to compensate the patent holder through a longer patent life — in the extreme a narrow patent with infinite life may be optimal. Gallini [7] shows that if increasing breadth is understood to mean that a patent is more difficult to work around, then patents with narrow breadth and long patent life may encourage investment in "work-arounds" that constitute wasteful imitation, suggesting that a patent should be short and broad so as to eliminate wasteful imitation. Takalo [26] considers the impact of patent width and patent length on profitability relative to social welfare. Denicolo [5] examines strategic interaction between competing firms in terms of a patent race. Firms compete in a product market and make strategic investments in R&D in a race for a patentable invention. The breadth of the patent impacts post innovation profit and the flow of social welfare. Profit gain and competitive threat from loss in the patent race determine the equilibrium level of investment. Kotowitz and Schure [14], consider a model where investment determines the probability of success in innovation and evaluate the optimal patent length in terms of the trade-off between expected profitability and risk (since innovation is not guaranteed). In this context, lower expected profitability or higher risk justify a longer patent length. Another strand of literature considers patent protection in the context of cumulative innovation where the natural flow of innovation has one innovation build on another. Each innovation in the flow is one of many so the issue of allocating appropriate incentives cannot be considered for each innovation in isolation (see Green and Scotchmer[10] and Scotchmer [22]). O'Donohue, Scotchmer and Thisse [17] study patent performance in a dynamic environment where breadth of a patent can impede both imitation of the innovation and improvement of the innovation. Bessen and Maskin [3] provide a broad discussion of the types of incentives that arise in environments with sequential and cumulative innovation.

This paper studies the optimality of patent policy from a different perspective: how restrictiveness of access to IP (intellectual property) affects the investment incentive for firms with differentiated (and possibly non-comparable) technologies, and how this feeds through to aggregate investment. In this setting there is ongoing pressure to remain competitive or gain some lead time in new technology through investment in innovation where firms are differentiated by technology. In the model, restrictiveness is characterized as reduced access or ability of an innovator to exploit existing technologies (of other firms.) In practical terms, this is represented as reduced profitability and reduced ability to innovate. So, for example, increasing patent length reduces a firms capacity to incorporate extant technology either in its own products or in its development process. But, similarly, increased breadth of patents would impose a similar restriction on firms. Whatever the interpretation (length will be the interpretation here), the key assumptions relate to profitability and innovativeness — how reduced access to IP impacts current profit and success in innovation.

The main features of the model are the following. A large number of innovators invest and generate profit from period to period in an environment with heterogeneous technologies.[1] Profitability and the innovation success depends on a firms' own level of technological advancement, its level of investment and the distribution of competitors characteristics. Different firms have distinct technologies (which are generally not comparable) and variations in patent length or restrictiveness affect these firms investment incentives differently: the impact on the incentives of a firm depend on its technology. The cumulative effect alters the competitive environment and changes the need for (or pressure on) a firm to invest. And because firms have different technologies, the effect varies from firm to firm. At the aggregate level, a policy which directly encourages innovation investment raises competitive pressure which in turn may encourage or deter investment, making the overall effect difficult to determine. The paper studies how these combined forces affect the level of individual and aggregate investment, and hence the overall technologically improvement or dis-improvement. In the environment considered here, in general there is a positive externality to investment so that in equilibrium, investment is below the socially optimal level, for any given patent length. In such circumstances, shifts in aggregate investment may be related to changes in social welfare.

The model is described in section 2 — the structure of technology and of innovation over time, the profitability of the firm and equilibrium in the model. Section 2.1 considers the process of innovation and the major factors determining innovation success. The remainder of the section presents the model details. Section 3 motivates the results developed in section 4. Two cases are identified where the impact is unambiguous. In one, the impact of tightening IP access has the effect of depriving weaker firms of technology which leads them to "substitute" by developing or improving their own technology: innovation substitutes for lack of access to technology. In the second (and possibly less likely) case it is the better firms that are most impacted by the tightening of IP access and this turns out to adversely impacts the level of investment in innovation. In this case, good firms are advantaged and protected from competition by patented technology. Section 5 considers the issues from a welfare perspective and section 6 concludes.

[1]The large numbers assumption has advantages and disadvantages. It simplifies the study of dynamics over time since individual firm decisions do not affect aggregate levels, so the determination of optimal decisions at the firm level is simplified; and it avoids the complexities of strategic behavior that arise with small numbers. At the same time, this limits the scope to study such issues as preemption or strategic blocking of competitors that require more detailed modeling at the microeconomic level.

2 The Model

Firms earn profit each period and invest to maintain future profitability. At any point in time a firm has its own current technology and makes investment decisions that improve its technology over time. The population of technologies defines the competitive environment for the firm. Technology is protected by a patent regime, but firms may use technologies outside patent protection and may benefit from the presence of other technologies (by limited imitation, adaptation and so forth.) To allow for the possibility that one firm may be better than another in some respects, and worse in others, technology is multidimensional. This formulation allows different firms or products to have different strengths and weaknesses.[2]

Denote a firm's technology by $\alpha \in \Lambda$, where Λ is the set of all possible technologies. There are a continuum of firms with distribution of technologies in the market denoted μ, or μ_t to denote the distribution of technologies at time t, a probability measure on Λ.[3] The firm operates in an environment represented by the current and historical distribution of technologies in the population. The history of technology distributions is given by $\boldsymbol{\mu}_t = \{\mu_\tau\}_{\tau=t}^{-\infty} = (\mu_t, \mu_{t-1}, \ldots)$.

Technology evolves over time. A firm's technology α, the level of its investment, i, and the prevailing record of technology affect the quality of innovation of the firm. In addition, restricted access to patented technology limits the use of other technologies in the population. This is measured by patent length ℓ, where ℓ may also be viewed more generally as a measure of restrictedness of IP policy.[4]

2.1 Technology and Innovation

How should innovation be modeled? The following (somewhat lengthy) discussion describes key features of the innovation process which in turn suggests the specification used in this paper. In sum, innovation for a firm is (a) generally history dependent, (b) has complementary or interdependent spillover effects from other innovators and (c) is multidimensional in nature.[5] Ultimately, consideration of these features leads to a Markovian model of innovation below.

Most innovations are minor, possibly improving or modifying an existing idea or product and often move quickly from conception to use. From time to time a major innovation arrives with great impact (such as transistors, microchips, nanotubes). Major innovations typically impact a large range of industries and often take years or even decades to move from discovery of the innovation to application.[6] The modern cell

[2]So, for example, SDRAM has a range of technical features such as bandwidth, latency, read time and so on. Similarly, technical features of a cellphone include memory size, screen pixel density, graphical interface, supporting applications and so forth.

[3]If α and α' are in the support of μ_t, they represent two technologies in operation at time t.

[4]In the paper, ℓ will be thought of as patent length, with larger values corresponding to greater IP protection. More generally, ℓ may be interpreted as a measure of the degree to which access to IP is restricted. For the purpose of the paper, what matters is that increases in ℓ impact profit and innovation in specific ways.

[5]In differentiated product environments, innovation is often modeled as a Poisson arrival rate (potentially a function of current investment), so that the innovation arrival is independent across producers and independent of history.

[6]Watt's idea for an external steam condenser to improve steam engine efficiency took nearly 10 years from the granting in the patent in 1769 to develop a satisfactory operational model. The theory for transistors was developed in the 1920's, demonstrated in early designs in the late 1940's and not commercially significant for another 20 years. The theory underlying nanotubes began in 1952 with ongoing research through the 1990's and rapid commercialization in the 2000's. Techniques for

phone is considered a major recent innovation, but the core supporting technologies have been in existence and evolving over many years: capacitive touch screens began use at CERN in 1973, instant messaging first appeared in 1986, the operating systems OS-X and Android are based on BSD Unix (1977) and Linux (1991) respectively (See [1, 25]). In such cases the benefit to the innovator may come from being ahead of competitors in terms of experience working with the new product or process, or from the reputation effects as the originator of the innovation. See [11, 27] for detailed discussion of the invention/innovation process.

Innovation success depends on accumulated knowledge and complementary research. Furman and Stern [6] refer to the importance of the cumulative and overlapping aspects of innovation in promoting growth. Cumulative innovation can occur in a variety of ways. A single innovation may support or be essential to multiple subsequent innovations; multiple innovations may be utilized as a group to support a single subsequent innovation; or innovation may progress as a ladder with each innovation building on its predecessor.[7] See Scotchmer [23] for a categorization and discussion.

Innovation in one field is informed by developments in related fields. Crossover of ideas is part of the innovation process. Poetz, Franke and Schreier [19] note how innovations in the mining industry were applied to escalator installation in shopping malls. Scotchmer [21] points to the importance externalities and spillovers in the pace of innovation. The key idea in an innovation can often find application in other disparate fields. For example, improvements in battery technology allow for advances across the entire range of mobile electronic devices. Thus, for example, one might expect the rate of innovation to be correlated across similar fields as innovators learn from each other. Even with complete intellectual property protection, advances in one field may improve innovative progress in other fields. Considering innovation in small and medium sized enterprises in Poland, Stanisławski and Lisowska [24] examine the extent to which lack of openness (exchange and use of ideas from other companies or sources) limits growth in Poland. The value of obtaining problem insights or solutions by drawing on external problem solvers who may be "contextually distant" (similar problems faced in different environments) is noted by Poetz and Prúgl [18].

Innovation depends not only on investment, but also on the resources available to do research, the quality and experience of the research team, available equipment and so forth. These factors are fundamental to success for many projects. The state of an innovator or firm's technology affects the capacity to generate innovation so that the incremental rate of innovation is increasing in the cumulative level of investment. Innovation is often multidimensional with a set of problems to be solved to achieve the end result. Improvements in microchip design occur at the same time across a range of measures such as bandwidth, response time, parallel processing capacity, power consumption, physical size, heat generation, and so forth, with different facets of the innovation unavoidably connected.

Patents and possibly other restrictions limit the extent to which a firm may exploit the population of

production of graphene were demonstrated in the early 2000's, but so far commercialization is not significant. "Pushmail", an essential foundation of Blackberry's success in the early 2000's was developed in the 1980's (and a commercial failure at the time).

