Rewarding Sequential Innovators:
Prizes, Patents and Buyouts

Gerard Llobet

CEMFI

Hugo Hopenhayn

University of Rochester and Universidad Torcuato Di Tella

Matthew Mitchell

University of Minnesota and Federal Reserve Bank of Minneapolis

Abstract

This paper presents a model of cumulative innovation where firms are heterogeneous in their research ability. We study the optimal reward policy when the quality of the ideas and their subsequent development effort are private information. The optimal assignment of property rights must counterbalance the incentives of current and future innovators. The resulting mechanism resembles a menu of patents that have infinite duration and fixed scope, where the latter increases in the value of the idea. Finally, we provide a way to implement this patent menu by using a simple buyout scheme: The innovator commits at the outset to a price ceiling at which he will sell his rights to a future inventor. By paying a larger fee initially, a higher price ceiling is obtained. Any subsequent innovator must pay this price and purchase its own buyout fee contract.

1 Introduction

A central feature of innovative activity is that research is cumulative. This is relevant to the way in which research is rewarded. If research is rewarded through the granting of particular property rights, as for instance in a patent, the cumulative structure leads to the natural question of what to do when the next improvement arises. How can the property rights of the previous state-of-the-art be compatible with rewarding the recent improvement with its own property rights?

A variety of methods are available to reward innovators. Two of the most commonly discussed (in Wright (1983), for instance) are patents and research prizes. The latter is simply a transfer to the innovator for development of a particular invention. The former consists of the granting of some sort of market power to the innovator, perhaps in exchange for a fee. It is well understood that, when information is complete, it is optimal to choose a prize as the reward, since it does not result in any of the distortions that may accompany market power. When the principal charged with rewarding innovators does not have complete information about the benefits of an invention, however, it has been shown, for instance in Scotchmer (1999), that it may be optimal to grant a patent, since the value of the reward is then tied to the innovation's value through its potential profits in the market.

In this paper we study the optimal mechanism to reward innovations when new ideas arrive continually, and there is both moral hazard and adverse selection. We allow the patent office a variety of instruments, but force them to operate with limited information about the potentially patentable innovations. The optimal reward, it turns out, has a relatively simple form: it is a patent, with no statutory expiration date, but rather providing a constant amount of protection against future improvements forever. Under

1

very plausible conditions, the optimal patent policy involves different types of protection for different innovations. We show that the optimal mechanism can be instituted through a system of mandatory buyout fees which can be interpreted as compulsory licensing. This mechanism generates information which reduces the burden on courts of determining infringement. The menu of contracts offered by the patent office in order to implement the optimal reward is not time or history dependent, which is particularly appealing in terms of realistically using such a method.

When the value of an innovation is known, so that there is adverse selection but no moral hazard, the problem of competing property rights in a cumulative innovation setting may never arise. In that case, even if costs are unknown, it may be possible to provide incentive entirely through a simple cash prize. Then, the problem is not compounded by cumulative innovation, since the reward is paid in full immediately, and is therefore not relevant when the next improvement arrives. Scotchmer (1999) and Cornelli and Schankerman (1999), however, point to the usefulness of patents when the *value* of an innovation is unknown, resulting in a moral hazard problem. If the quality of the innovation is unobserved, the regulator cannot offer a reward which depends on quality. Therefore the regulator uses a patent. The monopolist pays a greater fee for a greater time period of patent protection.

When the model is extended to include multiple innovations, however, the optimal policy can change considerably. When value is unknown, so that patents are employed, the problem of competing property rights arises. A promise of patent rights to one innovator might be in conflict with offering another patent to a future innovator, and might discourage future innovators, as in O'Donoghue, et al. (1998). Lines of what constitutes a "sufficient" improvement to warrant a new patent must be drawn.

2

We characterize the optimal reward system in a cumulative innovation context with incomplete information on the part of the patent office. Like Scotchmer (1999), the patent office offers "better" patent protection in exchange for a higher fee. Here that protection means that a greater percentage of possible future innovations are precluded by the innovator's patent. As in other work such as O'Donoghue, et al. (1998) this ability of a patent to preclude future innovations is labeled the patent's "breadth." We use an extreme model where, if only one innovator is ever to arrive, the optimal patent policy implements the efficient level of research. We then show that in the same model, but with multiple innovators arriving in sequence with cumulative innovations, it is impossible to achieve the efficient level of research. This reinforces the idea that the cumulative nature of innovation is very relevant to policy.

By using a mechanism design approach, the problem of enforcement of patent breadth, a contentious one in practice, is studied alongside the optimal breadth itself. The optimal policy introduced here generates information about the quality of innovations. That is, all that the government must determine in order to dole out property rights is if an innovation is related, in the sense of being on the same "ladder" as the previous innovation. We require of the policy that it generates enough information to solve problems of "infringement" through the self-selection from a menu of patents. This is in contrast to current policy, where the courts must determine more than just if two innovations are quality improvements over the other, but also how much of an improvement has been made. In the mechanism studied here, infringement is determined by the reports of the innovators, lowering the burden on the patent system.[1]

---

[1]See Llobet (1999) for a study on how the legal environment affects cumulative research.

An interesting feature of the optimal policy is that each patent offered consists of a given breadth, maintained forever. That is, the patent expires only when something better than the constant threshold comes along; the optimal policy does not prescribe that the amount of protection increase or decrease over time. Current patent policy has a clear sense in which level of protection declines suddenly, at the end of the patent's statutory life. The optimal policy here suggests that patents should end only because something better arrives, and not because of some imposition of a statutory time limit for the protection.

We also show that, under plausible conditions, the menu offers a variety of patent breadths to different innovators, depending on their costs and the resulting quality of their innovations. Bigger improvements get greater protection. Interestingly, it has been claimed that, in fact, the patent courts do provide additional protection for products which represent large improvements. This result also shows that it may be optimal to provide for a variety of breadths, which the current US statute does not explicitly allow for.

Given that the optimal policy calls for a variety of patent breadths, it might seem that implementing the optimal policy would require a very complicated system. To the contrary, we show that the optimal mechanism can be implemented through mandatory buyout prices or licensing fees. The idea of mandatory licensing dates back at least as far as the 1800's, when such a rule passed the House of Lords (Machlup and Penrose(1950)). Innovators, as part of the granting of a patent, must commit to a price at which they will relinquish their rights. This commitment mitigates any bargaining power a patent holder might exert on future innovators. Tandon (1982) uses similar buyout fees in a complete information model to mitigate monopoly costs. Here we add the fact that, by offering a menu of compulsory licensing fees

4

to innovators, the fees can be useful in generating information about the innovators. The fee acts as the breadth of the patent: the bigger the fee, the greater the patent's implied breadth, since future innovators will need a substantial improvement to justify the larger buyout fee.

