WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 01-13
PATENTABILITY, INDUSTRY STRUCTURE, AND INNOVATION

Robert M. Hunt
Federal Reserve Bank of Philadelphia

August 2001

FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • (215) 574-6428 • www.phil.frb.org

PATENTABILITY,
INDUSTRY STRUCTURE,
AND INNOVATION

Robert M. Hunt*
Federal Reserve Bank of Philadelphia

August 2001

* Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106. Email: bob.hunt@phil.frb.org

The views expressed here are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. The author wishes to thank George Mailath, Stephen Coate, Dennis Yao, Leonard Nakamura, Kamal Saggi, Alberto Trejos, Nikolaos Vettas, Chris Chalmers, Juuso Valimaki, and Rohit Verma for their insights on previous drafts of my work on this topic. Any remaining errors are, of course, my own.

PATENTABILITY,
INNOVATION,
AND INDUSTRY STRUCTURE

Robert M. Hunt
Federal Reserve Bank of Philadelphia

August 2001

Abstract

This paper evaluates a model of sequential innovation in which industry structure (the number of firms) is endogenous and a standard of patentability determines the proportion of all inventions that qualify for protection. In equilibrium, the number of firms actively engaged in R&D is the primary determinant of an industry's rate of progress. The patentability standard affects firms' entry decisions because it affects the expected profits associated with innovations.

There is a unique, interior standard that maximizes the rate of innovation in a given industry by maximizing the number of firms actively engaged in R&D. This "optimal" standard is more stringent for an industry that tends to innovate rapidly and less stringent for an industry that tends to innovate slowly. If a single standard is applied to heterogeneous industries, it will encourage entry, and therefore innovation, in some industries while discouraging it in others. Finally, the effect of any given standard of patentability on industry structure and innovation will vary as conditions in the industry change. These results lead to interesting insights about changes in patent law in the U.S. and abroad.

1. Introduction

A significant strand of the industrial organization literature is devoted to investigating which market structures are most conducive to innovation (Cohen and Levin 1989, Scherer 1992), and the results could be described as inconclusive. More recently, economists have investigated, in the context of cumulative innovation, the relationship between the availability of patent protection and the rate of innovation (O'Donoghue 1998, Hunt 1999a, and Bessen and Maskin 2001). The general conclusion is that an industry's rate of innovation is maximized by protecting some inventions, but not others.

This paper presents a model in which industry structure (the number of firms engaged in R&D) depends on the share of all discoveries that qualify for protection, that is by the stringency of criteria used to examine applications for a patent. In the model, the number of firms actively engaged in R&D is the primary determinant of an industry's rate of progress. This in turn depends on the fixed cost of establishing a research facility, the productivity of R&D, and the resulting profits generated in the output market. Patentability criteria affect expected profits because they determine the likelihood that a firm's invention will lead to a competitive advantage and the speed with which that advantage will be eroded.

In the model, industry structure is characterized by a single firm in the output market that is eventually replaced by a firm that develops a patentable innovation. The arrival rate of these innovations is a non-monotonic function of the standard used to determine the patentability of inventions. There is a unique, interior standard that maximizes the rate of innovation in a given industry by maximizing the number of firms that choose to engage in

R&D. This "optimal" standard is more stringent for an industry that tends to innovate rapidly and less stringent for an industry that tends to innovate more slowly.

If the goal of patent policy is to maximize the rate of innovation, a social planner would adopt a more stringent patentability standard for those industries that innovate more rapidly and a less stringent one for those industries that innovate more slowly. In practice, however, patent laws generally use a single set of criteria to evaluate inventions across all industries. Subject to this constraint, more stringent patentability standards should be associated with more participants in rapidly innovating industries and fewer participants in industries that innovate more slowly. Thus, so long as a single standard is employed, the choice patentability criteria is also a choice about industrial policy.

The results lead to two additional implications. First, so long as intellectual property rights are generally defined for all technologies, the optimal patentability standard will vary across countries depending on their mix of industries and the relative sizes of those industries. Thus, efforts towards the harmonization of patent law that involve a common set of patentability criteria will affect countries differently. Second, if the "optimal" standard depends on industry characteristics that influence the rate of innovation, this standard will vary as conditions in the industry change. That suggests optimal patentability criteria may look more like a balancing test than a statutory threshold.[1]

The remainder of the paper is organized as follows. Section 2 introduces the model and contrasts it with the previous literature. Section 3 presents the equilibrium, describes its

[1] Practitioners of patent law would likely object to this characterization of optimal standards (see the conclusion) This paper does not address the arguments underlying such objections. Rather, it illustrates the economic trade-offs implicit in the selection of a single set of criteria to be used by all industries, or even

properties, and compares it to the first best. Section 4 describes the relationship between patentability standards and the rate of innovation and derives the optimal standard as a function of an industry's propensity to innovate. Section 5 concludes. The appendix contains the proofs of all the propositions.

2. The Model

2.1 An Infinite Sequence of Stochastic Patent Races

Time is continuous and the horizon is infinite. Let $r > 0$ denote the discount rate. Discoveries occur at different points in time. It is convenient to divide time into the intervals between these discoveries and call them patent races. Because there is randomness in the process that generates discoveries, the actual duration of patent races will vary.

At any point in time there are n+1 firms in the industry, where $n \geq 0$ is determined according to a free entry constraint that reflects the expected return to innovation and the fixed cost k of setting up an R&D lab. This cost is sunk on commencement of the firm's first patent race, and new fixed investments are not required thereafter. Firms are indexed by the superscript i. At the beginning of a race, firms simultaneously choose their R&D intensity, denoted $h^i \in [0, \bar{h}]$, where $\bar{h}$ is a very large, but finite point of saturation. Firms maintain their research intensity until a discovery occurs and the current race ends. The flow cost of conducting R&D, denoted $C(h^i)$, is strictly increasing and twice continuously differentiable in R&D intensity.

---

countries.

All firms share the same R&D technology. A firm's discoveries arrive through time according to a Poisson process, where the arrival rate is determined by its R&D intensity. Thus the arrival rate of ideas for firm i in patent race q is λh_q^i, where λ is an industry-specific productivity parameter. The probability that firm i discovers an invention before date t in the patent race q is $1\text{-}e^{-\lambda h_q^i t}$. The firm faces a constant *rival hazard rate* $\lambda a_q^i \equiv \lambda \sum_{j \neq i} h_q^j$. The probability that firm i wins is $h_q^i / [h_q^i + a_q^i]$, the ratio of firm i's hazard rate to the hazard rate for the entire industry.

2.1.1 A Passive Incumbent. A firm that owns a patented invention will be called an *incumbent*. The other firms will be called *challengers*. The model contains an additional assumption about the nature of technological competition: A firm that makes a patentable discovery does not compete in the subsequent patent race. This ad hoc restriction considerably simplifies the model and subsequent analysis, but it does not affect the qualitative properties of a model of patent races where the only difference between the incumbent and other firms are the rents it earns. In models of this type, being successful in a given patent race does not convey a natural advantage over rivals in subsequent races. It can be shown that the incumbent will race less aggressively than other firms, because it takes into account the fact that its R&D may replace profits it already earns (Reinganum 1985). In other models (Grossman and Helpman 1991), incumbents do not race at all.[2]

[2] In that model firms borrow to finance their R&D investments and the arrival rate of innovations is linear in firms' investments. In this case, the incumbent is at such a disadvantage vis a vis its rivals that it cannot finance subsequent innovations. It should be noted that in models that contain more asymmetry between firms, the assumption made here would significantly affect the properties of the resulting equilibria.

2.2 The Nature of Inventions and a System of Property Rights

A discovery is an improvement in product quality. The extent of an improvement is denoted $u_q \in [0, \bar{u}]$; where $\bar{u} < \infty$.[3] The magnitude of improvements is random, unknown until the time of invention, and common knowledge thereafter. For each invention, u is drawn from the continuous density $f(u)$ with corresponding cumulative density $F(u)$. This distribution is constant through time and unaffected by the level of a firm's R&D spending.

Once a discovery has been made, it can be reverse-engineered at zero cost by all other firms. If a patent is granted, the inventor receives an exclusive right to produce and sell that invention. The statutory life of the patent is infinite. Not all inventions will be protected, however. Let $s \in [0, \bar{u}]$ denote the minimum extent of improvement for which the patent office is willing to grant a patent. This represents the standard of nonobviousness. An invention whose extent is less than s is not protected and becomes part of the public domain of product improvements. Let $\theta(s) = 1 - F(s)$ denote the ex-ante probability of obtaining patent protection, given the standard of nonobviousness s.