[7] Bessen and Maskin [4] examine how both the sequential and cumulative aspects of innovation may increase the rate of discovery.

technologies. Patent protection limits direct duplication of a technology, but the possibility remains for partial imitation, work-arounds and exploitation of ideas in legally protected technology. Thus, not only are (old) ideas outside the period of patent protection available to a firm, but currently protected ideas or methods may be exploitable to a degree by a firm to raise profit or augment its rate of innovation. The degree to which this is possible depends on the extent to which intellectual property (IP) rights are assigned and enforced. Let ℓ be a real number measuring the intensity of IP rights enforcement (policy) with larger values of ℓ denote more effective enforcement.[8]

These considerations argue for a general model of innovation at the firm level — where the firm's IP environment, its current state of technology and its level of investment jointly determine its innovation capacity. The pair $(\boldsymbol{\mu}_t, \ell)$ describes the IP structure facing a firm; and that along with the firm's own state of technology, α, describe firm α's knowledge environment. The firm's technology evolves stochastically, depending on $(\boldsymbol{\mu}_t, \ell)$, the firms technology, α, and it's level of investment, i — denote the distribution over α's next period technology by $P(\cdot \mid \boldsymbol{\mu}_t, \ell, \alpha, i)$.[9]

Remark 2.1: This formulation of innovation allows for complex history dependence, interaction of ideas between innovators, the modeling of improvements in existing technologies and also, the arrival of new technologies. If, for example, each firm has a technology given by an n-dimensional vector, $\alpha = (\alpha_1, \ldots, \alpha_n)$ technology improvement may be represented by $\tilde{\alpha} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n)$, $\tilde{\alpha}_i \geq \alpha_i$ whereas a new technology may be modeled as an increase in dimension $(\alpha_1, \ldots, \alpha_n) \to (\alpha_1, \ldots, \alpha_n, \alpha_{n+1})$. To permit comparison of different firms technologies, take A to be an ordered space, with order $\succeq$ (see appendix I for discussion). Throughout the paper technology distributions are compared in terms of first order stochastic dominance. Assume that the space of technologies, Λ is an ordered set. Given two measures $\mu, \nu \in \mathcal{P}(\Lambda)$, μ first order dominates ν, written $\mu \succcurlyeq \nu$ if and only if for all measurable increasing functions $g : \Lambda \to \Re$, $\int g d\mu \geq \int g d\nu$. Note that "$\succeq$" is an ordering on technologies, Λ, whereas "$\succcurlyeq$" is an ordering on distributions over technologies $\mathcal{P}(\Lambda)$. ■

Throughout the discussion it is assumed that: (i) having a better technology or investing more raises the probability of drawing a better technology, (ii) having less access to technology (through longer patent life, ℓ) lowers the probability of drawing a good technology, and (iii) better ambient technology (μ_t) improves a firm's ability to innovate. Formally, $P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$ is weakly increasing in α, and i; and weakly decreasing in ℓ (in terms of first order stochastic dominance).[10] Assumption (ii) captures the impact of patent length on the firm's ability to innovate as the firm's freedom to incorporate other technologies is reduced. (As, for example, when firms must create "workarounds" to achieve a function available in a patented technology.) Define $\boldsymbol{\mu}'_t \succeq \boldsymbol{\mu}'_t$ coordinate-wise to mean that $\mu'_{t-j} \succeq \mu_{t-j}$ for all $j \geq 0$. The kernel, $P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$, is assumed to be increasing in $\boldsymbol{\mu}_t$ — in the sense that if $\boldsymbol{\mu}'_t$ dominates $\boldsymbol{\mu}_t$ coordinate-wise, written $\boldsymbol{\mu}'_t \succcurlyeq \boldsymbol{\mu}_t$, then other things equal, a better distribution is drawn conditional on $\boldsymbol{\mu}'_t$ than $\boldsymbol{\mu}_t$. Better technology in the

[8]Take ℓ to be a non-negative real number.

[9]Large firms may be modeled by having atoms in the distribution μ_t or arise endogenously if the kernel, P, may have atoms at some or all profiles $(\boldsymbol{\mu}_t, \ell, \alpha, i)$ (in the simplest case with finite support.) However, that would greatly increase the technical complexity of analyzing the model as individual firms can then affect the aggregate state, adding an extra layer of complexity.

[10]For example, $\alpha' \succeq \alpha$ implies first order stochastic dominance: $P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha', i, \ell) \succcurlyeq P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$.

population and better technology in the public domain improves the firm's success in innovation. Finally, although a firm can fall behind competitors, a firm's technology cannot dis-improve over time: if drawing α' is possible for α, then $\alpha' \succeq \alpha$.[11,12] Assume that the measure, μ_t, has no atoms: each firm has probability or measure 0, so that no firm has strategic market power. This implies that each firm is negligible and eliminates strategic considerations from the model. Furthermore, since the model is concerned with the dynamics of investment over time, this assumption makes state variables independent of individual behavior and simplifies the computations.

The next sections (2.2 and 2.3) describe the determinants of profit and formulate the model of innovation. Following that, the firm's optimizing problem and equilibrium behavior are considered (sections 2.4 and 2.5).

2.2 Profit and revenue

These variables, $(\boldsymbol{\mu}_t, \ell, \alpha)$, affect a firm in two ways, through the firm's profit and through its innovation. Profit, π, is modeled as a function of the same parameters, $\pi(\boldsymbol{\mu}_t, \ell, \alpha)$, the profit resulting from market equilibrium.

Remark 2.2: At this level, these functions, π and P, represent a reduced form model. This has advantages and disadvantages. While the model lacks a detailed description of the environment, it accommodates a fairly broad class of models of innovation. ■

Examples (2.1) and (2.2) illustrate the derivation of a profit function in a single market environment and in multi-market environment with cross price effects.

Example 2.1: This example derives a profit function from a simple single market model. Let firm α have cost given by $\frac{1}{2}c(\alpha, \boldsymbol{\mu}_t)q^2$, where for any $(\boldsymbol{\mu}_t, \alpha)$, $c(\alpha', \boldsymbol{\mu}_t) \leq c(\alpha, \boldsymbol{\mu}_t)$ if $\alpha' \succeq \alpha$, and $c(\alpha, \boldsymbol{\mu}'_t) \leq c(\alpha, \boldsymbol{\mu}_t)$ if $\boldsymbol{\mu}'_t \succcurlyeq \boldsymbol{\mu}_t$. Therefore, firm α's profit at price P is given by $\max_q Pq - \frac{1}{2}c(\alpha, \boldsymbol{\mu}_t)q^2$. The solution, q, satisfies $P - c(\alpha, \boldsymbol{\mu}_t)q = 0$, determining individual output q: $q(\alpha) = \frac{P}{c(\alpha, \boldsymbol{\mu}_t)}$. To simplify further, let cost be given by $c(\alpha, \boldsymbol{\mu}_t) = [\varphi(\alpha)g(\boldsymbol{\mu}_t)]^{-1}$ where $\varphi(\alpha)$ is a scalar quality-efficiency index of technology (increasing in α), and g an increasing real valued function of $\boldsymbol{\mu}_t$ reflecting the impact of ambient technology on a firm's efficiency. Thus, $q(\alpha) = P\varphi(\alpha)g(\boldsymbol{\mu}_t)$. Aggregate supply at price P is

$$Q = \int q(\alpha)\mu_t(d\alpha) = P\int \frac{1}{c(\alpha, \boldsymbol{\mu}_t)}\mu_t(d\alpha) = Pg(\boldsymbol{\mu}_t)\int \varphi(\alpha)d\mu_t = Pg(\boldsymbol{\mu}_t)\bar{\varphi}(\mu_t) \qquad (1)$$

where $\bar{\varphi}(\mu) = \int \varphi(\alpha)d\mu$. Let market demand at time t be $P_d(Q, \boldsymbol{\mu}_t) = d(\boldsymbol{\mu}_t)Q^{-\beta}$, $\beta > 0$, so that elasticity of demand is $\frac{1}{\beta}$ and where $d(\boldsymbol{\mu}_t)$ measures the impact of ambient technology on demand (for example,

[11] If $\alpha' \in \operatorname{supp} P(\cdot \mid \boldsymbol{\mu}, \alpha, i, \ell)$, then $\alpha' \succeq \alpha$, where given a measure υ on Λ, supp υ is the support of υ

[12] If $P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$ has support $\{\alpha\}$ when $i = 0$, then the firm cannot improve without investment.

capturing the impact of quality on demand). Then, from equation (1), market clearing gives

$$Q = P_d(Q, \boldsymbol{\mu}_t)g(\boldsymbol{\mu}_t)\bar{\varphi}(\mu_t) = d(\boldsymbol{\mu}_t)Q^{-\beta}g(\boldsymbol{\mu}_t)\bar{\varphi}(\mu_t)$$

with equilibrium quantity $Q = [d(\boldsymbol{\mu}_t)g(\boldsymbol{\mu}_t)\bar{\varphi}(\mu_t)]^{\frac{1}{1+\beta}}$, and price $P = d(\boldsymbol{\mu}_t)[d(\boldsymbol{\mu}_t)g(\boldsymbol{\mu}_t)\bar{\varphi}(\mu_t)]^{-\frac{\beta}{1+\beta}}$. Profit for firm α is

$$\pi(\boldsymbol{\mu}_t, \alpha) = Pq(\alpha) - \frac{1}{2}c(\alpha, \boldsymbol{\mu}_t)q(\alpha)^2 = [P - \frac{1}{2}c(\alpha, \boldsymbol{\mu}_t)q(\alpha)]q(\alpha) = \frac{1}{2}P^2 q(\alpha)$$

Thus,

$$\begin{aligned}
\pi(\boldsymbol{\mu}_t, \alpha) &= \frac{1}{2}d(\boldsymbol{\mu}_t)^2[d(\boldsymbol{\mu}_t)g(\boldsymbol{\mu}_t)\bar{\varphi}(\mu_t)]^{-\frac{2\beta}{1+\beta}}g(\mu_t)\varphi(\alpha) \\
&= \gamma(\boldsymbol{\mu}_t)\varphi(\alpha), \ \text{ with } \ \gamma(\boldsymbol{\mu}_t) = \frac{1}{2}[d(\boldsymbol{\mu}_t)^2 g(\boldsymbol{\mu}_t)^{1-\beta}\bar{\varphi}(\mu_t)^{-2\beta}]^{\frac{1}{1+\beta}}
\end{aligned}$$

This determines current profit $\pi(\boldsymbol{\mu}_t, \alpha)$, or making the dependence on ℓ of γ gives $\pi(\boldsymbol{\mu}_t, \ell, \alpha)$. Expected future profit depends on the transition kernel. Over time a firm's technology evolves according to the transition kernel $P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$ so that over time the profit flow depends on this and the profit function. To illustrate, suppose the transition kernel is a weighted average of two distributions:

$$P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) \quad = \quad \rho(\boldsymbol{\mu}_t, \alpha, i, \ell)F(d\tilde{\alpha}) + [1 - \rho(\boldsymbol{\mu}_t, \alpha, i, \ell)]G(d\tilde{\alpha})$$

with $F \succcurlyeq G$, and $\rho(\boldsymbol{\mu}_t, \alpha, i, \ell) \in [0, 1]$ increasing in $\boldsymbol{\mu}_t$, α and i. The expected profit one period ahead is:

$$\begin{aligned}
\int \pi(\boldsymbol{\mu}_{t+1}, \tilde{\alpha}, \ell)P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) &= \gamma(\boldsymbol{\mu}_t)[\rho(\boldsymbol{\mu}_t, \alpha, i, \ell)\bar{\varphi}_F + (1 - \rho(\boldsymbol{\mu}_t, \alpha, i, \ell))\bar{\varphi}_G] \\
&= \gamma(\boldsymbol{\mu}_t)[\rho(\boldsymbol{\mu}_t, \alpha, i, \ell)(\bar{\varphi}_F - \bar{\varphi}_G) + \bar{\varphi}_G] \qquad (2)
\end{aligned}$$

with $\bar{\varphi}_F = \int \varphi dF$, $\bar{\varphi}_G = \int \varphi dG$ and where $\boldsymbol{\mu}_{t+1} = (\mu_{t+1}, \boldsymbol{\mu}_t)$ with μ_{t+1} determined by aggregate investment (see section (2.3)). ◇

EXAMPLE 2.2: Consider an environment where each firm, identified by its technology α, faces a demand which depends on the pricing profile of other firms and has constant marginal cost $c(\alpha)$. Suppose a one-period demand facing firm α is:

$$q_d(\alpha) = a(\alpha) + b(\alpha)p(\alpha) + \int \gamma_\alpha(\tilde{\alpha})p(\tilde{\alpha})\mu(d\tilde{\alpha})$$

and supply is a function of the price of firm α's product, $p(\alpha)$: $S(p(\alpha), \alpha)$. In equilibrium,

$$S(p(\alpha), \alpha) = a(\alpha) + b(\alpha)p(\alpha) + \int \gamma_\alpha(\tilde{\alpha})p(\tilde{\alpha})\mu(d\tilde{\alpha})$$

with equilibrium price function $p^*(\alpha)$. Equilibrium quantity is then $q(\alpha) = a(\alpha) + b(\alpha)p^*(\alpha) + \int \gamma_\alpha(\tilde{\alpha})p^*(\tilde{\alpha})\mu(d\tilde{\alpha})$. Making the dependency of the equilibrium price and quantity functions on the aggregate distribution explicit, $q(\boldsymbol{\mu}_t, \alpha, \ell)$, $p^*(\boldsymbol{\mu}_t, \alpha, \ell)$, and allowing cost to depend on the current aggregate parameters, $q(\boldsymbol{\mu}_t, \alpha, \ell)$, the profit of α is then:

$$\pi(\boldsymbol{\mu}_t, \alpha, \ell) = p^*(\boldsymbol{\mu}_t, \alpha, \ell)q(\boldsymbol{\mu}_t, \alpha, \ell) - c(\boldsymbol{\mu}_t, \alpha, \ell)q(\boldsymbol{\mu}_t, \alpha, \ell)$$

$\diamond$

The next definition concerns monotonicity of the profit function in technology.