The set of patents, then, offered under the compulsory licensing agreement is a menu of buyout fees, accompanied with prices paid to the government. In order to take the lead, an innovator must pay the prearranged buyout amount to the prior innovator, choose its own buyout fee, and make the appropriate payment to the government as prescribed by the menu. The menu of contracts offered by the government has a simple, stationary form. The government has a constant set of posted prices for patents with various buyout amounts, and innovators choose their favorite whenever they want to take the market lead.

Other authors have studied the trade-off between patents and prizes. Notably, Wright (1983) argues that prizes may mitigate problems associated with patent races. In our formulation, with ideas private to a single innovator, this argument for prizes is not present. Shavell and van Ypersele (1999) suggest offering an optional reward, so that some patents would be replaced with rewards. They do not consider the possibility that this might lead to adverse selection when the quality of innovations is unobservable and endogenous. In single innovation formulations such as Scotchmer (1999), it is true that the optimal reward is a prize if the value is known. The single innovation case with asymmetric information and general reward mechanisms is considered in Chiesa and Denicolo (1999), with similar results.

Previous models of optimal patent policy have for the most part not considered cumulative research. Exceptions include Scotchmer and Green (1990), Green and Scotchmer (1995), and O'Donoghue, et al. (1998). The

5

model employed in the latter is most similar to the one here; a central difference is that in their model the patent office is fully informed and therefore could offer a prize to reward innovators. We consider a case where research is cumulative and the patent office's information is incomplete, and show that it differs from the optimal mechanism in the one innovation case, in the sense that it departs from the logic of the regulation literature. On the other hand, they consider heterogeneity on the part of the consumers as well as a form of bargaining between innovators. They consider both infinite length, finite breadth patents and infinite breadth, finite length patents. We consider the set of all possible breadths and lengths (in the sense that length can be thought of as breadth reduced to zero after a certain time), and find that infinitely lived patents are optimal.

The next section sets the stage by studying a version of the model where there will only ever be one innovator. When costs are unknown but value of the project is observable, the reward for an innovator is a cash prize. When value is unobservable, patents are employed as a reward. In the model below, the efficient level can be implemented in either case. These results mirror the findings of several papers in the patent literature and are similar to the spirit of mechanisms used in the regulation literature. The main points come in the third section. There, the general model of many innovations is introduced and the optimal policy is contrasted with the one for the single innovator. The patent agency must concern itself with how to provide a reward for one innovator without discouraging future innovators.

2 The Environment

2.1 Preferences and Technology

There is a single good differentiated by quality q and an infinite horizon of discrete time periods. A product of quality zero is sold competitively at marginal cost, normalized to zero. There is a single, infinitely lived consumer with time-additively separable preferences and per period utility $q - p$, where p is the price of the good. These preferences are standard as in, for instance, Anderson, et al. (1992), and can be justified by adding a composite good and quasilinear utility. The future is discounted according to a discount factor δ.

An "idea" is the private property of a single firm. It must be researched, though, to be made into a viable product. That product can be freely imitated unless the innovator is given some specific property rights (a patent). Suppose that firms are indexed by their capabilities to undertake research through a parameter z. We assume z is drawn from a known distribution $\Phi(z)$ with density $\phi(z)$. A higher z means that the firm can obtain improvements at a lower cost.[2] The cost of the improvement is a function not only of z but also of the size Δ of the improvement (in the quality space) over the state of the art, in this case $q = 0$, according to $c(\Delta, z)$. It is assumed that $c_1 > 0$, $c_2 < 0$, $c_{11} > 0$ and $c_{12} < 0$. The first assumption says that bigger improvements are more costly. The second says that the higher is the firm's z, the lower are its costs. Costs are convex. The last assumption is important: the higher is z, the *lower* are marginal costs c_1. Therefore the social planner prefers that firms which draw high z spend more on research. The higher is z, the more efficient is the firm at research.

We will, at various points, consider two market structures. The first is

---

[2]Alternative interpretations of z could be: quality of the idea that the firm obtains, experience obtained by marketing similar products, scale economies, know-how, etc.

simply competition, where $p = 0$ is the result of marginal cost pricing. In that case, consumers receive Δ units of surplus per period. On the other hand, if the innovator is given a monopoly right to the product, the product is sold for $p = \Delta$, the consumers are left with no surplus, and the firm makes Δ units of profits per period.

The size of the innovation Δ is also the amount of social surplus it generates each period. If the product is sold competitively at marginal cost 0, then the homogeneous consumer enjoys Δ units of surplus. If the product is sold at a higher price, profits rise one-for-one with lost consumer surplus. The First Best allocation that a fully informed social planner would choose $\Delta^*(z)$ solves

$$\Delta^*(z) = \arg \max_{\Delta} \frac{\Delta}{1 - \delta} - c(\Delta, z).$$

The implicit function theorem can be applied to show that $\frac{\partial \Delta^*}{\partial z} = -\frac{c_{12}}{c_{11}} > 0$. In general, not all ideas will generate an increase in welfare that justifies the cost. There will be a threshold value $\underline{z}^*$ so that only $z > \underline{z}^*$ will be worthwhile. This minimum size is defined by

$$\frac{\Delta^*(\underline{z}^*)}{1 - \delta} - c(\Delta^*(\underline{z}^*), \underline{z}^*) = 0.$$

2.2   A Single Innovation

In order to set the stage for the model with cumulative innovations, it is useful to start by addressing how to reward an innovation if it is the only innovation that will ever arise.

First, suppose that the regulator can observe q, but cannot tell what the costs of innovation for the good were. That is, $z \in R_+$ is private information to the innovator.

What can the regulator do to encourage innovation in this case? The patent office can reward the innovator with a fee $F(z)$ (which may be nega-

tive, a prize), or with a patent lasting $T(z)$ periods. A firm with improvement Δ sells it for price $p = \Delta$ during the patent term. As in Baron and Myerson's (1982) monopoly regulation problem, it is sufficient here for the government to encourage innovators to choose the efficient level of spending through a transfer program, which might be thought of as a prize system. Let $T = 0$, consider the following reward:

$$
F(z) = \begin{cases} 0 & z < \underline{z}^* \\ -c(\Delta^*(\underline{z}^*), \underline{z}^*) & z = \underline{z}^* \\ F(\underline{z}^*) - \int_0^z \frac{\partial \Delta^*}{\partial i} c_1(\Delta^*(i), i) di = \frac{\Delta^*(z)}{1-\delta} & z > \underline{z}^*. \end{cases} \tag{1}
$$

Since F is negative, this corresponds to a prize. The nature of the optimal prize is identical to the one Baron and Myerson (1982) use to regulate a monopolist with an unknown cost. For any given z, the regulator knows the efficient Δ^*. For the lowest type, the prize offered exactly offsets research costs, $F(\underline{z}^*) = -c(\Delta^*(\underline{z}^*), \underline{z}^*)$. For a higher type z, which has lower marginal cost, it offers a larger prize in return for more innovation. This looks attractive to a low marginal cost firm, but not so to a high marginal cost firm. It is easy to verify that $F(z)$ leads to $\Delta^*(z)$ being an optimal innovation choice.