2.2.1 Reverse Engineering. Patent claims are defined as the improvement itself, so each improvement does not infringe a patent on another improvement. But when, and under what conditions, will an inventor be able to use prior generations of improvements in her product? For example, the firm might be required to license all prior improvements from their inventors. At the other extreme, an inventor could use all prior discoveries without

[3] Alternatively, we can express innovations as some percent reduction in the cost of producing some final good. In that case $\bar{u} < 1$. The analysis is consistent with the quality improvement representation used in the paper as long as we assume that cost reductions are perfectly compatible, so that a cost reduction applied to different vintages of technology achieves the same percent reduction in cost.

obtaining a license. In this paper, we adopt an intermediate case: if an invention satisfies the standard of nonobviousness, the inventor may use all prior discoveries without licensing them. However, if the standard is not satisfied, the prior discoveries remain proprietary. One implication of this specification is that there is always, at most, one protected invention. Thus while the statutory length of patent protection is infinite, the economic life of a patent is the amount of time until the next patentable invention.[4]

Lach and Rob (1996) adopt an alternative approach, where firms embody new technology in vintage-specific capital goods. In a model of Cournot competition, the introduction of new technologies leads to a more gradual erosion of profits until the older firms exit altogether. O'Donoghue (1998) incorporates the effect of licensing games through an exogenous transactions cost. The higher this cost, the more stringent the standard of patentability set by a social planner. That intuition would also apply to this model.

2.3 The Output Market and Flow Profits

All consumers are identical and aggregate demand is normalized to one. The reservation value of the final product to consumers is simply the level of its quality, multiplied by *p*, the price of the final good relative to the R&D inputs (we'll suppress *p* until it becomes important in the comparative statics).[5] Let u_q^p denote the extent of the innovation protected during race *q*. This is not necessarily the invention from the previous race.

[4] This definition is consistent with the "reverse engineering" defense Congress established for *mask rights*, a sui generis form of intellectual property protecting the physical layout of computer chips (Hunt 1999a). Similar features were adopted for semiconductor topographies in the 1989 Treaty on Intellectual Property in Respect of Integrated Circuits and in the agreement on trade-related aspects of intellectual property rights (TRIPs) adopted in the Uruguay round of GATT negotiations.

[5] These assumptions make the consumer's problem stationary. If we characterize innovations as cost

Firms compete in prices and the cost of production is zero. The system of property rights described above implies the incumbent can offer a product incorporating the best available technology, embodying all of the improvements that have already occurred. Challengers can offer a product that embodies all previous innovations except for the patented one. Then, the equilibrium price of the final good is $u_q^p,$ the incumbent earns flow profit $u_q^p,$ and all challengers earn a flow profit of zero.

Flow Profits Earned during Patent Race *q*+1

	Innovation *q* was	
The Firm is	Patentable	Unpatentable
The leader from race (*q*-1)	0	$u_q^p,$
The winning challenger	u_q	0
A losing challenger	0	0

Firm i's flow profits during the next patent race depend on two outcomes that occur in the current one: (1) who invents first; and (2) whether or not the invention is important enough to be patentable. If firm i discovers an improvement, and it satisfies the standard of nonobviousness, it earns a flow profit of u_q during race $q+1$. If the improvement satisfies the nonobviousness standard, but is discovered by another firm, i's flow profits will be zero. If the improvement is found to be obvious, the flow profits of each firm in the next race will be the same as those earned in the current race.

2.4 The Existing Literature

reductions, we get the same behavior by assuming a constant elasticity of demand function with an elasticity of one.

The theoretical literature on innovation and intellectual property design is voluminous. This section reviews only the work most closely related to the model presented here.

2.4.1 Patent Races and Endogenous Growth. The model builds on an extensive literature on stochastic patent races (Loury 1979, Dasgupta and Stiglitz 1980, Lee and Wilde 1980, and Reinganum 1985). The resulting equilibrium is similar to ones analyzed in certain models of endogenous growth (Aghion and Howitt 1992 and Grossman and Helpman 1991). The model developed here introduces randomness in the magnitude of discoveries and an explicit specification of intellectual property rights in which not all inventions qualify for protection. One can interpret the previous models as an extreme case of the model constructed here, when all innovations satisfy the nonobviousness requirement and every discovery eliminates the rents associated with the prior one. By allowing for heterogeneity in the magnitude of discoveries, and conditioning patent protection on the basis of this magnitude, we can examine the effect of variations in intellectual property protection.

2.4.2 Optimal Patent Design. The early literature on optimal patent design focused on the trade-off between providing an incentive to innovate and the distortions that result from monopolistic pricing. The initial work (Nordhaus 1969) focused exclusively on patent length. More recent work considers the optimal combination of patent length and *breadth* (Gilbert and Shapiro 1990 and Klemperer 1990). Breadth is the degree to which a product or process must differ from a patented one to avoid infringement of the patent.

Patent breadth and nonobviousness are distinct concepts. Patent breadth affects the likelihood that a new invention will infringe the patent on a prior one. The nonobviousness requirement distinguishes between proprietary and non-proprietary discoveries. An invention may be obvious and yet may not infringe an existing patent. Conversely, an invention may satisfy the standard of nonobviousness and yet still infringe the claims of a prior patent.

2.4.3 Patent Design with Cumulative Innovation. Another line of research (Green and Scotchmer 1995 and Scotchmer 1996) examines the role of patents in the context of cumulative innovation, i.e., where inventions build on each other. These papers examine how patent breadth may be used to allocate rents between an initial and subsequent innovators. This paper focuses exclusively on the standard of nonobviousness, abstracting from issues of licensing and the division of rents to focus on how the proportion of proprietary and non-proprietary discoveries affects the economic value of patents. Costly licensing would affect the results of this paper in the obvious way, increasing the advantages of maintaining a relatively strict standard of nonobviousness.

Several papers have reported a non-monotonicity result similar to proposition 5 of this paper (Cadot and Lippman 1995, O'Donoghue, 1998, and Bessen and Maskin 2001). Unlike those papers, the model in this paper allows for an endogenous industrial structure and considers how patentability criteria influence the willingness of firms to enter into R&D competition. The model also allows us to contrast the effect of patentability criteria across industries differentiated by their propensity to innovate and demonstrates that more stringent criteria are appropriate for industries that innovate more rapidly.

3. Equilibrium

3.1 The Stage Games

In this model, the leading firm is a passive recipient of rents earned on its previous patentable discovery. Eventually an innovation will occur, ending the current race and possibly the incumbent's rents. During the current race, challengers select the R&D intensity that maximizes expected current cash flow plus the expected present value of competing optimally in future races. The exact magnitude of flow profits associated with a patentable discovery is not known until the discovery has actually occurred. Firms take into account the expected invention magnitude of patentable discoveries $(\tilde{u})$ when choosing their R&D intensity. The challengers move simultaneously, taking the number of their rivals as given (the participation constraint is addressed explicitly in the next section).

Let $V^i(h^i, a^i)$ denote the value function for the challenger i. Let V^w and V^l denote, respectively, the continuation values associated with playing optimally in all future races when the firm wins or loses the current one (the time subscripts have been suppressed in the text). Firms incur R&D expenses until the first discovery occurs. The value of competing actively in the current race, after sinking the cost of establishing an R&D lab, is then

$$[1] \qquad V^i(h^i, a^i) = \int_0^\infty \left\{ \lambda h^i V^w + \lambda a^i V^l - C(h^i) \right\} e^{-\lambda[h^i + a^i]t - rt} dt = \frac{\lambda h^i V^w + \lambda a^i V^l - C(h^i)}{\lambda h^i + \lambda a^i + r}.$$

The first order condition of the firm problem identifies the level of R&D where the marginal cost of additional effort is just equal to the marginal benefit of winning, rather than losing, the current race, that is $C'(h^i) = \lambda[V^w \text{-} V^i(h^i, a^i)]$.

3.2 The Stationary Symmetric Equilibrium of the Game

A strategy of a firm in the game is a specification of a feasible R&D intensity to be played in each race, for each possible history of the game preceding that race. At the beginning of each race, each firm knows the play of all firms in the prior races and the outcomes of those races. When the firm is the incumbent, its only feasible R&D intensity is zero. Whenever the firm is a challenger, the set of feasible R&D intensities is always the same subset of $\mathbb{R}$. There are likely to be many equilibria of the game, but we focus on stationary equilibria where firms choose identical strategies. In the appendix, we prove the following:

Proposition 1 - Suppose the R&D cost function satisfies the following assumptions:

(i) $C'(h) > 0,\ C''(h) > 0,\ \forall h > 0$;

(ii) $C'(h)h - C(h) \geq 0\ \forall h > \hat{h} \in [0, \infty)$;

(iii) $\bar{h},\ C(\bar{h}) < \infty$;

(iv) $\text{Lim}_{h\to\infty}\ C'(h) = \infty$;

(v) $\text{Lim}_{h\to 0}\ C(h)/h = \text{Lim}_{h\to 0}\ C'(h) = 0$;

Then, there exists a unique, stationary, symmetric equilibrium of the game.