Definition 2.1: *Profit is non-decreasing in technology if* $\pi(\boldsymbol{\mu}_t, \alpha, \ell)$ *is non-decreasing in* $\boldsymbol{\mu}_t$.

REMARK 2.3: In example (2.1), $d(\boldsymbol{\mu}_t)$ is a product improvement factor that raises demand, $g(\boldsymbol{\mu}_t)$ is a cost factor reflecting the impact of the aggregate distribution on cost and $\bar{\varphi}(\mu_t)$ is the mean of the individual specific component of cost. Profit is increasing in the efficiency and demand values g and d. Therefore, if the impact of innovation raises demand or reduces common costs the effect is to raise profit. In contrast, if the effect is to make firms individually more cost competitive, $\bar{\varphi}$ increases, each individual firm is relatively worse off and the impact on its profit is negative. $\blacksquare$

Given the technology described in section (2.1), section (2.3) describes how technology evolves over time.

2.3 The Evolution of Technology

The investment strategies of firms in conjunction with the transition kernel, P, move the state of the system forward over time. Firms investment strategies are represented by a joint distribution, τ, on $(i, \alpha) \in I \times \Lambda$, written $\tau \in \mathcal{M}(I \times \Lambda)$. Conditioning on α, $\tau(di \mid \alpha)$, gives the distribution over investment of firm α. Given the extant distribution over technologies, μ, for consistency, if $\tau \in \mathcal{M}(I \times \Lambda)$ the marginal distribution of τ on Λ should coincide with μ: $\text{marg}_\Lambda \tau = \mu$. Let $\mathcal{C}(\mu) = \{\tau \mid \text{marg}_\Lambda \tau = \mu\}$, the set of distributions on $I \times \Lambda$ with marginal μ on Λ. The distribution of technologies evolves as:

$$\mu_{t+1}(\cdot) = \psi(\cdot \mid \boldsymbol{\mu}_t, \tau_t, \ell) \overset{\text{def}}{=} \int_{i_t, \alpha_t} P(\cdot \mid \boldsymbol{\mu}_t, \alpha_t, i_t, \ell)\tau_t(di_t \mid \alpha_t)\mu_t(d\alpha_t) = \int_{i_t, \alpha_t} P(\cdot \mid \boldsymbol{\mu}_t, \alpha_t, i_t, \ell)\tau_t(di_t \times d\alpha_t) \quad (3)$$

So, given the current distribution on technologies, μ_t, if α_t invests according to the strategy $\tau_t(\cdot \mid \alpha_t)$, then next period the aggregate distribution on technologies is given by μ_{t+1}.[13]

2.4 The Firm's decision.

Firms make period by period decisions on production, and in addition, make investment decisions to develop future technology. The current production decision arises in the period by period market equilibrium and determines current profit, π. The investment decision generates current cost but improves the competitive position of the firm in subsequent periods. For some of the discussion, it is useful to write the present value at time t of the payoff flow to a firm, α, optimizing in each period from this point on — given the distribution up to the present, $\boldsymbol{\mu}_t$, and a sequence of aggregate distributions $\boldsymbol{\tau}^t = \{\tau_s\}_{s=t}^{\infty}$ as $v(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)$. The individual optimization problem may expressed in a Bellman equation as:

$$v(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell) = \max_i \{\pi(\boldsymbol{\mu}_t, \alpha, \ell) - r(i) + \delta \int v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell) P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)\} \tag{4}$$

where $\boldsymbol{\mu}_{t+1} = (\boldsymbol{\mu}_t, \mu_{t+1})$ with μ_{t+1} given by equation (3). The function v is increasing in α: a firm with higher α can imitate the investment strategy of one with lower α but enjoy lower cost and stochastically better technology draws.

2.5 Equilibrium.

For i to be an optimal solution in equation (4) requires (assuming an interior solution):

$$-r'(i) + \delta \lim_{i' \to i} \left[\frac{1}{i' - i}\right] \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)[P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i', \ell) - P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)] = 0$$

The first order condition for i is:[14]

$$-r'(i) + \delta \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell) \Delta_i P(\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) = 0. \tag{5}$$

[13] In this environment, improvement in technology overall results from the flow of individual discoveries — with each individual discovery insignificant relative to the overall volume of discovery. One possible extension of this model is to allow for "paradigm shift" discoveries which revolutionize an industry. The formulation used here can accommodate such an extension provided that big breakthroughs are unanticipated and do not result in market power (technically where a firm becomes an atom in the distribution). In such a formulation, there is positive probability of a breakthrough discovery in any period (some firm will have a major discovery or development), but no single firm can guarantee that it will have such a discovery with positive probability. In this case, revolutionary innovations are unanticipated and hence don't directly affect the investment incentives of firms.

[14] Assume that $\frac{1}{i'-i}[P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i', \ell) - P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)]$ converges weakly to a signed measure $\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$ as $i' \to i$. Further, for the second order condition, assume that $\left[\frac{1}{i'-i}\right][\Delta_i P(B \mid \boldsymbol{\mu}_t, \alpha, i', \ell) - \Delta_i P(B \mid \boldsymbol{\mu}_t, \alpha, i, \ell)]$ converges weakly to a signed measure, $\Delta_{ii} P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$, as $i' \to i$.

The second order condition for an optimum is then:

$$-r''(i) + \delta \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)\Delta_{ii} P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) < 0. \tag{6}$$

Considering equation (5), since v is increasing in $\tilde{\alpha}$, and P in α (in first order stochastic dominance terms), the optimal value of i increases in α.

A sequence of strategies $\bar{\tau}^t = \{\bar{\tau}_{t+j}(\cdot \mid \alpha_{t+j})\}_{j\geq 0}$ is an equilibrium if for any t, for each $j \geq 0$, $\bar{\tau}_{t+j} \otimes \mu_{t+j}$ has support $\{(i_{t+j}(\alpha), \alpha) \mid \alpha \in A\}$, where $i_t(\alpha)$ solves (5) at $\bar{\tau}^t$ and $i_{t+j}(\alpha)$ solves the analogous condition at time $t + j$. Establishing the existence of equilibrium is straightforward using arguments from [2] or [13].

3 Investment and IP access: Motivating discussion.

Considering the impact of patent length or restrictiveness of IP access on investment, there are a number of channels through which this occurs. Individual incentives are altered as variations in technology access affect profitability and efficacy of investment; and at the aggregate level, population distributions adjust to changes in individual behavior again impacting incentives. At the firm level (setting aside aggregate effects), the covariation of the pair (ℓ, i) — the movement of investment in response to variation in patent length or patent restrictiveness — may be viewed from the perspective of super or sub-modularity. For example, an increase in ℓ which induces an increase in i reflects positive co-variation: investment increases in response to higher values of ℓ.

To clarify this point, fixing the aggregate distributions, let

$$\bar{v}(\ell, i, \alpha) = \int v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell) P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)\}$$

be the continuation payoff to α and consider the direct impact of (ℓ, i) variations on $\bar{v}$.

Suppose that $\bar{\imath} > \underline{i}$ and $\bar{\ell} > \underline{\ell}$. Given (μ, α), $\bar{v}(\ell, i, \alpha)$ is supermodular in (ℓ, i) if $\bar{\ell} \geq \underline{\ell}$, $\bar{\imath} \geq \underline{i}$, then:

$$\bar{v}(\bar{\ell}, \bar{\imath}, \alpha) + \bar{v}(\underline{\ell}, \underline{i}, \alpha) \geq \bar{v}(\bar{\ell}, \underline{i}, \alpha) + \bar{v}(\underline{\ell}, \bar{\imath}, \alpha)$$

or equivalently, $\bar{v}(\bar{\ell}, \bar{\imath}, \alpha) - \bar{v}(\bar{\ell}, \underline{i}, \alpha) \geq \bar{v}(\underline{\ell}, \bar{\imath}, \alpha) - \bar{v}(\underline{\ell}, \underline{i}, \alpha)$. The function is submodular if the inequality is reversed. For a supermodular function, an increase in the value of ℓ raises the marginal value of investment, i. In terms of derivatives, supermodularity gives $\frac{\partial}{\partial \ell}\{\frac{\partial \bar{v}}{\partial i}\} \geq 0$, so that in a sense, the co-movement is complementary. However, in the present context increasing ℓ is a *tightening* of patent policy that encourages an increase in investment by raising the marginal value of investment. Therefore, from an incentive perspective, extra investment is substituting for the tightening of patent policy: investment is *substituting* for IP outside the firm. Conversely, with submodularity, $\frac{\partial}{\partial \ell}\{\frac{\partial \bar{v}}{\partial i}\} \leq 0$, so that a reduction in access to IP outside the firm (an increase in ℓ) lowers the marginal value of investment and gives the firm an incentive to reduce investment, or, alternatively, an increase in access to external IP causes an increase in investment, so that investment is *complementary* to IP access. Thus, it is natural to use the term substitutability to reflect a situation when tightening patent policy encourages an increase in investment, and complementarity in the case where tightening patent policy encourages a reduction of investment. In section (4.1) the case where investment substitutes for IP access is discussed; section (4.2) considers the case where the two are complementary.

To gain additional insight into the conditions of sections (4.1) and (4.2), suppose that P has a density, f, so that

$$\bar{v}(\mu, \ell, i, \alpha) = \int v(\tilde{\alpha}, \mu, l) f(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha}$$

Then assuming differentiability, $\frac{\partial \bar{v}(\mu,\ell,i,\alpha)}{\partial \ell} = \int v_l(\tilde{\alpha}, \mu, l) f(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha} + \int v(\tilde{\alpha}, \mu, l) f_l(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha}$ and

$$\begin{aligned}\frac{\partial^2 \bar{v}(\mu, \ell, i, \alpha)}{\partial i \partial \ell} &= \int v_l(\tilde{\alpha}, \mu, l) f_i(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha} + \int v(\tilde{\alpha}, \mu, l) f_{li}(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha} \\ &= \int [v_l(\tilde{\alpha}, \mu, l) f_i(\tilde{\alpha} \mid \mu, \ell, \alpha, i) + v(\tilde{\alpha}, \mu, l) f_{li}(\tilde{\alpha} \mid \mu, \ell, \alpha, i)] d\tilde{\alpha}\end{aligned}$$

so that the substitute/complement conditions depend on the behavior of $v_l(\tilde{\alpha}, \mu, l) f_i(\tilde{\alpha} \mid \mu, \ell, \alpha, i)$ and $v(\tilde{\alpha}, \mu, l) f_{li}(\tilde{\alpha} \mid \mu, \ell, \alpha, i)$ such that $\frac{\partial^2 \bar{v}(\mu,\ell,i,\alpha)}{\partial i \partial \ell}$ has an unambiguous sign.