Since our formulation disregards the static costs of monopoly, it would also be possible for the government to support the first best with patents of varying lengths, increasing in Δ. We focus here on the pure-transfer system because it minimizes monopoly power (there is none; innovations are bought for $F(z)$ and provision is competitive).

In reality, the true quality of an innovation is hard to ascertain, and therefore a prize system may be difficult to implement effectively. As a result, Scotchmer (1999) shows that it may be useful to offer patents, since the value of a patent is tied to the value of the innovation. The solution is

similar to the one employed by Lewis and Sappington (1988a, 1988b). There, when regulating a monopolist, the planner offers to allow a higher price in exchange for payment of a fee. Only a monopolist with high demand will find it worthwhile to pay the fee. A patent works like the price: a longer patent is valuable the greater is quality (demand).

Consider the model presented above, but where Δ is unobservable. A prize structure like the one in (1) will not be effective, since everyone will claim to have the highest z to collect the largest prize, and there will be no way to verify if the high level of research demanded for that z has actually been undertaken. However, since the monopolist can extract all the surplus from the consumers, the efficient outcome can be attained, but only through the granting of monopoly power. In fact, if $T(z) = \infty$, it is easy to verify that the innovator chooses $\Delta^*(z)$, since the profits of the firm are exactly equal to the social surplus, and so profit maximization and the definition of Δ^* coincide.

Because of the inelastic demand, even unknown Δ may not prevent the government from implementing the efficient level of innovation. However, that result can only be attained through the granting of patent rights, as opposed to the earlier case with known Δ, which could be implemented entirely with cash payments. This is the important point of this section. In the next section, we take up the main point of the paper: what happens when innovation is cumulative. When patents are essential (i.e. when Δ is unobserved), there will come the question of how to weigh the property rights of one innovator with that of the next. One of the important differences in the next section is that the optimal policy will be different and will be unable to attain the efficient level of research due to the sequential structure.

3 Multiple Innovations

Suppose that each period a new firm arrives with a new idea z, allowing an improvement over the current quality in the amount Δ with cost $c(\Delta, z)$ as before. If Δ is observable, the government could simply employ the prize system of the last section in sequence to each innovator. The conclusions of the last section would remain unaltered.

When Δ is unobserved, the optimal mechanism for a single innovator involved an infinite patent. Offering a patent, however, is quite different when innovation is cumulative, as pointed out, for instance, in Scotchmer and Green (1990) and O'Donoghue, et al. (1998). The optimal patent for the sequential case must take into account the fact that property rights granted to the first innovator might preclude some future improvements. The patent, in order to be economically meaningful, may have to preclude small improvements, or else the original innovator's ability to profit from research may be short lived.

If ex ante contracting were perfect, the optimal policy would remain unchanged: simply grant an infinite property right to the first innovator, and allow him to license that patent to future innovators. Since there is no static cost of monopoly, the licensing would allocate innovation efficiently.

Given the history of innovations and their rewards, though, it is likely that disputes would emerge. For instance, there may be a *hold-up* problem. If the patentee cannot commit beforehand to sell a license to an infringer before the invention is obtained, then once the development cost is sunk, the patentee can raise the price for the license. This discourages investment in innovations. Moreover, there may be disputes about what innovations pertain to which products being sold, since one product may contain many distinct innovations. These problems lead to potential dynamic costs of monopoly

rights, which is our focus here, as in O'Donoghue et al. (1998).

We assume that the arbiter of the disputes, which we take to be the government, has no way to learn anything truthful about the dispute from the parties involved. As such, all that the arbiter can do is assign property rights based on the reports of z that are made by the innovators. Moreover, the arbiter can make a determination about this mechanism at the outset, as if "owning" the entire ladder; no useful information can be generated from the parties. Here, the role of the government is one of commitment and independence: it allocates rewards only on the basis of reports, according to a policy that is fixed ex ante.

Of course, our assumption that there is no way for the patentor to learn about the innovations from the innovators is extreme. Llobet (1999) studies a simpler model but where the litigation technology is meaningful. Here, if some perhaps noisy signal about z or Δ could be generated, it would simply result in the arbiter using a combination of the mechanism described here and a reward based on the signal of the underlying information, using the signal more if it is more precise.

In principle, the government can choose the market structure, in terms of which innovations may be sold and whether or not monopoly rights are granted, at any point in time, along with transfer payments (or fees) for all the innovators that have arrived up to that time. The set of possible instruments we allow, then, is very large. Fortunately, we find a relatively simple characterization of the optimal mechanism, and then focus our attention on studying that recursive problem.

From the innovator's perspective, all that matters is how long monopoly power will last for the innovation. The innovation can make profits Δ in any period that it is the highest quality product allowed to be sold, and

12

zero otherwise.[3] Denote the expected discounted duration of the monopoly power granted under the patent for type z by $d(z)$. Since a unit of d gives the innovator the right to make Δ profits for one immediate period, the innovator's expected revenue from sales of the product are $d(z)\Delta$, the product of $d(z)$, the expected discounted duration of monopoly power, and Δ, the per period profits of producing the innovation. The innovator solves

$$\Delta^d(z, d) = \arg \max_{\Delta} \ d(z)\Delta - c(\Delta, z) - F(z). \tag{2}$$

We make one additional assumption which, as can be seen in Appendix A, simplifies matters considerably. It is that $\frac{\partial^2 \Delta^d}{\partial^2 d} > 0$. With the assumption, under the optimal mechanism, a new innovation is either implemented and given the right to produce immediately or is never given the right to produce.[4]

Proposition 1 *The optimal mechanism gives each innovator a duration of incumbency $d(z)$, in such a way that whenever a new innovator is awarded a $d(z) > 0$ the previous one is immediately replaced.*

This allows the state variable of the dynamic program to consist simply of information from the current market leader. Without it, the analysis would be similar but the state variable would have to include all of the promised future monopoly rights.[5]

The previous proposition says nothing about how duration should be allocated; in fact, since $d(z)$ is expected duration, the property right might be allocated stochastically. The next proposition shows how duration is

---

[3]This assumes that monopoly power for the leading innovation does not preclude earlier innovations from being sold. That assumption is not essential; the details are below.