The R&D technology, the distribution of invention magnitudes, and the relationship between patented technology and expected profits do not vary across races. The expected outcome of the races, then, varies only if firms choose different R&D intensities over time. If all challengers choose the same R&D intensity in all patent races, the probabilities of winning and losing, together with the expected length of races, will be the same in each race. Similarly, the continuation values associated with being the incumbent or a challenger, denoted $V^I(h)$ and $V^C(h)$, are the same across patent races. Ex ante then, each race looks the same to the challengers. We examine an equilibrium where firms respond to the same conditions in the same way through time.

For a challenger, the continuation value associated with losing the current race is the expected value of being a challenger in the subsequent race, i.e., $V^{l} = V^{C}(h)$. For the winner of the current race, the expected value of making the first discovery is a weighted average of the continuation values associated with starting the next race as the incumbent or as a challenger. The weights depend on the probability that the invention is patentable. Thus, $V^{w} = \theta V^{I}(h) + (1-\theta)V^{C}(h)$. If firms choose the same R&D intensity in each patent race, the continuation values are simply a weighted average of the expected flow profit enjoyed by incumbents and the R&D expenditures of challengers:

$$V^{I}(h) = \frac{[r + \theta\lambda h]\tilde{u} - \theta\lambda nhC(h)}{r[r + \theta\lambda(n+1)h]} \qquad \text{and} \qquad V^{C}(h) = \frac{\theta\lambda h\tilde{u} - [r + \theta\lambda nh]C(h)}{r[r + \theta\lambda(n+1)h]}.$$

Substituting these continuation values into the first order condition, we have

[2] $$C'(\sigma) = \theta\lambda[V^{I}(\sigma) - V^{C}(\sigma)] = \theta\lambda\left(\frac{u^{e} + C(\sigma)}{r + \theta\lambda(n+1)\sigma}\right).$$

Challengers select the R&D intensity, denoted σ, that equates marginal cost with marginal benefits, which is the difference in the expected value of the cash flows associated with starting the next race as the incumbent rather than as a challenger. The denominator on the right hand side of [2] is a measure of the *economic* life of patents. As $\theta\lambda(n+1)\sigma$ becomes larger, patentable discoveries occur more frequently. The incumbent enjoys her rents for less time, on average, so the present value of the rents is smaller.

3.2.1 The Participation Constraint. The first order condition specifies a challenger's R&D intensity as a function of exogenous parameters and the number of rivals, which is endogenous. The number of rivals is determined by a participation constraint - firms enter a patent race so long as $V(\sigma,(n-1)\sigma)-k\geq 0.$ This implies that

[3] $$\left(\frac{C(\sigma)+rk}{\sigma}\right)\leq\theta\lambda[V^{I}(\sigma)-V^{C}(\sigma)].$$

When the participation constraint binds, average R&D costs are just equal to the marginal benefit associated with winning the current patent race. This implies that average and marginal costs are just equal. Using equation [2], equation [3] can also be expressed as

[4] $$C'(\sigma)\leq\frac{\theta\lambda\left[\tilde{u}-rk\right]}{r+\theta\lambda n\sigma}.$$

Thus there will be no active patent races unless the revenues generated in the output market can amortize fixed R&D costs. But even when $k=0,$ so long as expected revenues are finite, *n* is uniquely determined and finite.

3.3 Properties of the Equilibrium

In a stationary equilibrium where there are no random shocks, firms that wish to compete will sink their fixed R&D investments at the beginning of the first patent race.[6] Thereafter, if we consider marginal changes in certain parameters, the number of firms engaged in R&D would not change because the expected value of actively competing in subsequent races is strictly positive (so long as $k > 0$).

Instead, we consider two games involving industries with a different value for a single exogenous parameter. The industries are otherwise identical. In each industry, the participation constraint binds. We compare the resulting symmetric stationary equilibria. Firms take into account the exogenous parameters when deciding whether to incur the fixed cost of an R&D lab. In the appendix we show the following:

> **Proposition 2** - The R&D intensity of individual firms does not vary with differences in output prices (p) and the productivity or R&D (λ). But more firms will engage in R&D in the industry with either the higher output price or more productive R&D. Consequently the industry-wide rate of innovation will be higher.

[6] This is not to say there is no uncertainty in the model. We are just pointing out that after firms decide to enter the industry in the first race, there are no shocks to firms' research productivity or costs that would imply subsequent entry or exit.

Proposition 3 - The R&D intensity of individual firms varies with differences in the discount rate (r) and the cost of setting up an R&D lab (*k*). Higher discount rates or higher fixed R&D costs are associated with more R&D at the level of individual firms. But fewer firms will engage in R&D and the resulting industry-wide rate of innovation will be lower.

Propositions 2 and 3 tell us that the industry-wide rate of innovation is higher when the output price or the productivity of R&D is higher, but lower when the discount rate or the fixed cost of establishing an R&D facility is higher. In each of these cases, the determining factor is the number of firms engaged in R&D.

3.3.1 Marginal vs. Average Cost. Holding σ and *n* constant, differences in *p* or λ imply differences in the value of winning the current race, $\theta\lambda[V^I - V^C]$ across industries. Holding, *n* constant, firms will select their R&D intensity so that marginal cost and marginal benefit are equal. Without a difference in the number of firms, there is an inequality between marginal and average cost in one or both industries. This in turn implies that $V^C \neq k$. If *n* is allowed to vary, and the entry constraint binds, the equality between marginal and average cost is restored. The difference in the number of firms across the industries is such that $\theta\lambda[V^I - V^C]$ is the same for in both industries. As a result, marginal cost, average cost, and marginal benefit in these two industries are the same. Since we've assumed *r* and *k* are the same, firms in both industries choose the same R&D intensity. Differences in the industry-wide rate of innovation are thus due entirely to differences in the number of firms.

Differences in *r* or *k* imply differences in average costs across industries. This also implies there is a difference in R&D investments at the firm level, because the binding participation constraint implies marginal cost and average cost are the same. If firms' R&D investments in the two industries are different, the marginal benefit from winning the current race must also be different. This results from a difference in the number of firms, which implies a different industry wide rate of innovation. The industry with a higher average cost will have fewer firms, and this more than offsets the fact the remaining firms are each doing more research.

3.3.2 Market Structure and Innovation. The parameters described in propositions 2 and 3 are sufficient to describe a wide variety of industry structures and innovation rates. These propositions are sufficient to show that different industry structures, in themselves, cannot be ordered in terms of their rate of innovation.

In the environment most closely analogous to perfect competition, fixed costs are zero so there are many firms, each engaged in a little R&D. Whether we consider this industry to be highly innovative will depend on the productivity of R&D and its relative cost. Greater concentration, in the sense of relatively few firms competing to innovate, occurs when fixed costs are relatively high. Individual firms will do more R&D than we would see in the previous case. But whether we would consider this industry highly innovative again depends on the productivity of R&D and its relative cost.

Other things equal, we would expect the industry with lower fixed costs to be more innovative (proposition 3). But when other things are not equal, it is possible that a more concentrated industry would innovate more rapidly. This might occur, for example, when an industry has a relatively high fixed cost to establishing research facilities but relatively low marginal costs of using them more intensively.

3.3.3 Contrast with Fixed Industry Structure. The properties of this model can be contrasted with the model in Hunt (1999a), which is similar, but which assumes the number of firms is fixed. In that model, per firm R&D intensity increases with the relative price of output. R&D intensity declines as the productivity of R&D rises, but effective R&D ($\lambda\sigma$) increases. A higher discount rate is associated with less R&D at the firm level, and therefore the industry level. An exogenous increase in the number of firms reduces the R&D intensity of individual firms, but industry wide R&D rises as long as a stability condition is satisfied.