From this expression one sees that key determinants of the sign of $\frac{\partial^2 \bar{v}(\mu,\ell,i,\alpha)}{\partial i \partial \ell}$ are: (a) the marginal productivity of investment, f_i, (b) the marginal impact on profitability of varying ℓ on different firms (whether v_ℓ is increasing or decreasing in α), (c) the impact on marginal productivity as ℓ varies, f_{li}.[15] These conditions are developed in detail below.The main additional consideration in these calculations is that the consequent aggregate distributional shifts need to be factored in to the calculations, complicating the analysis significantly.

4 The Impact of Patent Length on Investment.

The effect of lengthening patent life is to limit the available technology for use. What is the impact of such a change on welfare? Because the socially optimal level of investment is higher than that arising in competitive equilibrium, whether lengthening patent length is beneficial or not depends on the impact such changes have on investment. The results to follow identify two cases.

When low technology firms are more dependent than high technology firms on the use of technology protected by patent, then, subject to conditions, the impact of lengthening patent life is to force those firms to greater research effort (by depriving them of access to previously unrestricted technology.) And this has a knock-on effect of increasing the competitive pressure on good firms, leading them to also raise

[15] For example, suppose that the marginal productivity of investment is positive ($\int g(\tilde{\alpha}) f_i(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha}$ for g increasing). If the impact of tightening IP policy is to lower payoffs more for weaker firms (so that v_ℓ is increasing in α), then $\int v_l(\tilde{\alpha}, \mu, l) f_i(\tilde{\alpha} \mid \mu, \ell, \alpha, i) d\tilde{\alpha}$ is positive implying an incentive to raise investment.

investment. As a result, overall investment in R&D increases and this raises social welfare (because of positive externalities in investment, see section (4.1)). Put differently, technically weak firms are making greater use of publicly available technology. Reducing access to older technology forces those firms to invest. Firms must "substitute" investment to compensate for reduced access to technology. In the second case, this situation is reversed and increasing patent length has greater impact on better firms.

These results suggest that patents are beneficial when, as a result of the need to compete, they spur R&D and hence innovation among weaker firms. To the extent that disallowing a firm from use of the discovery of others ultimately forces that firm to greater investment in R&D the effect of patents is beneficial because it also forces better firms (through competition) to invest more. When this pressure is absent, the opposite occurs.

4.1 Investment and Patented Knowledge as Substitutes.

When the (negative) impact of lengthening patent life and reducing access to technology is greatest on low technology firms, technology improvement may be considered a substitute for the patented technology, and since investment improves technology, investment becomes a substitute for patented technology. In the assumptions to follow, this is expressed by having firms with weaker technology suffer greater impact both in terms of profitability and quality of innovation (S-i). Furthermore, if increasing patent length raises the marginal product of investment then investment can compensate from the loss of access to patented technology (S-ii). In such circumstances, it turns out that the overall effect of increasing patent length is to encourage weak firms to invest and this encourages better firms to invest, resulting in a rise in aggregate investment (Theorem (4.1)).

(S-i) Increasing patent life, ℓ, or worsening technology, $\boldsymbol{\mu}$ has a greater impact on weaker firms.[16]

(a) An increase in ℓ or fall in $\boldsymbol{\mu}$ lowers current profits of better firms less than weaker firms:

$$\pi(\boldsymbol{\mu}', \alpha, \ell') - \pi(\boldsymbol{\mu}, \alpha, \ell), \text{ is increasing in } \alpha, \text{ for } \ell' \geq \ell, \ \boldsymbol{\mu}' \preccurlyeq \boldsymbol{\mu}$$

(b) An increase ℓ or fall in $\boldsymbol{\mu}$ worsens the technology draw of weaker firms more. For g increasing:

$$\int g(\tilde{\alpha}) P(d\tilde{\alpha} \mid \boldsymbol{\mu}', \alpha, i, \ell') - \int g(\tilde{\alpha}) P(d\tilde{\alpha} \mid \boldsymbol{\mu}, \alpha, i, \ell), \text{ is increasing in } \alpha, \text{ for } \ell' \geq \ell, \ \boldsymbol{\mu}' \preccurlyeq \boldsymbol{\mu}$$

(S-ii) Increasing patent life, ℓ, raises the marginal productivity of investment. For g increasing:

$$\int_{\tilde{\alpha}} g(\tilde{\alpha}) \Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}', \alpha, i, \ell') \geq \int_{\tilde{\alpha}} g(\tilde{\alpha}) \Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}, \alpha, i, \ell), \quad \text{for } \ell' \geq \ell$$

Figure (1) illustrates assumptions (S-i)(a) and (S-ii). Considering (S-i)(a) and taking $\boldsymbol{\mu}' = \boldsymbol{\mu}$, the marginal impact on profit of an increase in ℓ is less for a firm with better technology (Since α is not a real number, "α" denotes an axis of ordered α's).[17] Similarly, if the ambient aggregate distributions are worse, the effect

[16] Weaker in terms of α. Note that $\pi(\boldsymbol{\mu}', \alpha, \ell') - \pi(\boldsymbol{\mu}, \alpha, \ell)$ is negative, so that larger values correspond to smaller profit reduction.

[17] For a firm, the impact on profit of increasing ℓ is to reduce access to patented technology, so the firm will be worse off. (The effect is always negative, but more severe for a weaker firm.)

is greater on weaker firms. Assumption (S-i)(b) likewise assumes that better firms are more advantaged in these terms in the technology draw. So, for example, increasing patent length has a more detrimental effect on lower quality firm's in terms of success in drawing a new technology: limited access to technology or poorer aggregate technology impacts weaker firms more negatively. Assumption (S-ii) asserts that the marginal productivity of investment for all firms is increased when access to patented technology is reduced: there is more value to investing when publicly available technology is reduced.

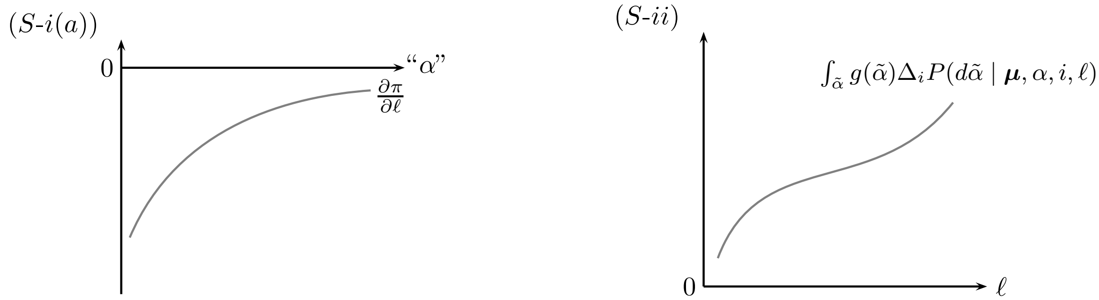


Figure 1: Patented Knowledge Substitutes for Technology and Investment

Together, these conditions imply that (when patent length is increased) there is greater pressure on weak firms to improve in terms of profitability; and there is greater reward to investment after improvement. With these assumptions:[18]

Theorem 4.1: *Suppose that profit increases with technology improvement and assumptions (S-i) and (S-ii) are satisfied. Then lengthening patent life improves the aggregate distribution of technologies in successive periods.*

The mechanism by which investment increase is through pressure on weaker firms to raise investment because the impact of reduced IP access is more detrimental to weak firms than strong ones, in terms of profitability or success in innovation. Weaker firms compensate with extra investment pressuring better firms to also increase investment. In the next section incentives work in the opposite direction resulting in a reduction of investment.

Remark 4.1: If investment increases, then the 'quality' of the future aggregate distribution increases. Whether this reinforces the increased reward to investment (or mitigates this effect), depends on how better distributions translate into profit for a firm. As an assumption (monotonicity of profit in the aggregate distribution), this is not innocuous since, as mentioned, there are conflicting effects: while ambient technological improvements raise a firm's efficiency (lower cost in the example), improve the innovation success and may raise demand — thus tending to raise profit, they also strengthen competition between firms and this works in the opposite direction. However, as equation (2) in the example shows, aggregate technological improvement has a direct effect on immediate profit and positive effect on expected future profit. It

[18] Proofs are in the appendix.

is the combination of these effects that determine monotonicity of the present value of the profit flow (see equation (4) of section (2.5)), and it is this property which is actually required for the results and for which monotonicity of the profit function is sufficient, but not necessary. ■

4.2 Investment and Patented Knowledge as Complements.

The effect of lengthening patent life is to reduce the publicly available technology, and when the impact of this is greatest on high technology firms, good technology is complemented by the patented discovery. Furthermore, if increasing patent length reduces the marginal product of investment, then that information is also a complement to investment.

So, in contrast to the previous assumptions (S-i) and (S-ii), suppose instead that better firms are more dependent on patented information to generate current profit and support innovation, so that such information is a complement to the quality of a firms' technology. Suppose also that increasing patent length removes from use information which raises the marginal product of investment (such information is complementary to investment). These conditions are formalized next.

(C-i) Increasing patent life, ℓ, or worsening technology, $\boldsymbol{\mu}_t$, has a greater impact on better firms.

- (a) An increase in ℓ, or fall in $\boldsymbol{\mu}$ lowers profits of better firms more than weaker firms:
  $\pi(\boldsymbol{\mu}', \alpha, \ell') - \pi(\boldsymbol{\mu}, \alpha, \ell)$, is decreasing in α, for $\ell' \geq \ell$, $\boldsymbol{\mu}' \preccurlyeq \boldsymbol{\mu}$
- (b) An increase in ℓ, or fall in $\boldsymbol{\mu}$ worsens the technology draw of better firms more. For g increasing:
  $\int g(\tilde{\alpha})P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t', \alpha, i, \ell') - \int g(\tilde{\alpha})P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$, is decreasing in α, for $\ell' \geq \ell$, $\boldsymbol{\mu}' \preccurlyeq \boldsymbol{\mu}$

(C-ii) Increasing patent life lowers the marginal productivity of investment. For g increasing:

$$\int_{\tilde{\alpha}} g(\tilde{\alpha})\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}', \alpha, i, \ell') \leq \int_{\tilde{\alpha}} g(\tilde{\alpha})\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}, \alpha, i, \ell), \quad \text{for } \ell' \geq \ell$$

These assumptions are depicted in figure (2).