[4]We require that the mechanism be immune to simple bribes across agents. The details are contained in Appendix A.

[5]Other questions would arise if innovations were encouraged but forced to wait; for instance, do subsequent innovators build on the state-of-the-art invention (patent) or the most advanced product being sold in the marketplace?

allocated: a patent can be defined by its "breadth" $\underline{z}'(z)$. That is, a firm reporting that it is of type z receives sole rights until a report of $\underline{z}'$ is made by another innovator.

Proposition 2 *Given a duration $d(z)$, it is optimal to grant a constant breadth, infinitely lived patent $\underline{z}'(z)$.*

In addition to making the analysis substantially simpler, this result points to an important intuition about optimal property rights for repeated innovators. Take, for instance, the example of a fixed breadth $\underline{z}$ for T periods followed by zero breadth after that, which could be thought of as current US policy. If a longer patent is offered, less breadth needs to be offered in each period. The extra protection being extended through a longer time is given for the smallest z, whereas the breadth it reduces comes, at the margin, from large z near $\underline{z}$. The patentee cares only about the cumulative probability of being a leader, but the patentor cares about the size of the next innovation which is allowed. Therefore, this switch from protection against high z now to lower z later can improve social welfare, while at the same time maintaining the expected profits to the innovator. This force pushing for longer, lower breadth patents in the cumulative context is not an artifact of any particular modeling assumption, but rather a natural result of the different aims of patentor and patentee. This complements the result in Gilbert and Shapiro (1990) which suggests that, in terms of static cost of monopoly, long lived patents may be best.

This formulation assumes that when a new innovation arrives, older ones may still be produced; in equilibrium of the pricing game they are not sold in positive amounts, though, although they do have the effect of limiting profits to the incremental quality Δ rather than the entire quality of the produced good. This assumption is not material; all of the results remain

14

unchanged if a patent allows the holder to dominate the entire industry with no fear of competition. The model is formulated as it is, though, to capture the idea that although legally still a patented product and fit for sale, the effective life comes to an end because something else comes along and makes the innovation effectively obsolete.[6]

It remains to be decided how to allocate duration for different types of z; the rest of the section is concerned with that problem. This mirrors the problem above: in the one innovation case, it is optimal to provide full protection, i.e. duration of $\sum_{i=0}^{\infty} \delta^i$. With multiple innovations there is the additional trade-off that larger duration means excluding more future ideas. In order to figure out the optimal $\underline{z}'(z)$, and hence the optimal duration, first an inventor's response, in terms of innovation, to a given breadth $\underline{z}'$ must be calculated. From equation (2) a given firm chooses Δ to solve

$$\Delta^p(z, \underline{z}') = \arg\max_{\Delta} \frac{\Delta}{1 - \delta\Phi(\underline{z}')} - c(\Delta, z) - F(z),$$

where $\frac{1}{1-\delta\Phi(\underline{z}')}$ reflects the expected duration given an infinitely lived patent of breadth $\underline{z}'$. The corresponding first order condition is

$$\frac{1}{1 - \delta\Phi(\underline{z}')} = c_1(\Delta^p, z) \tag{3}$$

where again Δ^p denotes the optimal choice made by the firm. Due to the assumptions on the cost function, it is immediate that the function $\Delta^p(z, \underline{z}')$ is increasing in both its arguments. That is,

$$\frac{\partial \Delta^p}{\partial z} = \frac{c_{12}(\Delta^p, z)}{-c_{11}(\Delta^p, z)} > 0, \tag{4}$$

$$\frac{\partial \Delta^p}{\partial \underline{z}} = \frac{\frac{-\delta\phi(\underline{z}')}{(1-\delta\Phi(\underline{z}'))^2}}{-c_{11}(\Delta^p, z)} > 0. \tag{5}$$

---

[6]In O'Donoghue, et al. (1998), the consumer side is set up so that the two best products are sold in positive quantities, thereby making the obsolescence gradual.

The more breadth is granted, the more innovation will be undertaken, since it is likely to pay off for a longer effective patent life. Of course, the cost is that breadth precludes future innovations that might be worthwhile.

Although we consider breadth in terms of the quality of the idea z, it is equivalent to think of breadth in terms of the size of the innovation Δ required for a noninfringing innovation since there is a strictly increasing relationship between the size of the invention achieved, Δ, and the parameter z. Breadth can be thought of in the usual sense: a higher $\underline{z}'$ means that it will take a larger improvement for a subsequent product to be produced.

Note that, as before, the first best requires that

$$\frac{1}{1-\delta} = c_1(\Delta^*, z).$$

If anything less than an infinite breadth patent is granted, innovation will not be at the efficient level. But precluding future innovations will not, in general, be optimal; patents will be weakened to allow for subsequent innovations at the cost of inefficient levels of research for each innovation.

The condition for truthful revelation given the two instruments, $\underline{z}'$ and F, is for the report $\hat{z}$ to be the one that solves

$$W(z, \underline{z}') = \max_{\hat{z}} \frac{\Delta^p(z, \underline{z}'(\hat{z}))}{1 - \delta\Phi(\underline{z}'(\hat{z}))} - c(\Delta^p(z, \underline{z}'(\hat{z})), z) - F(\hat{z}).$$

Checking the Spence-Mirrlees condition we obtain

$$\frac{\partial^2 W}{\partial \underline{z}' \partial z} = -\delta\phi(\underline{z}')c_1(\Delta, z)\frac{c_{12}(\Delta^p, z)}{c_{11}(\Delta^p, z)} > 0.$$

For the lowest type $\underline{z}$ which is undertaken, profits are minimized to zero, i.e. $W(\underline{z}, \underline{z}'(\underline{z})) = 0$, and every other type z receives positive profits as a result. The expression for $F(z, \underline{z})$ is obtained in the standard way as follows:

$$F(z, \underline{z}) = \frac{\Delta^p(z, \underline{z}')}{1 - \delta\Phi(\underline{z}')} - c(\Delta^p(z, \underline{z}'), z) + \int_{\underline{z}}^{z} c_2(\Delta^p(x, \underline{z}'(x)), x)dx.$$

16

Because $\frac{\partial W}{\partial \underline{z}'} = \frac{\delta \phi(\underline{z}(\hat{z}))\Delta^p(z,\underline{z})}{\left(1-\delta\Phi(\underline{z}(\hat{z}))\right)^2} > 0$ it must be that F is strictly increasing in z, so that better inventors obtain more protection at a higher price.