3.4 Comparing the Equilibrium to the First Best

It is useful to compare the private equilibrium R&D intensity and number of firms to the social planner's solution. Unlike firms, society enjoys a permanent benefit to each innovation regardless of its magnitude – $\tilde{v}^S \equiv p\tilde{u}(0)/r$, where $\tilde{u}(0)$ is the mean of the distribution of invention magnitudes. Expected social welfare at the beginning of the game is

[5] $$W = \underset{h,n}{Max}\left\{\frac{n}{r}\left(\lambda h\tilde{v}^S - C(h) - rk\right)\right\}.$$

The first derivatives are

$$\frac{\partial W}{\partial h} = \frac{n}{r}\left[\lambda\tilde{v}^S - C'(h)\right], \qquad \frac{\partial W}{\partial h} = \frac{1}{r}\left[\lambda h\tilde{v}^S - C(h) - rk\right].$$

Thus the social planner would specify a firm level R&D intensity that equalizes marginal cost and the marginal social benefit of the next discovery. This R&D intensity is strictly greater than the level attained in the private equilibrium because

$$\frac{p\tilde{u}(0)}{r} > \frac{\theta[p\tilde{u} - rk]}{r + \lambda nh}.$$

As for the optimal number of firms, there is a razor edge result. Generically, the marginal social benefit of innovation will be either greater or less than a firm's average cost at the first best R&D intensity. Consequently the social planner will either desire an infinite number of firms, or no firms, to establish R&D facilities. If we assume the average cost is less than the marginal social benefit to innovation, it is apparent the social planner would wish to subsidize both entry and R&D effort.

4. The Standard of Nonobviousness and the Rate of Innovation

4.1 General Results

We typically think of the U.S. patent system as applying a common set of criteria to inventions in all technology fields and industries. In this section, however, we construct a hypothetical in which two otherwise identical industries are subject to different standards of nonobviousness. Firms take this standard into account when deciding whether to sink the fixed cost of an R&D lab. In this way we allow for the possibility that patentability criteria affect the number of firms engaged in R&D.

In equilibrium, firms equate the marginal cost of additional R&D effort to the expected gain associated with inventing first. This gain is affected by the nonobviousness standard in two ways. First, there is the likelihood that any given invention by a firm qualifies for protection. Second, there is a relationship between this probability and the number of rivals a firm competes with. In the appendix we show the following:

Proposition 4 - In the stationary symmetric equilibrium, differences in the standard of nonobviousness do not affect the R&D intensity of individual firms, but they do affect the number of firms actively engaged in R&D, and therefore the industry-wide rate of innovation.

The intuition behind propositions 2 and 4 is very similar. Differences in patentability criteria do not affect the left hand side of equations [2] and [4], so marginal and average cost in the two industries will be the same. This implies the marginal benefit to innovating first is also the same. This equality is the result of a difference in the number of firms in these industries. But which industry has more firms and therefore innovates faster? In the appendix, we show

Proposition 5 - There exists a unique standard of nonobviousness, denoted *s**, such that in the interval $[0, s^*)$, industry-wide R&D activity is strictly *increasing* in the standard of nonobviousness. In the interval $(s^*, \bar{u}]$, industry-wide R&D activity is strictly decreasing in the standard of nonobviousness.

Proposition 5 tells us that differences in the rate of innovation between two industries will depend on their patentability standards relative to each other and relative to *s**. Consider two otherwise identical industries with patentability standards s_1 and s_2. When $s^* \leq s_1 < s_2$, the first industry will innovate more rapidly than the second. This reflects the conventional wisdom that weaker patentability criteria can stimulate additional R&D activity. But when $s_1 < s_2 \leq s^*$, the industry with the more stringent patentability criteria innovates faster. Thus, without knowing *s** and the actual standard of nonobviousness relative to it, we cannot say a priori whether a change in the standard will increase or decrease the rate of innovation.

4.1.1 Deriving the Optimal Standard of Nonobviousness. In the appendix, we show that s^* is implicitly defined by the following equation:

[6] $$\Psi(s) \equiv \left(\frac{\theta\lambda n\sigma}{r+\theta\lambda n\sigma}\right)\left[p\tilde{u} - rk\right] - \left[ps - rk\right] = 0.$$

As the nonobviousness requirement is made more strict, firms encounter the following tradeoff. On the one hand, a firm that makes a marginal discovery would not obtain a patent. This makes the participation constraint [4] harder to satisfy, which is reflected in the second term in equation [6]. Call this the static effect of tighter patentability criteria. On the other hand, raising the nonobviousness standard increases the amount of time an incumbent can expect to remain the incumbent. This increases the value of the remaining patentable discoveries, which is reflected in the first term in equation [6]. Call this the dynamic effect of tighter patentability criteria.

When the nonobviousness standard is very weak (s = 0), the static effect is irrelevant because such marginal inventions do not affect the participation decision (the static effect is only important when $ps \geq rk$). In this range, an increase in the standard of nonobviousness will increase the number of firms actively engaged in R&D. But as the requirement is made more and more strict, the dynamic effect becomes smaller ($\theta\lambda n\sigma$ eventually declines as *s* increases) while the static effect becomes larger. When the nonobviousness requirement is very strict, the static effect dominates. There is only one standard of nonobviousness, where the two effects are exactly equal.

4.2.1 Implications. The most obvious implication of [6] is that the optimal standard (optimal in the sense that it maximizes the rate of innovation) depends on the characteristics of the industry that determine its underlying propensity to innovate. If those parameters change, so would *s**. If those parameters vary across industries, there would be a unique, but different standard that maximizes the rate of innovation in those industries. Industry structure, which is determined by *r* and *k* in this model, are relevant in the determination of *s** only to the extent that if influences the industry's rate of innovation.

Note that in [6], the higher the arrival rate of innovations, the more positive is the sign of $\Psi(s)$. This should imply that *s** will be more stringent in industries that innovate more rapidly than in industries that innovate less rapidly. In the appendix, we prove

> **Proposition 6** - The critical standard of nonobviousness, *s**, is increasing in the equilibrium rate of innovation. This implies that in rapidly innovating industries, a smaller proportion of inventions can be protected without causing the rate of innovation to decline.

Proposition 6 tells is the relative importance of the static and dynamic effects, described in the previous section, vary across industries distinguished by their propensity to innovate. Consider two industries 1 and 2, where industry 1 innovates less rapidly than industry 2.

Increasing the standard of nonobviousness in industry 1 increases an incumbent's expected tenure. But this tenure is already relatively long, so we are talking about adding profits earned in the distant future, and therefore heavily discounted. If we consider the same experiment in industry 2, extending the tenure of incumbents is adding profits that earned relatively soon, and therefore not discounted very much. So the dynamic effect is more important when firms innovate rapidly than when firms innovate more slowly.

When a common patentability requirement is applied to all industries, as is typically asserted for the U.S. patent system, it is more likely to lie in the dynamic range for a rapidly innovating industry than it is for an industry that innovates less rapidly. Adopting less stringent criteria is more likely to reduce the rate of innovation in industries that typically innovate rapidly while increasing the rate of innovation in industries that typically innovate more slowly.[7] The mechanism for these changes is an increase or decrease in the number of firms actively engaged in R&D in those industries.

[7] This property of rapidly innovating industries suggests that a number of changes in U.S. intellectual property law that effectively relaxed patentability criteria during the 1980s may have favored innovation in traditional industries over high technology industries. See Hunt 1999a and 1999b.

5. Discussion and Conclusions

This paper develops a model of cumulative innovation where the profitability of inventions is eroded by the introduction of new, competing technologies through time. When firms can readily duplicate each other's discoveries, the nonobviousness requirement plays an important role in determining the proportion of discoveries that affect the profits earned on proprietary discoveries. It also influences the degree of rivalry a firm encounters, because it plays a role in determining the number of firms actively engaged in R&D.

In such an environment, there exists a unique interior standard of nonobviousness that maximizes the rate of innovation in an industry by maximizing the number of firms that enter into R&D competition. The effect of changes in patentability criteria on the industry wide rate of innovation, then, depend on whether the initial standard is more or less stringent than this optimal value. Also, the optimal standard is more stringent for industries that innovate rapidly than for those industries that innovate slowly.

If a common set of patentability criteria are applied to all industries in an economy, the number of firms engaged in R&D in each of those industries will depend on the stringency of those standards. Generally speaking, when the standard is more stringent, there will be more firms in the industries disposed to innovate more rapidly and fewer firms in industries disposed to innovate less rapidly than there would be with a weaker standard. A social planner would take into account the relative size of these industries when setting the optimal standard. In setting patentability criteria, then, we are also setting industrial policy.