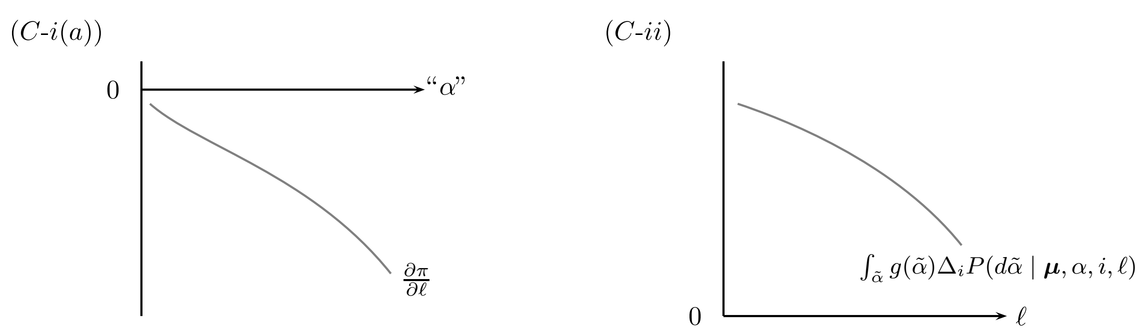


Figure 2: Patented Knowledge Complementary to Technology and Investment

Under these circumstances, lengthening patent life reduces investment. Assumption (C-i)(a) asserts that better technology firms are more negatively impacted by an extension in patent length, and would be more negatively impacted by a worsening of the aggregate technology. Assumption (C-i)(b) expresses a similar (greater negative impact on better firms), but in terms of the impact on innovation of variations in patent length or aggregate distribution quality. Finally, (C-ii) says that increasing the patent length or lowering the quality of the aggregate distribution raises the marginal productivity of investment. Then,

Theorem 4.2: *Suppose that profit increases with technology improvement and assumptions (C-i) and (C-ii) are satisfied. Then lengthening patent life worsens the aggregate distributions in successive periods.*

REMARK 4.2: Apart from these two cases, there are many other possibilities. For example, if conditions $(S-i)$ and $(C-ii)$ are both satisfied, then lengthening patent life puts pressure on weak firms to raise investment in innovation, but the benefit from investment in innovation is reduced, providing conflicting incentives. ∎

5 Equilibrium Investment and Welfare

In any equilibrium, the level of investment is inefficient because there are positive externalities from investment. This is easy to see considering equation (4). There, the individual firm maximizes the present value of profit less investment costs, taking as exogenous aggregate behavior of all firms. However, aggregate behavior enters the value function and while individual firms don't affect this distribution, their combined investment does, an effect not internalized at the level of the individual firm. Because the positive externality from investment does not appear in the individual firm's investment decision, equilibrium is inefficient. These observations are summarized in the following theorem.

Theorem 5.1: *In equilibrium, given ℓ, investment is below the efficient level, since the positive externalities of investment through improved technology distributions are not internalized.*

Thus, aggregate firm welfare (present value of profit net of investment) may be raised by a small increase in investment by each firm. Because such a policy improves the aggregate distribution of technologies over time, provided this raises consumer welfare, the overall effect is unambiguously positive. Note that the social welfare optimization problem must respect the same intellectual property rights as in the individual firm problem. However, this optimization problem does address the externality issues — given that access to technology is restricted according to the patent length, ℓ.

The previous discussion focuses on investment externalities and incentives, given patent length. A second perspective on efficiency arises from the prospect of varying ℓ and the resulting impact on welfare as equilibrium varies. In the case where investment is set to maximize welfare for a given value of ℓ (the socially optimally level rather than the market equilibrium investment level), the optimal value of ℓ is 0. This is discussed next.

Assume an environment where in each period there is a measure of consumer and producer welfare (total surplus, TS) which depends on the distribution of characteristics and the patent policy (length). Let total welfare generated at time t be denoted $TS(\boldsymbol{\mu}_t, \ell)$ and assume that:

$$TS(\boldsymbol{\mu}_t', \ell') - TS(\boldsymbol{\mu}_t, \ell) \geq 0, \quad \boldsymbol{\mu}_t' \succcurlyeq \boldsymbol{\mu}_t,\ \ell' \leq \ell \tag{7}$$

so that improving technologies or reducing patent protection on existing technologies raises current welfare. With this the social welfare maximizing problem may be defined:

$$V(\boldsymbol{\mu}, \ell) \quad = \quad \max_{\tau \in C(\mu)} \{TS(\boldsymbol{\mu}, \ell) - \int r(i) d\tau + \delta V(\boldsymbol{\mu}', \ell)\} \tag{8}$$

where $\boldsymbol{\mu}' = (\boldsymbol{\mu}, \mu')$, with μ' determined from $\mu'(\cdot) = \int P(\cdot \mid \alpha, \boldsymbol{\mu}_t, i, \ell)\tau_t(di \times d\alpha)$. The effect of improving the aggregate distributions is to raise demand and lower supply (marginal cost).

Theorem 5.2: *Social welfare, $V(\boldsymbol{\mu}, \ell)$, is decreasing in ℓ.*

The next example describes a particular choice for total surplus.

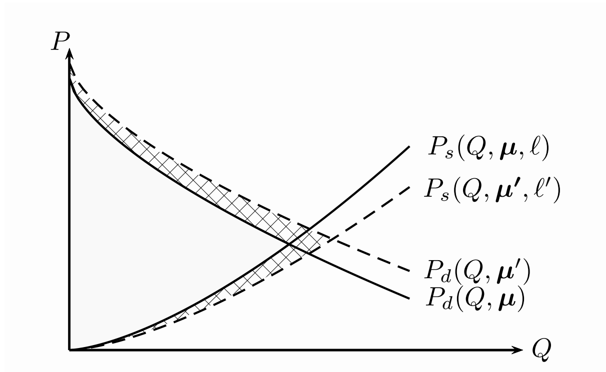
Example 5.1: The per period social welfare, TS may be illustrated in a supply and demand model. Consider a demand and supply model represented by $P_d(Q, \boldsymbol{\mu})$ and $P_s(Q, \boldsymbol{\mu}, \ell)$ respectively. Surplus is given by:

$$TS(\boldsymbol{\mu}, \ell) = \max_Q \int_0^Q [P_d(\tilde{Q}, \boldsymbol{\mu}) - P_s(\tilde{Q}, \boldsymbol{\mu}, \ell)] d\tilde{Q} \,,$$

Write $CS(\boldsymbol{\mu}_t, \ell)$ for consumer welfare at time t and $PS(\boldsymbol{\mu}_t, \ell)$ for producer welfare.

$$TS(\boldsymbol{\mu}_t, \ell) = PS(\boldsymbol{\mu}_t, \ell) + CS(\boldsymbol{\mu}_t, \ell)$$

Total surplus may be decomposed as follows. For the firm, profit is $\pi(\boldsymbol{\mu}, \alpha, \ell)$ so total firm surplus is $PS(\boldsymbol{\mu}_t, \ell) = \int \pi(\boldsymbol{\mu}, \alpha, \ell)\mu(d\alpha)$ and consumer surplus is then $T(\boldsymbol{\mu}, \ell) - PS(\boldsymbol{\mu}, \ell)$. This case is illustrated in the figure, along with the impact of reducing ℓ to ℓ' or improving $\boldsymbol{\mu}$ to $\boldsymbol{\mu}'$. Improving the distribution increases both supply and demand, reducing ℓ increases technological availability for firms and raises supply. In either case the overall effect is positive (as depicted by the crosshatched lines in the figure.)



◇

Writing $V_S(\mu, \ell)$ an $V_C(\mu, \ell)$ to denote equilibrium total surplus in the substitutes and complements cases respectively, figure (3) depicts how surplus varies in each case, and relative to the socially optimal case where investment is managed by a social planner.

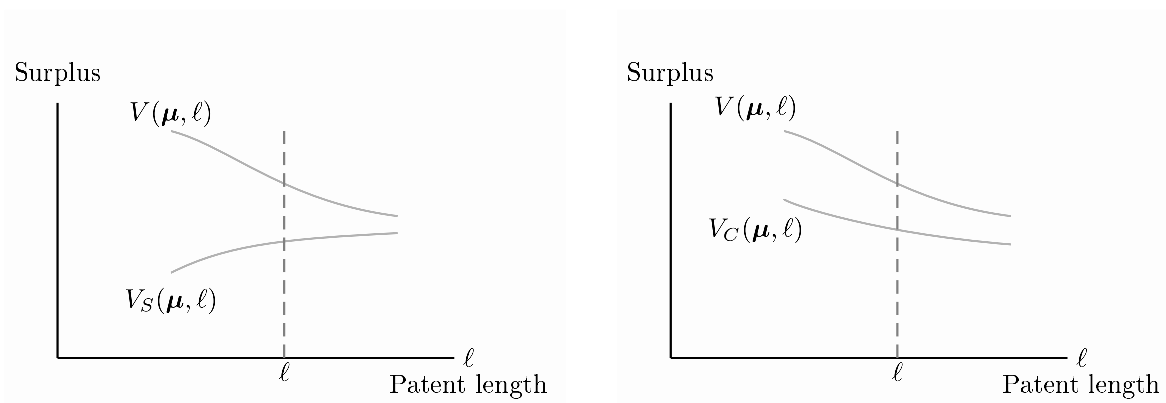


Figure 3: Welfare as patent length varies

6 Conclusion

The model developed in this paper is an idealized model with a focus on technological innovation and associated incentives among competitive firms to invest. The effect of patent policy may be to either encourage or discourage investment depending on the way firms respond to tighter or looser patent policy. And, it may not be possible to determine the impact of such policy at all — for example when the impact of a change in tightness, l, raises the marginal product of investment for some firms, and lowers it for others. In such circumstances, the impact on investment of varying patent tightness is ambiguous since it creates competing influences on different firms decisions to invest.

The framework here assumes a specific market structure (large numbers of agents) and competitive firm behavior. In an environment with a small number of firms, strategic issues become important as individual firms' decisions directly impact others. In this case, competition may be more oligopolistic than competitive in nature. The model can accommodate large firms (atoms in the distribution), but then the consideration of equilibrium behavior becomes much more complex.

Finally, regarding the focus of the model, much recent discussion has considered weaknesses in the patent granting process with the granting of trivial inventions and their use to extract rents, block competitors and generally obstruct the functioning of the system. Likewise, strategic behavior based on exploitation of the legal system (using patents to achieve hold-up or as blocking devices to impede competitors) is a significant concern. Such issues are beyond the scope of the model. Finally, the prospect of licensing is not considered here. The model is to a large extent a reduced form model developed to consider aggregate behavior and not well suited to the study of licensing and the associated strategic considerations.

Appendix I: Orderings and Technology Evolution

This appendix clarifies features of the ordering on technologies used in the paper. The subsequent appendix gives proofs of the results in the paper (See [28] for additional details).

Ordering on Technology

Λ is an ordered topological space where the relation $\succeq$ is reflexive ($\alpha \succeq \alpha$), transitive ($\alpha \succeq \alpha'$ and $\alpha' \succeq \alpha''$ imply $\alpha \succeq \alpha''$), and antisymmetric ($\alpha \succeq \alpha'$ and $\alpha' \succeq \alpha$ imply $\alpha = \alpha'$). (For example: $\Lambda = \{\alpha \mid \alpha : [a,b] \rightarrow \Re,\ \alpha \text{ measurable}\}$ where $\alpha' \succeq \alpha$ if $\alpha'(x) \geq \alpha(x), x \in [a,b]$.) If technology were characterized by a real number, the firm with the largest α would be the best firm, unequivocally, eliminating the possibility for different firms to have area specific strengths.

Ordering on Distributions over Technology

For the following review, take as given: (a1) Λ, a completely regular topological space (for example, Λ a metric space), (a2) $\mathcal{B}_\Lambda$ the Borel field on Λ, (b) $\succeq$, an order on Λ (reflexive, transitive and antisymmetric), (c) $C_b(\Lambda)$, the set of continuous bounded real-valued functions on Λ, (d) $\mathcal{M}_+(\Lambda)$, the set of non-negative measures on Λ, and (e) $\mathcal{P}(\Lambda)$ the set of probability measures on Λ. (A topological space Λ is completely regular if and only if when A is closed in Λ and $\alpha \notin A$, there is a continuous function, f, $f : \Lambda \rightarrow [0,1]$ such that $f(\alpha) = 0$ and $f(A) = 1$.)