The mechanism designer, then, must choose a protection level $\underline{z}'(z)$ for any level z that might arise. The problem has a relatively simple recursive structure. The dynamic problem of the principal is

$$V(q,\underline{z}) = \Phi(\underline{z})\left[q+\delta V(q,\underline{z})\right] + \tag{6}$$
$$\max_{\underline{z}'(z)} \int_{\underline{z}} \left[(q+\Delta^p(z,\underline{z}')) - c(\Delta^p(z,\underline{z}'),z) + \delta V^p(q+\Delta^p(z,\underline{z}'),\underline{z}')\right]\phi(z)dz$$

subject to the constraints

$$\text{(IC)} \quad \underline{z}'(z) \text{ increasing in } z \quad \text{for all } z > \underline{z},$$
$$\text{(IR)} \quad W(z,\underline{z}'(z)) \geq 0 \quad \text{for all } z > \underline{z}.$$

The state of the economy is the leading edge quality q and the promised breadth $\underline{z}$. The first term reflects the fact that, if an idea comes along less than $\underline{z}$ (which happens with probability $\Phi(\underline{z})$), consumer plus producer surplus equals q, since no improvement can be allowed to arrive, and the state is unchanged. If an idea of type greater than $\underline{z}$ arrives, though, the patent authority can offer a new patent with breadth $\underline{z}'$. That encourages the innovator to generate an invention of size $\Delta^p(z,\underline{z}')$, leaving the leading edge quality at $q + \Delta^p(z,\underline{z}')$. The formulation assumes that funds are costlessly obtainable by the government; it is completely straightforward to add a cost μ of acquiring funds, as in Laffont and Tirole (1986). None of the results presented are adversely affected by the inclusion of such a cost.

The two restrictions correspond to the incentive compatibility and individual rationality constraints. The second one implies that the planner will fulfill the promise made to the previous innovator despite the private information on the quality of the next idea.

A natural question is whether the optimal contract involves a uniform patent or if the principal will be interested in separating different firms with different capabilities. The answer is that the optimal patent contract calls for greater breadth for higher z if offering more breadth has a greater effect on research for innovators with better ideas, i.e. $\frac{\partial^2 \Delta^p}{\partial \underline{z}' \partial z} > 0$.[7]

Proposition 3 *If $\frac{\partial^2 \Delta^p}{\partial \underline{z}' \partial z} > 0$, the optimal mechanism satisfies $\underline{z}'(z)$ strictly increasing in z.*

The intuition here is straightforward. The cost of higher breadth in terms of lost future projects does not depend on z, but the marginal benefit is increasing if $\frac{\partial^2 \Delta^p}{\partial \underline{z}' \partial z} > 0$, since it has more of an effect on incentives to innovate.

There is some evidence that courts follow something like this rule. The most common way to invalidate a patent is to show the courts that it is not a very "big" improvement. In such cases, the patent may be invalidated (i.e. zero breadth), or it may be that it is quite easy for other products to be produced. Allison and Lemley (1998) study a sample of 299 patents litigated in 239 cases. These represent all the suits in the period 1989-1996 started by competitors in order to invalidate existing patents. They find that the most argued reason to limit the original innovator's property right is the obviousness of the patented invention, used in 42% of the cases. In this model, this idea is captured by the size of the innovation, Δ. When $\underline{z}'(z)$ is strictly increasing in z, small improvements get less protection, while larger inventions get more. This additional protection, is, of course, costly. The proof of the prior proposition, together with the first order condition, implies that $\frac{\partial V}{\partial \underline{z}'}(q + \Delta(z, \underline{z}'), \underline{z}') < 0$, so promises of breadth by the patent office decrease future prospects.

---

[7]This condition of Δ^p amounts to a condition on a third order cross derivative of c. An example of a simple function satisfying this assumption follows.

An important feature of the optimal mechanism is that $\underline{z}'(z)$ does not depend on the previous threshold, and the interaction comes exclusively through $F(z, \underline{z})$. The reason this is important is that it means that, in practice, the decision to award a new innovator market leadership depends only on the report of the current leader and the current innovator. In particular, F is increasing in $\underline{z}$ in such a way that the higher was the previous threshold, the lower will be profits for the subsequent innovator, implementing the higher threshold $\underline{z}'$.

One might imagine what cost function might satisfy the assumption made on Δ^p. The following demonstrates one.

Example 1 *Consider the cost function $c(\Delta, z) = \frac{\Delta^{\alpha}}{z} + k$ with $\alpha > 1$ and $k \geq 0$. In this case,*

$$\Delta^p = \left[\frac{z}{\alpha(1 - \delta\Phi(\underline{z}'))} \right]^{\frac{1}{\alpha - 1}}.$$

Clearly,

$$\frac{\partial \Delta^p}{\partial z} = \frac{1}{\alpha - 1} z^{\frac{2 - \alpha}{\alpha - 1}} \left[\frac{1}{\alpha(1 - \delta\Phi(\underline{z}'))} \right]^{\frac{1}{\alpha - 1}} > 0,$$

$$\frac{\partial \Delta^p}{\partial \underline{z}'} = \frac{1}{\alpha - 1} \frac{z^{\frac{1}{\alpha - 1}}}{\alpha^{\frac{1}{\alpha - 1}}} \left[\frac{1}{1 - \delta\Phi(\underline{z}')} \right]^{\frac{2 - \alpha}{\alpha - 1}} \frac{\delta\phi(\underline{z}')}{1 - \delta\Phi(\underline{z}')} > 0,$$

and we can compute,

$$\frac{\partial^2 \Delta^p}{\partial z \partial \underline{z}'} = \left(\frac{1}{\alpha - 1} \right)^2 \frac{1}{\alpha^{\frac{1}{\alpha - 1}}} \left[\frac{z}{1 - \delta\Phi(\underline{z}')} \right]^{\frac{2 - \alpha}{\alpha - 1}} \frac{\delta\phi(\underline{z}')}{1 - \delta\Phi(\underline{z}')} > 0.$$

Finally, when Δ is not observable, less inventions will be implemented and they will have a smaller value.

Proposition 4 *The optimal mechanism when Δ is not observable has $\Delta^p(z) \leq \Delta^*(z)$ and $\underline{z}'(z) > \underline{z}^*$ for all z.*

To show that the results are independent of the market structure chosen consider for example the case where the producer of the last quality has control over the whole ladder. Per period profits in this case will be equal to $q + \Delta$, and therefore, the value of the innovation will correspond to,

$$W(z, \underline{z}') = \max_{\widehat{z}} \frac{q + \Delta^p(z, \underline{z}'(\widehat{z}))}{1 - \delta\Phi(\underline{z}'(\widehat{z}))} - c(\Delta^p(z, \underline{z}'(\widehat{z})), z) - F(\widehat{z}).$$

We can define in this case a new fee $\widetilde{F}(\widehat{z}) = F(\widehat{z}) + \frac{q}{1 - \delta\Phi(\underline{z}'(\widehat{z}))}$, and all the previous results follow for the contract $\left\{\underline{z}'(z), \widetilde{F}(z)\right\}$.