In a similar fashion, a social planner in another country, with different mixes of industries, would likely adopt a different patentability standard. In general, the country whose high technology industries (i.e. those that innovate rapidly) represent a larger share of the economy would adopt the more strict standard. One implication is that, under an optimal standard, patents would be easier to obtain in less developed countries than in more developed ones. Adopting an international standard of patentability may increase the rate of innovation in some countries, but reduce it in others.[8]

Efforts towards patent harmonization have thus far concentrated on issues such as establishing uniform priority, a minimum patent length, fewer subject matter exceptions, adequate remedies for infringement (damages, injunctions), and adequate administrative and judicial infrastructures. One exception was the proposed Patent Harmonization Treaty, abandoned in the mid 1990s, which included a specification of patentability standards (Moy 1993). In March 2001, the U.S. Patent and Trademark Office proposed to include, among other things, an American style nonobviousness test in its agenda for future international negotiations on patent harmonization.

[8] More general statements about welfare require a model that allows for trade, foreign direct investment, and licensing. Excellent surveys of the relevant development literature can be found in Maskus (2000) and Saggi (2001). A general conclusion is that the ongoing trend towards stronger patent rights around the world may lead to not insignificant transfers of wealth from users of patented technologies in developing economies and patent holders, often located in the most developed economies. But there is also some empirical evidence of dynamic gains in the form of increased imports of high technology goods, greater foreign direct investment, and increased opportunities to license advanced technologies. A key factor in determining whether the benefits of these changes exceeds the cost is the degree to which innovation is increased in developed economies, developing economies or both.

Given that the optimal standard is a function of industry characteristics that influence the industry's rate of innovation, this standard is likely to vary as those characteristics change. An economy-wide increase in the productivity of R&D, for example, might suggest that patentability criteria should be strengthened in order to obtain the maximum possible benefit of this new found productivity. If the productivity increase occurred in a single industry, a social planner would likely adopt a more stringent standard, but doing so would reduce the rate of innovation in the other industries. This might suggest that the standard of nonobviousness should take the form of a balancing test rather than a statutory threshold.[9]

Many practitioners and scholars of the patent system would support relatively stable patentability criteria and an equal treatment of all patentable technologies. A principal argument in this vein is that the patent system is already costly and additional complexity would only increase these costs while also increasing uncertainty about future returns. In practice, it is often difficult to distinguish where one industry or technology ends and another begins, so implementing a multiplicity of criteria may not be feasible.

[9] The nonobviousness requirement was incorporated into U.S. law by the 1952 Patent Act. Before then, the courts employed a similar test, which was derived from judicial precedents over the previous century (Hunt 1999b).

But it is worth pointing out that the U.S. engaged in precisely these experiments during the 1980s, relaxing the standard of nonobviousness for patents and adopting a new form of intellectual property (mask rights) to protect the physical layout of semiconductor chips (Hunt 1999a, 1999b). Many argued these changes would stimulate innovation in America's high technology industries. The results of this paper suggest the opposite might well be true, because relaxing a uniform standard of nonobviousness is relatively more likely to encourage innovation in industries that innovate less rapidly by increasing entry into that industry. Conversely, such a change is more likely to discourage innovation in rapidly innovating industries, by discouraging entry into the industry.

The final assessment of the changes adopted in the 1980s is of course an empirical question. This model suggests at least one testable implication: historical patterns of entry and exit from industries should may have changed in some systematic way – with relatively more net entry into industries that innovate less rapidly and relatively less net entry into industries that innovate more rapidly. A topic of future research, then, is to establish whether this conjecture is borne out in the data.

REFERENCES

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction," *Econometrica*. Vol. 60: 323-51.

Bessen, James and Eric Maskin. 2001. "Sequential Innovation, Patents, and Imitation," MIT Department of Economics Working Paper No. 00-01.

Cadot, Olivier and Lippman, Steven A. 1995. "Barriers to Imitation and the Incentive to Innovate," mimeo, INSEAD.

Cohen, Wesley M. and Richard C. Levin. 1989. "Empirical Studies of Innovation and Market Structure," in Schmalensee, Richard and Robert D. Willig, eds., *Handbook of Industrial Organization*. Amsterdam: North Holland, pp. 1059-1107.

Dasgupta, Partha and Joseph Stiglitz. 1980. "Uncertainty, Industrial Structure, and the Speed of R&D," *The Bell Journal of Economics*. Vol. 11: 1-28.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. Cambridge, MA: The MIT Press.

Gilbert, Robert and Carl Shapiro. 1990. "Optimal Patent Length and Breadth," *The RAND Journal of Economics*. Vol. 21:106-12.

Green, Jerry and Suzanne Scotchmer. 1995. "On the Division of Profit in Sequential Innovation," *The RAND Journal of Economics*. Vol. 26: 20-33.

Grossman, Gene M., and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy*. Cambridge, MA: The MIT Press.

Hall, Bronwyn H., and Rosemarie H. Ziedonis. 2001. "The Patent Paradox Revisited: Determinants of Patenting in the US Semiconductor Industry, 1979-95," *The RAND Journal of Economics*. Vol. 32: 101-28.

Hunt, Robert M. 1999a. "Nonobviousness and the Incentive to Innovate: An Economic Analysis Of Intellectual Property Reform," Working Paper 99-3, Federal Reserve Bank of Philadelphia.

_____. 1999b. "Patent Reform: A Mixed Blessing for the U.S. Economy?" Federal Reserve Bank of Philadelphia *Business Review* (November/December).

Klemperer, Paul. 1990. "How Broad Should the Scope of Patent Protection Be?" *The RAND Journal of Economics*. Vol. 21: 113-30.

Lach, Saul and Rafael Rob. 1996. "R&D, Investment and Industry Dynamics," *Journal of Economics and Management Strategy*. Vol. 5: 217-49.

Lee, Tom and Louis L. Wilde. 1980. "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economics*. Vol. 94: 429-36.

Loury, Glenn C. 1979. "Market Structure and Innovation," *Quarterly Journal of Economics*. Vol. 93: 395-410.

Maskus, Keith. E. 2000. *Intellectual Property Rights in the Global Economy*. Washington: Institute for International Economics.

Moy, R. Carl. 1993. "The History of the Patent Harmonization Treaty: Economic self-interest as an Influence," *The John Marshall Law Review*. Vol. 26: 457-95.

Nordhaus, W. 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, MA: MIT Press.

O'Donoghue. 1998. "A Patentability Requirement for Sequential Innovation," *The RAND Journal of Economics*. Vol. 29: 654-679.

Reinganum, Jennifer F. 1985. "Innovation and Industry Evolution," *Quarterly Journal of Economics*. Vol. 100: 81-99.

Saggi, Kamal. 2001. "Trade, Foreign Direct Investment, and International Technology Transfer: A Survey," mimeo, Department of Economics, Southern Methodist University.

Scherer, Frederic M. 1992. "Schumpeter and Plausible Capitalism," *Journal of Economic Literature*," Vol. XXX: 1416-33.

Scotchmer, Suzanne. 1996. "Protecting Early Innovators: Should Second Generation Products Be Patentable?" *The RAND Journal of Economics*. Vol. 27: 322-31.

Scotchmer, Suzanne and Jerry Green. 1990. "Novelty and Disclosure in Patent Law," *The RAND Journal of Economics*. Vol. 21: 131-46.

APPENDIX

Proposition 1 - Suppose the R&D cost function satisfies the following assumptions:

(i) $C'(h) > 0,\ C''(h) > 0,\ \forall h > 0$;

(ii) $C'(h)h - C(h) \geq 0\ \forall h > \hat{h} \in [0, \infty)$;

(iii) $\bar{h},\ C(\bar{h}) < \infty$;

(iv) $\text{Lim}_{h \to \infty}\ C'(h) = \infty$;

(v) $\text{Lim}_{h \to 0}\ C(h)/h = \text{Lim}_{h \to 0}\ C'(h) = 0$;

Then, there exists a unique, stationary, symmetric equilibrium of the game.

Proof: The proof is constructed through the lemmas that follow.

Lemma 1 - Suppose $V^W_{q+1} \in (0, \infty)$ and $V^W_{q+1} - V^L_{q+1} > 0.$ If rivalry and the fixed R&D costs are sufficiently small, at least one challenger will choose to enter a stage game.

Proof: Note that we are treating the continuation values as exogenous parameters. Later we show that, in equilibrium, the continuation values satisfy the requirements set out in the lemmas. Consider the case where there is no rivalry and fixed R&D costs are zero. We need to show that $V^i_q(h^i, 0) \geq V^i_q(0,0).$ The inequality is satisfied when there exists some positive level of R&D intensity where $\lambda h^i_q V^W_{q+1} \geq C(h^i_q),$ which is satisfied if the minimum average cost of R&D is not too high. The fifth assumption above ensures there is at least one R&D intensity $\tilde{h} \in (0, \bar{h}]$ where the inequality is strict.