Definition 6.1: *A real valued function $f : \Lambda \rightarrow \Re$ is called increasing if $\alpha' \succeq \alpha$ implies that $f(\alpha') \geq f(\alpha)$ (and decreasing if $\alpha' \succeq \alpha$ implies that $f(\alpha') \leq f(\alpha)$). Write $\mathcal{I}_m(\Lambda)$ for the set of increasing measurable functions on Λ.*

A set $B \subseteq \Lambda$ is called increasing if $x, y \in \Lambda$, $x \in B$ and $y \succeq x$ imply that $y \in B$.

Definition 6.2: *Given $\mu, \nu \in \mathcal{P}(\Lambda)$, define a pre-ordering (reflexive and transitive relation) on $\mathcal{P}(\Lambda)$:*

$$\mu \succeq \nu \quad \text{if and only if} \quad \int f(\alpha)\mu(d\alpha) \geq \int f(\alpha)\nu(d\alpha),\ \forall f \in \mathcal{I}_m(\Lambda)$$

The natural generalization of a result on dominance in $\Re$ is (see Torres [28]):

Theorem 6.1: *$\mu \succeq \nu$ if and only if $\mu(A) \geq \nu(A)$ for every increasing measurable set A.*

The Evolution of Technology

The distribution $\mu_{t+1}(\cdot)$ depends on $\boldsymbol{\mu}_t$, ℓ and τ_t. This may be made explicit by writing:

$$\mu_{t+1}(\cdot) = \psi(\cdot \mid \boldsymbol{\mu}_t, \tau_t, \ell) \stackrel{\text{def}}{=} \int_{i_t, \alpha_t} P(\cdot \mid \boldsymbol{\mu}_t, \alpha_t, i_t, \ell)\tau_t(di_t \times d\alpha_t), \quad \tau_t \in \mathcal{M}(I \times \mathcal{A}), \tag{9}$$

where $\tau_t \in \mathcal{C}(\mu_t)$ and ψ is the function determining the one-period ahead distribution over technologies. With $\boldsymbol{\mu}_{t+1} = (\boldsymbol{\mu}_t, \mu_{t+1})$, and $\tau_{t+1} \in \mathcal{C}(\mu_{t+1})$, $\mu_{t+2}(\cdot) = \psi(\cdot \mid \boldsymbol{\mu}_{t+1}, \tau_{t+1}, \ell)$, and so on. If we fix a sequence of strategies $\tau^t = \{\tau_{t+j}\}_{j \geq 0}$, given μ_t, we may determine the sequence of distributions:

$$
\begin{aligned}
\psi_1(\cdot \mid \mu_t, \tau^t, \ell) &= \psi(\cdot \mid \boldsymbol{\mu}_t, \tau_t, \ell) \\
\psi_2(\cdot \mid \mu_t, \tau^t, \ell) &= \psi(\cdot \mid (\psi_1(\cdot \mid \mu_t, \tau^t, \ell), \boldsymbol{\mu}_t), \tau_{t+1}, \ell) \\
\psi_3(\cdot \mid \mu_t, \tau^t, \ell) &= \psi(\cdot \mid (\psi_2(\cdot \mid \mu_t, \tau^t, \ell), \psi_1(\cdot \mid \mu_t, \tau^t, \ell), \boldsymbol{\mu}_t), \tau_{t+2}, \ell) \\
\vdots \quad &= \quad \vdots \\
\psi_j(\cdot \mid \mu_t, \tau^t, \ell) &= \psi(\cdot \mid (\psi_j, \psi_{j-1}, \ldots, \psi_1, \boldsymbol{\mu}_t), \tau_{t+j}, \ell)
\end{aligned}
\tag{10}
$$

Over time, consistency requires that $\mathrm{marg}_\Lambda \tau_{t+j} = \mu_{t+j}$. Given $\boldsymbol{\mu}_t$ and τ_t, the history is determined in $t+j$ as $(\psi_j, \ldots, \psi_1, \boldsymbol{\mu}_t)$. Then profit to α in period $t+j$ is $\pi((\psi_j, \ldots, \psi_1, \boldsymbol{\mu}_t), \alpha, \ell)$ and the conditional distribution $P(\cdot \mid (\psi_j, \ldots, \psi_1, \boldsymbol{\mu}_t), \alpha, i, \ell)$. When $j > \ell$, variations in ℓ affect both the innovations in the public domain (from periods prior to $t+j-\ell$ and the current distribution $t+j$.) Since $(\psi_j, \ldots, \psi_1)$ is determined by $\boldsymbol{\mu}_t$ and τ^t, let $\pi_j(\boldsymbol{\mu}_t, \tau^t, \alpha, \ell) \stackrel{\text{def}}{=} \pi((\psi_j, \ldots, \psi_1, \boldsymbol{\mu}_t), \alpha, \ell)$ and $P_j(\cdot \mid \boldsymbol{\mu}_t, \tau^t, \alpha, i, \ell) \stackrel{\text{def}}{=} P(\cdot \mid (\psi_j, \ldots, \psi_1, \boldsymbol{\mu}_t), \alpha, i, \ell)$.

In this formulation, a variation in ℓ at time t impacts both the length of time for which patented discovery stays out of the public domain, but also impacts the aggregate distributions from period t onward, through the updating rule ψ_j, with the aggregate distributions, $\{\tau_{t+j}\}$ fixed.

Appendix II: Proofs

The following discussion presents a few results (lemma 1-lemma 4) which are used in the proofs of theorems (4.1) and (4.2) in the paper. The first lemma, lemma 1, confirms that the monotonicity of π in $\boldsymbol{\mu}$ carries over to the value function.

Lemma 1: Suppose that profit increases with technological improvement. Then $v(\boldsymbol{\mu}_t, \tau^t, \alpha, \ell)$ is increasing in $\boldsymbol{\mu}_t$: if $\bar{\boldsymbol{\mu}}_t = (\ldots, \bar{\mu}_{t-1}, \bar{\mu}_t)$ dominates $\hat{\boldsymbol{\mu}}_t = (\ldots, \hat{\mu}_{t-1}, \hat{\mu}_t)$ component-wise, then $v(\bar{\boldsymbol{\mu}}_t, \tau^t, \alpha, \ell) \geq v(\hat{\boldsymbol{\mu}}_t, \tau^t, \alpha, \ell)$ for all α. ◇

Proof: Since $\pi(\bar{\boldsymbol{\mu}}_t, \alpha, \ell) \geq \pi(\hat{\boldsymbol{\mu}}_t, \alpha, \ell)$ and $P(d\tilde{\alpha} \mid \bar{\boldsymbol{\mu}}_t, \alpha, i, \ell) \succcurlyeq P(d\tilde{\alpha} \mid \hat{\boldsymbol{\mu}}_t, \alpha, i, \ell)$ for each t, the result follows directly. ■

When patent length, ℓ, varies then apart from the direct effect on the payoff function and the transition kernel, there is the indirect effect of varying future distributions which impact the profit in those subsequent periods. Considering profit k periods on after the increasing of patent length in period t, the following calculations show that profit increases with ℓ, given any fixed sequence of strategies $\boldsymbol{\tau}^t = \{\tau_{t+j}\}_{j \geq 0}$. The next result uses the notation from Appendix I.

LEMMA 2: Suppose that (S-i)(a) holds. Then for all k,

$$\frac{\partial \pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)}{\partial \ell}$$

is increasing in α. $\diamond$

PROOF: Recall (S-i)(a): $\pi(\boldsymbol{\mu}',\alpha,\ell')-\pi(\boldsymbol{\mu},\alpha,\ell)$, is increasing in α, for $\ell'\geq\ell$, $\boldsymbol{\mu}'\preccurlyeq\boldsymbol{\mu}$. Let $\psi_1,\ldots,\psi_k$ be the distribution sequence determined by ℓ and $\psi'_1,\ldots,\psi'_k$ the sequence determined by $\ell'\geq\ell$. Consider

$$\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell')-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)$$

Since $(\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t)\preccurlyeq(\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t)$ (because the system with the longer patent life induces poorer distributions, period by period), and $\ell'\geq\ell$,

$$\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell')-\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell)$$

is increasing in α, from (S-i)(a). Therefore, with $\alpha^*\succeq\alpha$

$$\begin{aligned}\frac{1}{\ell'-\ell}&[\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha^*,\ell')-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha^*,\ell)]\\&\geq\frac{1}{\ell'-\ell}[\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell')-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)]\end{aligned}$$

and in the limit, for $\alpha^*\succeq\alpha$:

$$\frac{\partial \pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha^*,\ell)}{\partial \ell}\geq\frac{\partial \pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)}{\partial \ell}$$

$\blacksquare$

REMARK 6.1: Note that the variation in ℓ (to ℓ') has a direct effect on the subsequent distributions that appears in the calculation (ψ_{t+j} moves to ψ'_{t+j}). The calculations may be clarified by noting that

$$\begin{aligned}\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell')-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)\quad=&\\ \pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell')-\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell)&+\\ \pi((\psi'_k,\ldots,\psi'_1,\mu_t),\alpha,\ell)-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)&\end{aligned}$$

where both terms on the right are increasing in α by (S-i)(a):

$$\frac{1}{\ell'-\ell}[\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha^*,\ell')-\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha^*,\ell)]$$
$$\geq \quad \frac{1}{\ell'-\ell}[\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell')-\pi((\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t),\alpha,\ell)]$$

and

$$\frac{1}{\ell'-\ell}\pi((\psi'_k,\ldots,\psi'_1,\mu_t),\alpha,\ell)-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)$$
$$\geq \quad \frac{1}{\ell'-\ell}\pi((\psi'_k,\ldots,\psi'_1,\mu_t),\alpha,\ell)-\pi((\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t),\alpha,\ell)$$

Considering the first term, this give the variation in π for a given ℓ while the second gives the variation resulting from the impact on the aggregate distribution of technologies over time. ■

Again, given any fixed sequence of strategies $\boldsymbol{\tau}^t=\{\tau_{t+j}\}_{j\geq 0}$:

LEMMA 3: If (S-i)(b) holds, then for all k,

$$\frac{\partial}{\partial\ell}\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\psi^k,\alpha,i,\ell)$$

is increasing in α (where $\psi^k=(\psi_k,\psi_{k-2},\ldots,\psi_1)$). ◇

PROOF: With $\ell'\geq\ell$ and $\psi'^k=(\psi'_k,\psi'_{k-1},\ldots,\psi'_1)$, as in lemma (2), $(\psi'_k,\ldots,\psi'_1,\boldsymbol{\mu}_t)\preccurlyeq(\boldsymbol{\mu}_t,\psi_k,\ldots,\psi_1,\boldsymbol{\mu}_t)$ (because the system with the longer patent life induces poorer distributions, period by period), for $\alpha^*\succeq\alpha$,

$$\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi'^k,\boldsymbol{\mu}_t,\alpha^*,i,\ell')-\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha^*,i,\ell)=$$
$$\geq \quad \int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi'^k,\boldsymbol{\mu}_t,\alpha,i,\ell')-\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha,i,\ell)$$

Dividing by $\ell'-\ell$, with $\ell'>\ell$ and passing to the limit,

$$\frac{\partial}{\partial\ell}\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha^*,i,\ell)\geq\frac{\partial}{\partial\ell}\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha,i,\ell)$$ ■

REMARK 6.2: The variation from (ℓ,ψ^k) to (ℓ',ψ'^k) may be decomposed:

$$\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi'^k,\boldsymbol{\mu}_t,\alpha,i,\ell')-\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha,i,\ell)=$$
$$[\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi'^k,\boldsymbol{\mu}_t,\alpha,i,\ell')-\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi'^k,\boldsymbol{\mu}_t,\alpha,i,\ell)]$$
$$+[\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi'^k,\boldsymbol{\mu}_t,\alpha,i,\ell)-\int g(\tilde{\alpha})P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha,i,\ell)]$$