4 Decentralization through Compulsory Licensing

When patent breadth is studied in situations with complete information, there is no discussion about how to enforce the resulting optimal policy. Implicitly, it is assumed that the courts can be used to implement the appropriate breadth. In practice, though, the definition of breadth in the patent statute tends to be vague. Given that vast resources are spent in patent litigation, it seems that enforcement is not at all trivial. The mechanism design approach implicitly considers enforcement, since it induces agents to truthfully report z and choose the associated $\Delta(z)$. That is, an advantage of the proposal offered above is that the mechanism itself generates information about the quality of the innovations. If two products are on the same ladder, the report of the type z is sufficient to determine infringement and future protection. A system like the one in place in the United States requires the courts to determine precise infringement; that is, the courts must determine what qualities have arisen.

So far we have been using two instruments $\{\underline{z}, F\}$ to reward innovators. A natural question is, how do you implement a system where the market

leader is determined by something as seemingly nebulous as z. This seems particularly difficult in light of the fact that innovators which report a different z might get different protection. Fortunately, there is a mechanism to achieve the same allocations, one which has straightforward real-world analogues. This is important because an important message is that multiple patent breadths might be optimal.

A way to implement the optimal thresholds is through what we call a compulsory licensing or buyout fee mechanism. A firm with the invention Δ has to pay an amount τ to purchase the previous innovation and with it the right to produce. At the same time, purchasing a patent from the government with a price σ' the firm guarantees that anybody in the future wishing to produce will have to buy his invention for an amount τ'. We next focus on whether a compulsory licensing mechanism of this sort can implement the same kind of allocations as the original one.

The protection for an innovator comes from buyout fees $\{\tau'(z), \sigma'(z)\}$. Notice that the menu of contracts available to the innovator to protect *his* innovation does *not* depend on the type of patent in place, but only on the report of z. That is, the optimal contract is simply a set of guaranteed buyout fees $\tau'(z)$ offered at a fixed set of prices $\sigma'(z)$. Anyone wishing to sell a product must simply pay the existing fee τ to the current market leader, choose a contract, and pay σ' to the government. You can think of the innovator as choosing the fee $\{\tau', \sigma'\}$, without having to concern himself with anything like z.

Only the payoff to the prior innovator depends on the past. The set of contracts is not history dependent, which makes them particularly simple to set up, but potentially restrictive in what they can implement. We show that the optimal patent menu from the prior section can, in fact, be decentralized

with these simple sort of fees.

Denote the minimum entrant under those fees by $\underline{z}_\tau(z)$. A firm with capability z facing an existing innovation with protection $\{\tau, \sigma\}$ will obtain profits according to

$$W^\tau(z, \tau', \tau) = \max_{\widehat{z}} \frac{\Delta^p}{1 - \delta\Phi(\underline{z}_\tau(\widehat{z}))} + \frac{1 - \Phi(\underline{z}_\tau(\widehat{z}))}{1 - \delta\Phi(\underline{z}_\tau(\widehat{z}))}\delta\tau'(\widehat{z}) - c(\Delta^p, z) - \tau - \sigma'.$$

The first term reflects discounted expected profits from sales. The second is discounted expected receipt of the buyout fee τ'. Profits, then, are net of costs c and fees τ to the prior innovator and σ' to the government. Given $\underline{z}_\tau(\widehat{z})$, notice that τ, σ, and τ' do not affect the decision of which Δ^p to choose. Therefore, this mechanism will induce the same level of invention if $\underline{z}_\tau(z)$ is equivalent to the $\underline{z}'(z)$ obtained using the other mechanism. The following results show that a compulsory licensing contract can implement the mechanism described in the prior section.

Lemma 5 *The compulsory licensing contract $\{\tau'(z), \sigma'(z)\}$ is implementable and independent of τ if*

1. *$\tau'(z)$ is non-decreasing in z, and*

2. *$\sigma'(z)$ is obtained according to the following expression:*

$$\sigma'(z) = \frac{\Delta^p(z, \underline{z}'_\tau(z))}{1 - \delta\Phi(\underline{z}'_\tau)} + \frac{1 - \Phi(\underline{z}'_\tau)}{1 - \delta\Phi(\underline{z}'_\tau)}\delta\tau' + \int_{\underline{z}^0_\tau}^z \frac{\partial c}{\partial z}(\Delta^p(x, \underline{z}'_\tau(x)), x)dx$$

where $W^\tau(\underline{z}^0_\tau, \tau', 0) = 0$.

Lemma 1 is more restrictive than the usual Spence-Mirrlees condition. We focus on contracts that are not only incentive compatible but also independent of the previous τ. Still, this reduced family of contracts is enough to implement the second best.

Proposition 6 *Any incentive compatible contract $\{\underline{z}'(z), F(z)\}$ can be achieved using a compulsory licensing contract $\{\tau'(z), \sigma'(z)\}$ independent of τ. That is, $\underline{z}'(z) = \underline{z}'_\tau(z)$ for all z. Moreover, $\tau'(z)$ is non-decreasing in z.*

The implication of this is that contracts can be simply offered, at a price σ, by the government, which state that the holder of the contract has exclusive rights to produce on that ladder until such time as another innovator pays τ and a new price σ'. Better ideas (higher z) receive more protection through a higher buyout fee τ: they are protected more by the fact that the next innovator will have to pay more in order to produce.

The fact that the future contract is independent of the existing buyout τ makes this mechanism especially appealing, since the planner needs to gather little information on the current leader to offer future contracts, and stationary contracts can be offered.[8]

These compulsory licensing fees have a natural economic interpretation. As emphasized before, the inefficiency that sequential innovation generates is related to the temporary returns that an innovator obtains with respect to the permanent increase in social welfare created. These buyouts can be understood as a way to internalize this cost. When a future firm pays the prespecified buyout it is partially compensating the incumbent for being replaced. Therefore, the entrant is taking into account the externality created, and he will decide according to it. He will enter if his contribution is bigger than the foregone profits for the incumbent.

5 Summary

The patent statute involves a single, somewhat vague definition of breadth for all innovations, and leaves most of the job of deciding property rights to the courts. In constructing rewards for innovators, however, it is possible to generate information through the self-selection of patent protection. In fact,

---

[8]Of course, non-stationary contracts that implement the same allocation can be easily constructed.

we show here how patent breadth can be implemented through a system of compulsory licensing, replacing much of the burden now placed on the courts. This is especially important if one wishes to offer different patent breadth to different inventions. It is not surprising that in light of the heterogeneity of inventions, the optimal reward policy may reward different innovations with different breadth.