Now we consider a strictly positive fixed R&D cost k. In that case, a challenger chooses to enter so long as $\lambda h^i_q V^W_{q+1} \geq C(h^i_q) + [r + h^i_q]k.$ If $\lambda \hat{h}^i_q V^W_{q+1} > C(\hat{h}^i_q),$ there is also a level of fixed R&D cost where $\lambda \hat{h}^i_q V^W_{q+1} \geq C(\hat{h}^i_q) + [r + \hat{h}^i_q]k.$ Thus for k sufficiently small, there are always at least two firms, an incumbent and at least one challenger.

Now suppose there is some small positive level of rivalry. A challenger will enter the stage game if the following inequality holds:

[A.1] $$\lambda h^i_q \left\{ rV^W_{q+1} + \lambda a^i_q \left[V^W_{q+1} - V^L_{q+1} \right] \right\} \geq \left[r + \lambda a^i_q \right] C(h^i_q) \left[r + \lambda h^i_q + \lambda a^i_q \right] rk.$$

Applying the preceding argument to this inequality, for k and a^i_q sufficiently small, there is an R&D intensity in the interval $(0, \bar{h}\,]$ where this inequality is strict. So long as $V^W_{q+1} > 0$ and $V^W_{k+1} - V^L_{k+1} > 0$, the magnitude of the continuation values will always define a set of pairs $(a^i_q, k) \in \mathbb{R}^+$ where the participation constraint is satisfied. We can also define a level of fixed R&D cost, $\hat{k}(a^i_q)$ where the participation constraint just binds.

Lemma 2 - If $V^{W}_{q+1} \in (0,\infty)$, $V^{W}_{q+1} - V^{L}_{q+1} > 0$, and $k < \hat{k}(0)$, there exists an interior equilibrium of the stage game.

Proof: The proof of existence is a modification of the existence proof in Reinganum (1985). We continue to treat the continuation values as exogenous parameters, but take into account the effect of a firm's choice of R&D intensity on the likelihood of winning and the expected length of the patent race. Firms take their rival's research intensity as given. Fixed R&D costs must be sufficiently low so that at least one firm is willing to engage in R&D.

The derivative of the firm's objective function, $\partial V^{i}_{q} / \partial h^{i}_{q}$, is

[A.2] $$\frac{r\left[\lambda V^{W}_{q+1} - C'(h^{i}_{q})\right] + \lambda a^{i}_{q}\left[\lambda[V^{W}_{q+1} - V^{L}_{q+1}] - C'(h^{i}_{q})\right] + \lambda\left[C(h^{i}_{q}) - C'(h^{i}_{q})h^{i}_{q}\right]}{\left[r + \lambda(h^{i}_{q} + a^{i}_{q})\right]^{2}}.$$

The sign of [A1] is the sign of the numerator, which we will call $\phi^{i}(h^{i}_{q}, a^{i}_{q})$. Note that $\phi^{i}(h^{i}_{q}, a^{i}_{q})$ is strictly decreasing in R&D intensity:

$$\frac{\partial \phi^{i}(h^{i}_{q}, a^{i}_{q})}{\partial h^{i}_{q}} = - C''(h^{i}_{q}) \bullet \left[r + \lambda(h^{i}_{q} + a^{i}_{q})\right] < 0.$$

If the saturation point of R&D ($\bar{h}$) is sufficiently large, there will be a finite level of R&D effort where $\phi^{i}(h^{i}_{q}, a^{i}_{q}) = 0$. $V^{i}_{q}(h^{i}_{q}, a^{i}_{q})$ is maximized by this level of R&D effort. Let $h^{i}_{q}(a^{i}_{q})$ denote the firm's best response to the level of rivalry it encounters. The strict monotonicity of $\phi^{i}(h^{i}_{q}, a^{i}_{q})$ implies that this best response is unique. Firms never choose R&D intensities greater than $\tilde{h}^{i}_{k}$, so we can restrict the strategy space to a convex, compact, nonempty subset of $\mathbb{R}^{n}$, denoted $X \equiv \Pi^{n}_{i=1}[0, \tilde{h}^{i}_{q}]$. The vector $[h^{1}_{q}(a^{1}_{q}), h^{2}_{q}(a^{2}_{q}), \ldots h^{n}_{q}(a^{n}_{q})]$ maps X into itself continuously. Existence of an equilibrium then follows from Brouwer's fixed point theorem.

Lemma 3 - If $\lambda[V^{W}_{q+1} - V^{L}_{q+1}] - [C'(h_{q}) + h_{q}C''(h_{q})] < 0$, there exists a unique, symmetric equilibrium of the stage game.

Proof: Existence of a symmetric equilibrium follows from the firm's objective function and first order condition, which varies only by the level of rivalry encountered. In the symmetric equilibrium, $\phi^{i}(h^{i}_{q}, a^{i}_{q})$ becomes $\phi^{i}\left(h_{q}, (n-1)h_{q}\right)$. The corresponding first order condition is

$$r\left[\lambda V^{w}_{q+1} - C'(h_{q})\right] + \lambda(n-1)h_{q}\left[\lambda[V^{w}_{q+1} - V^{l}_{q+1}] - C'(h_{q})\right] + \lambda\left[C(h_{q}) - C'(h_{q})h_{q}\right] = 0.$$

The first and third terms are strictly decreasing in R&D effort. If the second term is also strictly decreasing, then only one level of R&D intensity satisfies the equality. Hence we require that

$$\lambda\left[V^{w}_{q+1} - V^{l}_{q+1}\right] - \left[C'(h_{q}) + h_{q}C''(h_{q})\right] < 0.$$

The symmetric equilibrium R&D intensity of the stage game with continuation values V^{W}_{q+1} and V^{L}_{q+1} is denoted $h_q(V^{W}_{q+1}, V^{L}_{q+1})$.

Lemma 4 - The game is continuous at infinity.

Proof: It is sufficient to show that total firm payoffs are a discounted sum of per period payoffs and that these per period payoffs are uniformly bounded [see Fudenberg and Tirole (1991), p. 110]. The per period payoff to firms is the present value of flow profits for the incumbent and the present value of R&D expenditures for challengers. The maximum per period return for an incumbent is $\bar{u}/r$. Per period returns for challengers are contained in the interval $[-C(\bar{h})/(r + \bar{h}), 0]$.

Lemma 5 - Lemmas 1 - 4 imply the existence of a stationary symmetric equilibrium of the game.

Proof: We return to the first order condition of the stage game, but assume that the continuation values associated with winning and losing the current race do not vary across races. Rearranging terms, we have:

[A.3] $$C'(h_q)\left[r + \lambda n h_q\right] = \lambda\left[rV^W + \lambda(n-1)h_q[V^W - V^L] + C(h_q)\right].$$

If firms take the continuation values as given, and these values are constant across races, it is a best response for each firm to choose the same R&D intensity in each race. Lemma 3 establishes the existence of such a best response for a given specification of the continuation values. The continuation values themselves take a simple recursive form:

$$V^I(h) = \frac{\tilde{u} + \lambda nh\left[\theta V^C(h) + (1-\theta)V^I(h)\right]}{r + \lambda nh} = \frac{\tilde{u} + \theta\lambda nh V^C(h)}{r + \theta\lambda nh}, \text{ and}$$

$$V^C(h) = \frac{\lambda h\left[\theta V^I(h) + (1-\theta)V^C(h) + (n-1)V^C(h)\right] - C(h)}{r + \lambda nh} = \frac{\theta\lambda h V^I(h) - C(h)}{r + \theta\lambda h}$$

Solving for $V^I(h)$ and $V^C(h)$, we have,

$$V^I(h) = \frac{[r + \theta\lambda h]\tilde{u} - \theta\lambda nh C(h)}{r[r + \theta\lambda(n+1)h]}, \text{ and}$$

$$V^C(h) = \frac{\theta\lambda h\tilde{u} - \left[r + \theta\lambda nh\right]C(h)}{r[r + \theta\lambda(n+1)h]}.$$

If we substitute for $V^I(h)$ and $V^C(h)$ in equation [A.3], the first order condition reduces to

[A.4] $$C'(h) = \theta\lambda\left(\frac{\tilde{u} + C(h)}{r + \theta\lambda(n+1)h}\right).$$

We use σ to denote the equilibrium R&D intensity that satisfies equation [A.4]. It can be verified that, using equation [A.4], the condition required in lemma 3 for the uniqueness of the symmetric equilibrium of the stage games is satisfied.