Form this perspective, the limit is the sum of two terms (the variation due to a change in ℓ given ψ, and the variation resulting from the distribution change at given ℓ.) ∎

LEMMA 4: Suppose that (S-i) holds. Then, for each t, $v_\ell(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)$ is increasing in α:

$$v_\ell(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha', \ell) \equiv \frac{\partial v(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha', \ell)}{\partial \ell} \geq \frac{\partial v(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)}{\partial \ell}, \ \alpha' \geq \alpha. \qquad \diamond$$

PROOF: Consider a two period problem where:

$$v^2(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell) \stackrel{\text{def}}{=} \max_{i_t}\{\pi(\boldsymbol{\mu}_t, \alpha, \ell) - r(i_t) + \delta \int \pi(\psi_1, \boldsymbol{\mu}_t, \tilde{\alpha}, \ell) P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell)\}$$

Differentiating with respect to ℓ (and using the optimality condition for i_t (see equation (5))):

$$\begin{aligned}\frac{\partial v^2(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)}{\partial \ell} &= \frac{\partial \pi(\boldsymbol{\mu}_t, \alpha, \ell)}{\partial \ell} + \delta \int \frac{\partial \pi(\psi_1, \boldsymbol{\mu}_t, \tilde{\alpha}, \ell)}{\partial \ell} P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell) + \\ &\qquad \delta \int \pi(\psi_1, \boldsymbol{\mu}_t, \tilde{\alpha}, \ell) \Delta_\ell P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell)\end{aligned}$$

Therefore,

$$\begin{aligned}&\frac{\partial v^2(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha', \ell)}{\partial \ell} - \frac{\partial v^2(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)}{\partial \ell} \\ &\quad = [\frac{\partial \pi(\boldsymbol{\mu}_t, \alpha', \ell)}{\partial \ell} - \frac{\partial \pi(\boldsymbol{\mu}_t, \alpha, \ell)}{\partial \ell}] + \\ &\qquad \delta \int \frac{\partial \pi(\psi_1, \boldsymbol{\mu}_t, \tilde{\alpha}, \ell)}{\partial \ell}[P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha', i_t, \ell) - P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell)] + \\ &\qquad \delta \int \pi(\psi_1, \boldsymbol{\mu}_t, \tilde{\alpha}, \ell)[\Delta_\ell P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha', i_t, \ell) - \Delta_\ell P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell)]\end{aligned}$$

Considering the three terms on the right, the first term is positive since since π_ℓ is increasing in α, from lemma (2). The second term is positive since $P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha', i_t, \ell) \succcurlyeq P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell)$. Finally, from (S-$i$)(b) the third term is positive — since $\pi(\boldsymbol{\mu}_{t+1}, \alpha, \ell)$ is increasing in α and $\Delta_\ell P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha', i_t, \ell)$ first order stochastically dominates $\Delta_\ell P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t, \ell)$], using lemma (3). Hence, $\frac{\partial v^2(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)}{\partial \ell}$ is increasing in α.

For notational convenience, observe that the k period valuation function, $v^k(\boldsymbol{\mu}_t, \tau^t, \alpha, \ell)$ may be expressed in terms of $\psi_k, \psi_{k-1}, \ldots, \psi_1$, the aggregate distribution in each of the k periods from t, determined by τ^t: $v^k(\psi_k, \ldots, \psi_1, \boldsymbol{\mu}_t, \alpha, \ell)$. Write $v^k_\ell(\psi_k, \ldots, \psi_1, \boldsymbol{\mu}_t, \alpha, \ell)$ for the partial derivative with respect to ℓ. Also, for convenience, let $(\psi_k, \ldots, \psi_1) = \psi^k$.

Suppose that $v^k_\ell(\psi^k, \boldsymbol{\mu}_t, \alpha, \ell)$ is increasing in α, then so is $v^{k+1}_\ell(\psi^{k+1}, \boldsymbol{\mu}_t, \alpha, \ell)$.

$$\frac{\partial v^{k+1}(\psi^{k+1},\boldsymbol{\mu}_t,\alpha,\ell)}{\partial\ell} = \frac{\partial\pi(\boldsymbol{\mu}_t,\alpha,\ell)}{\partial\ell}+\delta\int\frac{\partial v^{k}(\psi^{k+1},\boldsymbol{\mu}_t,\tilde{\alpha},\ell)}{\partial\ell}P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha,i_t,\ell)+ \delta\int v^k(\psi^{k+1},\boldsymbol{\mu}_t,\tilde{\alpha},\ell)\Delta_\ell P(d\tilde{\alpha}\mid\psi^k,\boldsymbol{\mu}_t,\alpha,i_t,\ell)$$

Thus, for any k, and $(\boldsymbol{\mu}_t,\tau^t,\ell)$, $v_\ell^k(\boldsymbol{\mu}_t,\tau^t,\alpha,\ell)$ is increasing in α.

Consider, for $\alpha'\succeq\alpha$, $\ell'\geq\ell$

$$v^k(\boldsymbol{\mu}_t,\tau^t,\alpha',\ell')-v^k(\boldsymbol{\mu}_t,\tau^t,\alpha',\ell)\geq v^k(\boldsymbol{\mu}_t,\tau^t,\alpha,\ell')-v^k(\boldsymbol{\mu}_t,\tau^t,\alpha,\ell)$$

Passing to the limit with $k\to\infty$,

$$v(\boldsymbol{\mu}_t,\tau^t,\alpha',\ell')-v(\boldsymbol{\mu}_t,\tau^t,\alpha',\ell)\geq v(\boldsymbol{\mu}_t,\tau^t,\alpha,\ell')-v(\boldsymbol{\mu}_t,\tau^t,\alpha,\ell)$$

Dividing by $\ell'-\ell$ and taking limits ($\ell'\to\ell$) gives:

$$\frac{\partial v(\boldsymbol{\mu}_t,\boldsymbol{\tau}^t,\alpha,\ell)}{\partial\ell}\quad\text{is increasing in}\quad\alpha.$$ ■

PROOF OF THEOREM 4.1:

Considering equation (5):

$$-r'(i)+\delta\int_{\tilde{\alpha}}v(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_i P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell)\equiv 0$$

the marginal impact of a variation in ℓ on firm α's investment level, ignoring the impact on the equilibrium distribution $\boldsymbol{\tau}^t$, is obtained by differentiating equation (5) with respect to ℓ.

$$-r''(i)\frac{di_t}{d\ell}+\delta\int_{\tilde{\alpha}}v(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_{ii}P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell)\frac{di_t}{d\ell} +\delta\int_{\tilde{\alpha}}v_\ell(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_i P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell) +\delta\int_{\tilde{\alpha}}v(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_{li}P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell)\equiv 0 \tag{11}$$

Therefore:

$$\frac{di_t}{d\ell}=\frac{\delta\int_{\tilde{\alpha}}v_\ell(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_i P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell)+\delta\int_{\tilde{\alpha}}v(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_{li}P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell)}{r''(i_t)-\delta\int_{\tilde{\alpha}}v(\boldsymbol{\mu}_{t+1},\boldsymbol{\tau}^{t+1},\alpha,\ell)\Delta_{ii}P(d\tilde{\alpha}\mid\boldsymbol{\mu}_t,\alpha,i,\ell)} \tag{12}$$

From the second order condition, the denominator is positive, so the sign of $\frac{di_t}{d\ell}$ is the same as that of the numerator. Since v is increasing in α and $\Delta_i P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell') \succcurlyeq \Delta_i P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$ for $\ell' \geq \ell$, from assumption (S-ii), $\int_{\tilde{\alpha}} v(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)\Delta_{li} P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) > 0$. From lemma 4, $v_\ell(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)$ is increasing in α so that $\int_{\tilde{\alpha}} v_\ell(\boldsymbol{\mu}(t), \tilde{\alpha}, \ell, t)\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) > 0$, for each t. Consequently, $\frac{di_t}{d\ell} > 0$.

This expression gives the variation in firm α's investment level that would result if ℓ were increased, and the actions of the population, $(\tau_t, \boldsymbol{\tau}^{t+1})$, held constant. However, increased investment resulting from an increase in ℓ occurs for all agents, alters future aggregate distributions, and alters optimal decisions in future periods.

These calculations ignore the impact of changes in investment behavior on the aggregate distribution. Recall the first order condition for i at technology α:

$$r'(i) = \delta \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \alpha, \ell)\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell).$$

With the increase in ℓ, from equation (12), the variation in i is upward. So, for any α, with $\ell' > \ell$, the expression:

$$r'(i') = \delta \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell')\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i', \ell').$$

has solution $i' > i$. However, τ^{t+1} cannot now be an equilibrium strategy as higher investment by each firm will impact μ_{t+j}, $j \geq 1$, raising the quality of the aggregate distribution in the next and subsequent periods. From lemma (1) ($v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)$ increasing in the aggregate distribution) and the fact that $\Delta_i P(B \mid \boldsymbol{\mu}_t, \alpha, i', \ell') \geq 0$, for all events B, $\int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell')\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i', \ell')$ increases so that the best response from α is to raise i further with consequent (further) impact on the aggregate distribution. Assuming $r'(x)$ is sufficiently large for large values of x, the iterative process will eventually converge to equilibrium. Consequently the impact of increasing ℓ is to raise the aggregate distribution quality in subsequent periods and hence the present value of surplus (welfare).

■

<u>PROOF OF THEOREM 4.2:</u>

The direct impact on investment is again given by equation (12). Assumption $(II - ii)$ gives $\Delta_i P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell') \preccurlyeq \Delta_i P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$ for $\ell' \geq \ell$ and implies that $\int_{\tilde{\alpha}} v(\boldsymbol{\mu}(t), \tilde{\alpha}, \ell)\Delta_{li} P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) < 0$. Considering the first term in the numerator of equation (12), from $(II - i)$, $\frac{\partial \pi(\boldsymbol{\mu}_t, \alpha, \ell)}{\partial \ell}$ is decreasing in α. Also, from $(II - i)$, $\Delta_\ell P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell') \preccurlyeq \Delta_\ell P(\cdot \mid \boldsymbol{\mu}_t, \alpha, i, \ell)$. With these observations, repeating the steps in lemma (4) implies that for each t, $v_\ell(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)$ is decreasing in α. Therefore, $\int_{\tilde{\alpha}} v_\ell(\boldsymbol{\mu}(t), \tilde{\alpha}, \ell)\Delta_i P(\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell) < 0$. Consequently, $\frac{di_t}{d\ell} < 0$.

As before, these calculations ignore the impact of changes in investment behavior on the aggregate distribution. So, reconsider the first order condition:

$$r'(i) = \delta \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \alpha, \ell)\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i, \ell).$$

After raising ℓ, the variation in i_t, holding τ^{t+1} fixed, is downward. With $\ell' > \ell$, $i' < i$ satisfies:

$$r'(i') = \delta \int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell')\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i', \ell').$$

Again, this expression ignores the fact that lower investment will impact the future distributions: τ^{t+1} cannot now be an equilibrium strategy as lower investment by each firm will impact μ_{t+j}, $j \geq 1$, reducing the quality of the aggregate distribution in the next and subsequent periods.