In this paper, the optimal policy is characterized, and it is found that in many cases such differentiation is optimal. This is an important practical point: it may be both optimal as well as feasible to offer multiple patent breadths. In addition, the optimal policy has patents of infinite statutory duration. The sort of patent policy described here would involve a more complicated set of patents offered, but a less complicated system to determine infringement, since the choice of patent would determine who had a right to produce.

A feature of the optimal policy is that patents have infinite statutory duration. They expire, effectively, only when something suitably better comes along. This results because it is always better to transfer profits to the leader in a state of the world when some amount of time has passed, but nothing good enough to supplant it has come along, rather than by giving extra breadth precluding a useful innovation. Whereas the patentee cares only about the probability of being the leader, the patentor cares about the size of the innovation when the new leader comes along.

This is important because it differs from the lessons of the one innovation case, which has been studied in slightly different environments. The cumulative research formulation used here suggests, as in other papers such as O'Donoghue et al. (1998), that breadth is a central part of the definition of a patent when further innovations will arrive. It may be more difficult

to achieve efficient research outcomes in the cumulative case. The optimal policy might require differentiation between innovations through breadth in the cumulative case.

The results also suggest an intuition regarding the question of the optimal length of patent protection. While this is a classic subject dating back to Nordhaus (1969) and Arrow (1969), among others, the formulation used here provides a new way to look at the role of statutory length when patents may become obsolete before the end of the statutory life of the patent. Long lived patents are beneficial in the sense that they shift the patent's enforcement to relatively low value projects, rather than precluding higher value projects for a smaller length of time.

Here, an infringing improvement is never able to be produced by way of some licensing agreement. This assumption is made to highlight the role of patents in dissuading future innovators. For the patent problem to be interesting, of course, licensing must be imperfect, lest the Coase theorem lead to an efficient outcome. It is possible to imagine that a compulsory licensing scheme such as the one suggested here might facilitate transactions of patents, since the protection they provide would be more clearly delineated than under the current patent law, where the outcome is left entirely to the court's discretion. It is clear that any policy which encourages licensing would have that as an extra benefit.

The idea that patents might help solve the hidden information problem faced in designing rewards for innovators is not new. In fact, John Stuart Mill argued in favor of patents on this basis, stating (from Machlup and Penrose (1950))

> ....an exclusive privilege, of temporary duration is preferable;
> because it leaves nothing to anyone's discretion; because the re-

25

ward conferred by it depends upon the invention's being found useful, and the greater the usefulness, the greater the reward....

When the rewarder's information is limited, prizes may not be useful because the rewarder cannot determine how much to reward. Monopoly rights, on the other hand, have a reward related to the value of the innovation. The mechanism proposed here leaves nothing to the discretion of the patentor; the patentee makes incentive-compatible choice of the appropriate protection. It may be, however, that the duration should not have a fixed length, but rather only end when a sufficiently good report arrives at the patent office.

Important questions remain. An important issue is that of strategic behavior by investors. Here each innovator in the sequence is different. This may overlook the fact that patentees routinely are thinking about future innovations that they will patent themselves when making research and patenting decisions. Another central question is the role of licensing. Incorporating some form of imperfect licensing would add an important element of the role of patents. All of these issues can be addressed within the structure introduced here, taking account of both the cumulative nature of research as well as the asymmetry of information that makes the rewarding of innovation a difficult task of government.

6 Appendix A: Recursive Formulation

In Propositions 1 and 2 we claim that the simple structure we study describes the optimal mechanism of a fairly more general problem. In this appendix we provide a sketch of the proofs.

6.1 The General Mechanism (Proof of Proposition 1)

The set of players is $\aleph$, the set of positive integers. A player's type is $z \in Z$. Let M denote the message space and $m_i(z) \in M$ the report of player i of type z. Player i reports his message m_i at time $t = i$. The information available to the player is his privately observed z_i and the reports of players $j < i$. The mechanism specifies for all t and as a function of messages m_i for $i \leq t$ a mixture σ_{it} over vectors $x_t = (F_{it}, a_{it}), i \leq t$, where F_{it} is a fee (possibly negative) paid by player i and $a_{it} \in \{0, 1\}$ indicates whether the player has a right to produce or not in the current period.

This mechanism defines a (Bayesian) sequential game. At time t, player $i = t$ makes its report $m_i(z)$ and chooses the level of investment $\Delta_i(z)$. These choices may be contingent on the history of past reports. Let $q_i = \sum_{j \leq i} \Delta_j$ denote the current frontier in the quality ladder at time i and normalize $q_0 = 0$. A player's payoff at time t is $-F_{it}$ if $a_{it} = 0$ or $a_{jt} = 1$ for $j > i$ and is equal to $q_i - q_j - F_{it}$ when $a_{it} = 1$ and $j = \max \{l \leq t | a_{lt} = 1\} < i$. The player's investment cost $c(\Delta_i, z_i)$ must also be subtracted from the payoffs for $t = i$.

For any history $m^t = (m_1, ..., m_t)$ of reports and for any $i \leq t$, let $\mu_j(m^t)$ denote the common beliefs that all players -excluding j- have on type Δ_j and let $\bar{q}_j$ denote the corresponding expected value of q_j. Let $\lambda_{ij}(m^t)$ denote the probability that $a_i = 1$ and that j is the second highest player (excluding i) that is allowed to produce in period t. Finally, let $\bar{F}_i(m^t)$ denote the expected

fee paid by player i in period t as a function of the history of reports. The strategy s_i of player i is a mapping from $M^{i-1} \times Z \to M \times \Re_+$, specifying the pair (m_i, Δ_i) as a function of all previous reports and the player's type z. The expected utility for player i is then:

$$W_i = \max_{\Delta_i} \sum_{t=i}^{\infty} \delta^{t-i} E_{(m_1,...,m_i)} \left(\sum_{j<i} \lambda_{ij}(m^t) (\Delta_i + \bar{q}_{i-1} - \bar{q}_j) - \bar{F}_{it}\left(m^t\right) \right) - c\left(\Delta_i, z_i\right).$$

The outcome is determined by a perfect Bayesian equilibrium of this game.