If we substitute for $V^I(h)$ and $V^C(h)$ in equation [A.1], the participation constraint is simply $V^C(h) \geq k.$ This in turn implies $C(h) + rk \leq \theta\lambda h\left[V^I(h) - V^C(h)\right] = C'(h)h.$ When the participation constraint binds, we can express it simply as

$$C'(h) = \theta\lambda\left(\frac{\tilde{u} - rk}{r + \theta\lambda nh}\right). \tag{A.5}$$

During each race, for every challenger, the R&D intensity σ is the unique best response to the continuation values $V^I(\sigma)$ and $V^C(\sigma).$ The strategy of playing σ in every race cannot be improved upon by choosing a different R&D intensity in one race and playing σ in all the others. If playing σ in every race cannot be improved upon by a deviation in one stage, and the game is continuous at infinity, choosing the R&D intensity σ in each race is a subgame perfect equilibrium of the game [see Fudenberg and Tirole (1991), p. 110].

Lemma 6 - The symmetric stationary equilibrium is unique.

Proof: It is sufficient to show that there is only one possible intersection of the curves described by [A.4]. At h = 0, $C'(h) = 0$ while $\theta\lambda[V^I(h) - V^C(h)] = \theta\lambda\,\tilde{u}/r.$ Thus at the first intersection, the marginal cost curve must be rising faster than $\theta\lambda[V^I(h) - V^C(h)].$ If we can rule out an intersection where $\theta\lambda[V^I(h) - V^C(h)]$ is rising faster than is marginal cost we are done. Define $\mathrm{M}^1 = C'(h)\left[r + \theta\lambda(n+1)\sigma\right] - \theta\lambda[\tilde{u} + C(h)] = 0$ and note that

$$\frac{\partial \mathrm{M}^1}{\partial h} = C''(h)\left[r + \theta\lambda(n+1)h\right] + C'(h)\theta\lambda n > 0. \tag{A.6}$$

This rules out an intersection where the marginal cost curve crosses $\theta\lambda[V^I(h) - V^C(h)].$ from above.

Proposition 2 - The R&D intensity of individual firms does not vary with differences in output prices (*p*) and the productivity or R&D (λ) . But more firms will engage in R&D in the industry with either the higher output price or more productive R&D. Consequently the industry wide rate of innovation will be higher.

Proposition 3 - The R&D intensity of individual firms varies with differences in the discount rate (r) and the cost of setting up an R&D lab (*k*). Higher discount rates or higher fixed R&D costs are associated with more R&D at the level of individual firms. But fewer firms will engage in R&D and the resulting industry wide rate of innovation will be lower.

Proof: We reintroduce the relative price of outputs in terms of inputs (*p*) and rewrite [A.4] and [A.5] in following form

$$M \equiv \begin{array}{l} \mathrm{M}^1 = C'(\sigma)[r + \theta\lambda(n+1)\sigma] - \theta\lambda[p\tilde{u} + C(\sigma)] = 0 \\ \mathrm{M}^2 = C'(\sigma)[r + \theta\lambda n\sigma] - \theta\lambda[p\tilde{u} - rk] = 0 \end{array}$$

We'll need the following derivatives:

$$\mathrm{M}^1_\sigma = C''(\sigma)[r + \theta\lambda\sigma] + \left[C''(\sigma)\sigma + C'(\sigma)\right]\theta\lambda n \qquad \mathrm{M}^2_\sigma = C''(\sigma)r + \left[C''(\sigma)\sigma + C'(\sigma)\right]\theta\lambda n$$

$$\mathrm{M}^1_p = -\theta\lambda\tilde{u} \qquad \mathrm{M}^2_p = -\theta\lambda\tilde{u}$$

$$\mathrm{M}^1_\lambda = -\frac{r}{\lambda}C'(h) \qquad \mathrm{M}^1_\lambda = -\frac{r}{\lambda}C'(h)$$

$$\mathrm{M}^1_r = C'(h) \qquad \mathrm{M}^2_r = C'(h) + \theta\lambda k$$

$$\mathrm{M}^1_k = 0 \qquad \mathrm{M}^2_k = r\theta\lambda$$

The Jacobian $|\mathrm{M} =| \mathrm{M}^1_\sigma \mathrm{M}^2_n - \mathrm{M}^2_\sigma \mathrm{M}^1_n = C''(\sigma)C'(\sigma)[\theta\lambda\sigma]^2 > 0.$

i. Increasing the output price:

$$\frac{\partial\sigma}{\partial p} = \frac{\mathrm{M}^2_p\mathrm{M}^1_n - \mathrm{M}^1_p\mathrm{M}^2_n}{|\mathrm{M}|} = \frac{C'(\sigma)\theta\lambda\sigma\left[\theta\lambda\tilde{u} - \theta\lambda\tilde{u}\right]}{C''(\sigma)C'(\sigma)[\theta\lambda\sigma]^2} = 0;$$

$$\frac{\partial n}{\partial p} = \frac{\mathrm{M}^1_p\mathrm{M}^2_\sigma - \mathrm{M}^2_p\mathrm{M}^1_\sigma}{|\mathrm{M}|} = \frac{\tilde{u}}{C'(\sigma)\sigma} > 0.$$

ii. Increasing the productivity of R&D:

$$\frac{\partial\sigma}{\partial\lambda} = \frac{\mathrm{M}^2_\lambda\mathrm{M}^1_n - \mathrm{M}^1_\lambda\mathrm{M}^2_n}{|\mathrm{M}|} = \frac{C'(\sigma)\theta\sigma\left[rC'(\sigma) - rC'(\sigma)\right]}{\lambda C''(\sigma)C'(\sigma)[\theta\lambda\sigma]^2} = 0;$$

$$\frac{\partial n}{\partial\lambda} = \frac{\mathrm{M}^1_\lambda\mathrm{M}^2_\sigma - \mathrm{M}^2_\lambda\mathrm{M}^1_\sigma}{|\mathrm{M}|} = \frac{r}{\theta\sigma\lambda^2} > 0.$$

iii. Increasing the fixed cost of setting up an R&D lab:

$$\frac{\partial\sigma}{\partial k} = \frac{\mathrm{M}^2_k\mathrm{M}^1_n - \mathrm{M}^1_k\mathrm{M}^2_n}{|\mathrm{M}|} = \frac{r}{C''(\sigma)\sigma} > 0;$$

$$\frac{\partial n}{\partial k} = \frac{\mathrm{M}^1_k\mathrm{M}^2_\sigma - \mathrm{M}^2_k\mathrm{M}^1_\sigma}{|\mathrm{M}|} = \frac{-r\left[C''(\sigma)\left[r + \theta\lambda(n+1)\sigma\right] + C'(\sigma)\theta\lambda n\right]}{C''(\sigma)C'(h)\theta\lambda\sigma^2} < 0.$$

The change in industry wide R&D is therefore

$$n\frac{\partial\sigma}{\partial k} + \sigma\frac{\partial n}{\partial k} = \frac{-r\left[r + \theta\lambda(n+1)\sigma\right]}{C'(h)\theta\lambda\sigma} < 0.$$

iv. Increasing the discount rate:

$$\frac{\partial\sigma}{\partial r}=\frac{\mathrm{M}_r^2\mathrm{M}_n^1-\mathrm{M}_r^1\mathrm{M}_n^2}{|\mathrm{M}|}=\frac{k}{C''(\sigma)\sigma}>0;$$

$$\frac{\partial n}{\partial k}=\frac{\mathrm{M}_r^1\mathrm{M}_\sigma^2-\mathrm{M}_r^2\mathrm{M}_\sigma^1}{|\mathrm{M}|}=-\left(\frac{C''(\sigma)\left[C'(\sigma)\sigma+\left[r+\theta\lambda(n+1)\sigma\right]k\right]+C'(\sigma)\theta\lambda\sigma k}{C''(\sigma)C'(h)\theta\lambda\sigma^2}\right)<0.$$

The change in industry wide R&D is therefore

$$n\frac{\partial\sigma}{\partial k}+\sigma\frac{\partial n}{\partial k}=-\left(\frac{C'(\sigma)\sigma+\left[r+\theta\lambda(n+1)\sigma\right]k}{C''(\sigma)C'(h)\theta\lambda\sigma^2}\right)<0.$$

Welfare - The social planner takes into account the permanent effect of an innovation, regardless of its patentability. Given the productivity and cost of R&D do not change over time, nor does the distribution of invention magnitudes, the planner has no incentive to choose a different R&D intensity or number of firms in different patent races. The welfare function is then

$$W=\underset{h,n}{Max}\left\{\frac{\lambda nhV^S-nC(h)}{r+\lambda nh}-nk\right\},\text{ where } V^S=\frac{p\tilde{u}(0)}{r}+\frac{\lambda nhV^S-nC(h)}{r+\lambda nh}.$$

This implies

$$W=\underset{h,n}{Max}\left\{\sum_{t=1}^{\infty}\left(\frac{\lambda nh}{r+\lambda nh}\right)^t\frac{p\tilde{u}(0)}{r}-\sum_{t=1}^{\infty}\left(\frac{\lambda nh}{r+\lambda nh}\right)^{t-1}\frac{nC(h)}{r+\lambda nh}-nk\right\},\text{ or}$$

$$W=\underset{h,n}{Max}\left\{\frac{n}{r}\left(\lambda h\frac{p\tilde{u}(0)}{r}-C(h)-rk\right)\right\}.$$

Proposition 4 - In the stationary symmetric equilibrium, differences in the standard of nonobviousness do not affect the R&D intensity of individual firms, but they do affect the number of firms actively engaged in R&D, and therefore the industry-wide rate of innovation.