For the same reasons as in theorem 4.1, $\int_{\tilde{\alpha}} v(\boldsymbol{\mu}_{t+1}, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell')\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i', \ell')$ decreases so that the best response from α is to reduce i further with consequent (further) impact on the aggregate distribution. Assuming $r'(x) \to 0$ as $x \to 0$, the iterative process will eventually converge to equilibrium. Consequently the impact of increasing ℓ is to worsen the aggregate distribution quality in subsequent periods and hence the present value of surplus (welfare).

Proof of theorem 5.1:

Let $i_t(\alpha)$ be the equilibrium investment strategy of α. Consider the aggregate expected payoff:

$$\begin{aligned}\int v(\boldsymbol{\mu}_t, \boldsymbol{\tau}^t, \alpha, \ell)\mu_t(d\alpha) = \int \pi(\boldsymbol{\mu}_t, \alpha, \ell)\mu_t(d\alpha) - \int r(i_t(\alpha))\mu_t(d\alpha) \\ +\delta \int v(\psi_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha), \ell)\mu_t(d\alpha)\}.\end{aligned}$$

(with ψ_1 given in equation (10)). Perturbing $i_t(\alpha)$ to $i_t(\alpha) + \epsilon h(\alpha)$ where $0 < h(\alpha) < c$ for some positive number c and ϵ small. Consider the variation in the aggregate expected payoff:

$$\begin{aligned}\Big\{ - \int r(i_t(\alpha) + \epsilon h(\alpha)) + \delta \int v(\tilde{\psi}_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha) + \epsilon h(\alpha), \ell)\Big\}\mu_t(d\alpha) \\ -\Big\{ - \int r(i_t(\alpha)) + \delta \int v(\psi_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha), \ell)\Big\}\mu_t(d\alpha)\end{aligned}$$

where $\tilde{\psi}_1$ is the $t+1$ period distribution given the strategy $i_t(\alpha)+\delta h(\alpha)$. Since the distribution $\tilde{\psi}_1$ dominates ψ_1 due to higher investment (see equation (3)),

$$\Delta v = v(\tilde{\psi}_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \alpha, \ell) - v(\psi_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \alpha, \ell) \geq 0, \forall \alpha$$

Rearranging,

$$\Big\{-\int [r(i_t(\alpha)+\epsilon h(\alpha)) - r(i_t(\alpha))]$$
$$+\delta \int v(\psi_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)[P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha)+\epsilon h(\alpha), \ell) - P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha), \ell)]\Big\}\mu_t(d\alpha)$$
$$+\delta \int \Delta \cdot P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha)+\epsilon h(\alpha), \ell)\mu_t(d\alpha)$$

Dividing by ϵ and letting $\epsilon \to 0$ gives:

$$\Big\{-\int r'(i_t(\alpha)) + \delta \int v(\psi_1, \boldsymbol{\mu}_t, \boldsymbol{\tau}^{t+1}, \tilde{\alpha}, \ell)\Delta_i P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha), \ell)\Big\}h(\alpha)\mu_t(d\alpha)$$
$$+\delta \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int \Delta \cdot P(d\tilde{\alpha} \mid \boldsymbol{\mu}_t, \alpha, i_t(\alpha)+\delta h(\alpha), \ell)\mu_t(d\alpha)$$

At the market equilibrium, the first term is 0, but the second term is strictly positive and captures the (positive) externalities from improving the aggregate distribution.

□

<u>Proof of theorem 5.2:</u>

Recall social welfare is measured as:

$$V(\boldsymbol{\mu}, \ell) \quad = \quad \max_{\tau \in C(\mu)} \{TS(\boldsymbol{\mu}, \ell) - \int r(i)d\tau + \delta V(\boldsymbol{\mu}', \ell)\} \tag{13}$$

where $\boldsymbol{\mu}' = (\psi_1, \boldsymbol{\mu})$ and where ψ_1 depends on τ (where necessary, this may be made explicit by writing $\psi_1(\tau)$). The following discussion shows that

$$V(\boldsymbol{\mu}', \ell') - V(\boldsymbol{\mu}, \ell) \geq 0, \quad \boldsymbol{\mu}' \succcurlyeq \boldsymbol{\mu},\ \ell' \leq \ell \tag{14}$$

so that, in particular, V is decreasing in ℓ. Define $V_1(\boldsymbol{\mu}, \ell) = TS(\boldsymbol{\mu}, \ell)$ and V_n inductively:

$$V_n(\boldsymbol{\mu}, \ell) \quad = \quad \max_{\tau \in C(\mu)} \{TS(\boldsymbol{\mu}, \ell) - \int r(i)d\tau + \delta V_{n-1}((\psi_1, \boldsymbol{\mu}), \ell)\} \tag{15}$$

Suppose that for some $n > 1$, V_{n-1} satisfies (15), then V_n satisfies (14). $V_n(\boldsymbol{\mu}, \ell) = V_{n-1}(\boldsymbol{\mu}, \ell)$. To see this, let τ^n solve (15), and consider a variation in ℓ to $\ell' < \ell$. With τ^n fixed, apart from direct impact, ψ_1 shifts to $\psi_1' \succcurlyeq \psi_1$, so that $V_{n-1}((\psi_1', \boldsymbol{\mu}), \ell') \geq V_{n-1}((\psi_1 \boldsymbol{\mu}), \ell)$ and similarly, $TS(\boldsymbol{\mu}, \ell') \geq TS(\boldsymbol{\mu}, \ell)$, so that

$$\begin{aligned} V_n(\boldsymbol{\mu}, \ell') = \max_{\tau \in C(\mu)} &\{TS(\boldsymbol{\mu}, \ell') - \int r(i) d\tau + \delta V_{n-1}(\psi_1(\tau), \boldsymbol{\mu}, \ell')\} \\ &\geq \; TS(\boldsymbol{\mu}, \ell') - \int r(i) d\tau + \delta V_{n-1}(\psi_1(\tau^n), \boldsymbol{\mu}, \ell')\} \\ &\qquad \geq \; TS(\boldsymbol{\mu}, \ell) - \int r(i) d\tau + \delta V_{n-1}(\psi_1(\tau^n), \boldsymbol{\mu}, \ell)\} = V_n(\boldsymbol{\mu}, \ell) \end{aligned}$$

Recalling (7), $V_1(\boldsymbol{\mu}', \ell') \geq V_1(\boldsymbol{\mu}, \ell)$ for $\boldsymbol{\mu}_t' \succcurlyeq \boldsymbol{\mu}_t$ and $\ell' \leq \ell$, this implies that $V_n(\boldsymbol{\mu}, \ell') \geq V_n(\boldsymbol{\mu}, \ell)$ by induction. Similar computations give a comparable result for $\boldsymbol{\mu}$ variation: $V_n(\boldsymbol{\mu}', \ell) \geq V_n(\boldsymbol{\mu}, \ell)$, $\boldsymbol{\mu}' \succcurlyeq \boldsymbol{\mu}$. Thus, if V_{n-1} satisfies (14), so does V_n. Thus, V_n satisfies (14) for each n, and taking the limit as $n \to \infty$, gives the property for V. In particular, this implies that V is decreasing in ℓ.

□

References

[1] F. Beck and C. Stumpe. "Two Devices for Operator Interaction in the Central Control of the New Cern Accelerator". In: *CERN discussion paper 73-76* (1973).

[2] J. Bergin and D. Bernhardt. "Anonymous sequential games: existence and characterization of equilibria". In: *Economic Theory* 5 (1995), pp. 461–489.

[3] J. Bessen and E. Maskin. "Sequential Innovation, patents, and imitation". In: *The RAND Journal of Economics* 40 (2009), pp. 611–635.

[4] J. Bessen and E. Maskin. "Sequential innovation, patents, and imitation". In: *The Rand Journal of Economics* 40, No. 4 (2009), pp. 611–635.

[5] V. Denicolo. "Patent Races and Optimal Patent Breadth and Length". In: *Journal of Industrial Economics* 44, No. 3 (1996), pp. 249–265.

[6] J. L. Furman and S. Stern. "Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research". In: *American Economic Review* 101(5) (2011), pp. 1933–63.

[7] N. Gallini. "Patent Length and Breadth with Costly Imitation". In: *The Rand Journal of Economics* 44 (1992), pp. 52–63.

[8] R. Gilbert and C. Shapiro. "Optimal Patent Length and Breadth". In: *The Rand Journal of Economics* 21, No. 1 (1990), pp. 106–112.

[9] J. Grandell. *Mixed Poisson Processes.* Monographs on Statistics & Applied Probability 77, Chapman and Hall, 1997.

[10] J. Green and S. Scotchmer. "On the Division of Profit in Sequential Innovation". In: *The Rand Journal of Economics* 26 (1995), pp. 20–33.

[11] C. Greenhalgh and M. Rogers. *Innovation, Intellectual Property, and Economic Growth.* Princeton University Press, 2010.

[12] E. Gutierrez-Pena and L. E. Nieto-Barajas. "Bayesian Nonparametric Inference for Mixed Poisson Processes". In: *Bayesian statistics 7: Proceedings of the seventh Valencia international meeting.* Ed. by J. M. Bernardo et al. Oxford University Press Oxford, 2003.

[13] B. Jovanovic and R.W. Rosenthal. "Anonymous sequential games". In: *Journal of Mathematical Economics* (1988), pp. 77–87.

[14] Y. Kotowitz and P. Schure. "The Optimal Patent Length". In: *mimeo, University of Victoria, Canada* (2006).

[15] M. Lefebvre. *Applied Stochastic Processes*. Springer, 2007.

[16] W. D. Nordhaus. "The Optimal Life of a Patent". In: *Cowles Foundation Discussion Paper No. 241* (1967).

[17] Ted O'Donohue, Susan Scotchmer, and Jacques-Francois Thisse. "Patent Breadth, Patent Life, and the Pace of Technological Progress". In: *Journal of Economics and Management* 7, No. 1 (1998), pp. 1–32.

[18] M. K. Poetz and R. Prügl. "Crossing Domain-Specific Boundaries in Search of Innovation: Exploring the Potential of Pyramiding". In: *Journal of Product Innovation Management* 27, No. 6 (2010), pp. 897–914.

[19] M. Poetz, N Franke, and M. Schreier. "Sometimes the Best Ideas Come from Outside Your Industry". In: *Harvard Business Review* (November 21, 2014).

[20] H. Privault. *Stochastic Finance, An Introduction with Market Examples*. Chapman and Hall/CRC Financial Mathematics Series, 2013.

[21] S Scotchmer. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law". In: *Journal of Economic Perspectives* 5(1) (1991), pp. 29–41.

[22] S. Scotchmer. "Protecting Early Innovators: Should Second-Generation Products be Patentable". In: *The Rand Journal of Economics* 27 (1996), pp. 322–331.

[23] S. Scotchmer. *Innovation and Incentives*. MIT press, 2004.

[24] Robert. Stanisławskia and R. Lisowska. "The Relations between Innovation Openness (Open Innovation) and the Innovation Potential of SME's". In: *Procedia Economics and Finance* 23 (2015), pp. 1521–1526.

[25] A. Sweeny. *BlackBerry Planet: The Story of Research in Motion and the Little Device that Took the World by Storm*. Wiley, 2009.

[26] T. Takalo. "On the Optimality of patent Policy". In: *Finnish Economic Papers* 14(1) (2001), pp. 33–40.

[27] *The Measurement of Scientific and Technological Activities: GUIDELINES FOR COLLECTING AND INTERPRETING INNOVATION DATA, Oslo Manual, third edition*. OECD, 2005.

[28] Ricard Torres. "Stochastic Dominance". In: *Northwestern Discussion paper* (1990).