Proof: We begin by calculating the derivatives

$$\mathrm{M}_s^1=\mathrm{M}_s^2=\lambda\left\{\frac{\partial\theta}{\partial s}\left[C'(\sigma)n\sigma+rk\right]-\frac{\partial\left(\theta\tilde{u}\right)}{\partial s}p\right\}.$$

Recall that $\theta=1-F(s)$ and $\theta\tilde{u}=\int_s^{\bar{u}}udF(u),$ which implies the preceding equation is simply which implies that $\mathrm{M}_s^1=\mathrm{M}_s^2=-\lambda f(s)\Psi(s),$ where

[A.7] $$\Psi(s) = \left\{\left[C'(\sigma)n\sigma + rk\right] - ps\right\}.$$

The comparative static calculations are thus

$$\frac{\partial\sigma}{\partial s} = \frac{\mathrm{M}_s^2\mathrm{M}_n^1 - \mathrm{M}_s^1\mathrm{M}_n^2}{|\mathrm{M}|} = \frac{f(s)\left[\Psi(s) - \Psi(s)\right]}{C''(\sigma)\theta\sigma} = 0;$$

$$\frac{\partial n}{\partial s} = \frac{\mathrm{M}_s^1\mathrm{M}_\sigma^2 - \mathrm{M}_s^2\mathrm{M}_\sigma^1}{|\mathrm{M}|} = \frac{f(s)\Psi(s)}{C'(h)\theta\sigma}.$$

The expression for $\Psi(s)$ used in the text is obtained by substituting for $C'(\sigma)\sigma$ using [A.5].

Proposition 5 - There exists a unique standard of nonobviousness, denoted s^*, such that in the interval $[0, s^*)$, industry wide R&D activity is strictly *increasing* in the standard of nonobviousness. In the interval $(s^*, \bar{u}]$, industry wide R&D activity is strictly decreasing in the standard of nonobviousness.

Proof: First, using [A.7] we check the slope of $\Psi(s)$:

$$\partial\Psi(s)/\partial s = \frac{\partial\sigma}{\partial s}\left[C''(\sigma)n\sigma + C'(\sigma)\sigma\right] + C'(\sigma)\sigma\frac{\partial n}{\partial s} - p = \frac{f(s)\Psi(s)}{\theta} - p.$$

Thus if there is a value $s^* \in [0, \bar{u}]$, where $\Psi(s^*) = 0$, we know that $\partial\Psi(s)/\partial s\big|_{s^*} = -p$. Thus there can be at most one extremum of $\Psi(s)$.

Next we check the values of $\Psi(s)$ as $s \to \bar{u}$ and $s \to 0$. These are evaluated most easily using the expression for $\Psi(s)$ used in the text: $\Psi(s) = \theta\lambda n\sigma\left[p\tilde{u} - rk\right]/\left[r + \theta\lambda n\sigma\right] - \left[ps - rk\right]$.

$$\mathrm{Lim}_{s\to\bar{u}}\Psi(s) = \left(\frac{0\cdot\lambda n(\bar{u})\sigma(\bar{u})}{r + 0\cdot\lambda n(\bar{u})\sigma(\bar{u})}\right)\left[p\bar{u} - rk\right] - \left[p\bar{u} - rk\right] < 0.$$

If we assume for the moment that the participation constraint is satisfied when $s = 0$ (i.e. $p\tilde{u}(0) - rk \geq 0$), the second limit is

$$\mathrm{Lim}_{s\to 0}\Psi(s) = \left(\frac{\lambda n(0)\sigma(0)}{r + \lambda n(0)\sigma(0)}\right)\left[p\tilde{u}(0) - rk\right] + rk > 0.$$

But it is possible that for a very weak standard of nonobviousness, the participation constraint [A.5] is violated (i.e. $p\tilde{u}(0) - rk < 0$). In that case $\Psi(s)$ does not exist at $s = 0$. Instead, define $\hat{s}$ s.t. $P\hat{s} - rk = 0$. Then $\Psi(\hat{s})$ exists and takes the sign

$$\Psi(\hat{s}) = \left(\frac{[1 - F(\hat{s})]\lambda n(\hat{s})\sigma(\hat{s})}{r + [1 - F(\hat{s})]\lambda n(\hat{s})\sigma(\hat{s})}\right)\left[p\tilde{u}(\hat{s}) - rk\right] - \left[p\hat{s} - rk\right] > 0,$$

because $-[p\hat{s}-rk]>0$. Existence of the extremum then follows from continuity of $\Psi(s)$ over $[\hat{s},\bar{u}]$.

Proposition 6 - The critical standard of nonobviousness s^*, is increasing in the equilibrium rate of innovation. This implies that in rapidly innovating industries, a smaller proportion of inventions can be protected without causing the rate of innovation to decline.

Proof: The critical standard of nonobviousness is defined by the equation $\Psi(s) = 0$. To examine how s^* varies with the industry wide rate of innovation, we compute comparative static derivatives with respect to the exogenous parameters explored in propositions 2 and 3:

[A.8] $$\frac{\partial s^*}{\partial z}=\frac{-\partial\Psi(s)}{\partial z}\Big/\frac{\partial\Psi(s)}{\partial s}=\frac{\partial\Psi(s)}{\partial z}\frac{1}{p},$$

where z is either p, λ, r, or k. Note also that

$$\frac{\partial\Psi(s)}{\partial z}=\left[C''(\sigma)\sigma+C'(\sigma)\right]n\frac{\partial\sigma}{\partial z}+C'(\sigma)\sigma\frac{\partial n}{\partial z}-\frac{\partial\left[ps-rk\right]}{\partial z}.$$

i. Higher output prices:

$$\frac{\partial s^*}{\partial p}=\frac{C'(\sigma)\sigma\tilde{u}(s^*)}{C'(\sigma)\sigma}-s^*=\tilde{u}(s^*)-s^*>0.$$

ii. More productive R&D:

$$\frac{\partial s^*}{\partial\lambda}=\frac{C'(\sigma)\sigma r}{\theta\sigma\lambda^2}-0=\frac{rC'(\sigma)}{\theta\lambda^2}>0.$$

iii. A higher discount rate:

$$\frac{\partial s^*}{\partial r}=\frac{\left[C''(\sigma)\sigma+C'(\sigma)\right]nk}{C''(\sigma)\sigma}-\frac{C''(\sigma)\left[C'(\sigma)\sigma+\left[r+\theta\lambda(n+1)\sigma\right]k\right]+C'(\sigma)\theta\lambda nk}{C''(\sigma)\theta\lambda\sigma}+k$$

$$\frac{\partial s^*}{\partial r}=k-\frac{C''(\sigma)\left[C'(\sigma)\sigma+\left[r+\theta\lambda\sigma\right]k\right]}{C''(\sigma)\theta\lambda\sigma}=-\left(\frac{C'(\sigma)\sigma+rk}{\theta\lambda\sigma}\right)<0.$$

iv. Higher fixed R&D costs:

$$\frac{\partial s^*}{\partial k}=\frac{\left[C''(\sigma)\sigma+C'(\sigma)\right]nr}{C''(\sigma)\sigma}-\frac{r\left\{C''(\sigma)\left[r+\theta\lambda(n+1)\sigma\right]+C'(\sigma)\theta\lambda n\right\}}{C''(\sigma)\theta\lambda\sigma}+r$$

$$=r-\frac{rC''(\sigma)\left[r+\theta\lambda\sigma\right]}{C''(\sigma)\theta\lambda\sigma}=\frac{-r^2}{\theta\lambda\sigma}<0.$$