Rewriting the utility function:

$$
\begin{aligned}
W_i &= \max_{\Delta_i} \left[\sum_{t=i}^{\infty} \delta^{t-i} E_{(m_1,...,m_i)} \sum_{j<i} \lambda_{ij}(m^t) \right] \Delta_i - c\left(\Delta_i, z_i\right) \\
&\quad + \sum_{t=i}^{\infty} \delta^{t-i} E_{(m_1,...,m_i)} \left(\sum_{j<i} \lambda_{ij}(m^t) (\bar{q}_{i-1} - \bar{q}_j) - \bar{F}_{it}\left(m^t\right) \right) \\
&= \max_{\Delta_i} d\left(m_1, ..., m_i\right) \Delta_i - c\left(\Delta_i, z_i\right) - F\left(m_1, ..., m_i\right),
\end{aligned}
$$

it follows that all mechanisms that induce the same functions d and F are equivalent. In particular, this implies that there is no loss in restricting the analysis to a mechanism with $\lambda_{ij} = 0$ for all $j < i-1$ and $F_{it} = 0$ for $t > i$. In this mechanism, F can be interpreted as an upfront fee and d as the expected discounted duration of a patent that entitles the firm to market its product but does not preclude firms with earlier patents from doing so.

Let $\lambda_i\left(m^t\right)$ denote the probability the firm i is the leading producer in period t and for $m = (m_1, m_2, ...)$ let $d_i\left(m\right) = \sum_{t=1}^{\infty} \delta^t \lambda_i\left(m^t\right)$. This gives an assignment to player i of the expected discounted duration (as of time zero) corresponding to the sequence of reports m, so that $\sum_{i=1}^{\infty} d_i\left(m\right) = \frac{1}{1-\delta}$ and $d\left(m_1, ..., m_i\right) = \delta^{-t} E_{(m_1,...,m_i)} d_i\left(m\right)$. For any path of reports m, the sequence of probabilities $\{\lambda\left(m^t\right)\}_{t=1}^{\infty}$ is *sorted* if for all t, $\lambda_i\left(m^t\right) > 0$ implies $\lambda_j\left(m^{t-1}\right) = 0$ for all $j > i$. It is easy to see that for any path $\{\lambda\left(m^t\right)\}_{t=1}^{\infty}$ that is not sorted, there exists a unique sorted path $\left\{\tilde{\lambda}\left(m^t\right)\right\}_{t=1}^{\infty}$ that gives rise to

identical $d_i(m)$ for all i. Since all payoff relevant information is contained in the assignments $d_i(m)$, without loss of generality we restrict to sorted paths.

Given the sequence of messages $(m_1, ..., m_{i-1})$, the problem confronted by agent i satisfies the conditions for the Revelation Principle: the principal can always design a mechanism that replicates the message corresponding to the equilibrium strategy for agent i in this Bayesian game, for every agent $i \in \aleph$. Hence, from now on we restrict the analysis to a revelation mechanism.

Letting $d(z_1, ..., z_i)$ denote the expected duration for player i, the profits obtained by a firm are

$$W_i(z_i) = \max_{\widehat{z}, \Delta} \Delta * d(z_1, ..z_i) - c(\Delta, z) - F(z_1, ..z_i).$$

Under the conditions imposed on d, incentive compatibility requires this to be nondecreasing in z_i. And we obtain $\Delta(z, d)$ as the size of innovation that the firm will undertake as a function of the idea and d.

This duration d_i can be expressed recursively as

$$d_i^t(z_1, ..., z_t) = \lambda_i(z_1, ..., z_t) + \delta \int_{z_{t+1}} d_i^{t+1}(z_1, ..., z_t, z_{t+1}) \phi(z_{t+1}) dz_{t+1}. \quad (7)$$

Focusing on sorted paths we can restrict attention to only those contracts in which innovators that are not producing in period t have either $d_i^t = 0$ or $d_i^t = d_i$. That is, suppose that at a certain moment t, the planner has a set $\Im$ of innovators for which duration has not been completely allocated (i.e $d_n^t > 0$ for all $n \in \Im$). Call $s_t = \{z_1, ..., z_t, d_1^t, ..., d_t^t\}$ the state variable in period t.[9] Hence, the present value of social welfare that the planner maximizes corresponds to

$$\widetilde{V}(s_t) = \max_{s'} \sum_{n \in \Im} \lambda_i^t(z^t) \left(\sum_{j \leq n} \Delta(z_j) \right) + \delta \widetilde{V}(s_{t+1})$$

$$\text{s.t.} \quad \sum_{n \in \Im} \lambda_i^t(z^t) = 1$$

---

[9]Although there can be only one innovator producing at a certain point in time, we can still have randomizations.

where s_{t+1} evolves according to (7).[10]

We can rewrite the dynamic programming problem (substracting the present value social welfare of the innovations already promised to be implemented) as,

$$V(d) = \max_{d_1(z),d_2(z)} \int \left\{ \left[\frac{1}{1-\delta} - d_1(z) \right] \Delta(z, d_2(z)) - c(\Delta, z) + \right.$$
$$\left. + \delta V(d_1(z) + d_2(z)) \right\} \phi(z)\, dz$$

s.t. $\quad$ (PK) $\quad d - 1 = \delta \int d_1(z) \phi(z)\, dz$

$$d_1(z) \geq 0, \; d_2(z) \geq 0$$

where $d = \sum_{i=1}^{t-1} d_i^{t-1}$, is the sum of all the durations allocated, $d_1(z)$ is the sum of these durations in the next period and $d_2(z)$ is the duration granted to the innovator appearing in period t.

Call μ the multiplier associated to the Promise Keeping Constraint. Then, the corresponding first order conditions will be,

$$-\Delta + \delta V'(d_1(z) + d_2(z)) + \delta\mu \leq 0 \qquad (= 0 \text{ if } d_1 > 0)$$
$$\left[\frac{1}{1-\delta} - d_1(z) - c_1(\Delta, z) \right] \frac{\partial \Delta}{\partial d_2} + \delta V'(d_1(z) + d_2(z)) \leq 0 \quad (= 0 \text{ if } d_2 > 0).$$

The next lemma gives a sufficient condition for not having innovators to be implemented. That is, the optimal rule consists of implementing the innovator that appears in period t immediately or never.

Lemma 7 *If $\frac{\partial^2 \Delta}{\partial^2 d} > 0$, $d_t^t > 0$ if and only if $d_i^t = 0$ for all $i < t$.*

Proof. We prove towards a contradiction that $d_1 > 0$ and $d_2 > 0$ cannot hold at the same time in histories that happen with positive probability. That is, suppose that this is the case for an interval of z's, $[z_a, z_b]$ with $z_b > z_a$. Then, the first order condition is satisfied with equality. Moreover, since $d_1(z)$ is a maximizer, it must be that $V'' < 0$ evaluated at the maximum. It

---

[10]The constraint says that the product is always served by a monopolist. Obviously this will be the case in the optimum, since we have too little duration.

is easy to verify that V is concave in d and that $d_2(z)$ is increasing in z and $d_1(z)$ is decreasing in z. Moreover, from the first equation we see that if z increases, Δ must increase, and therefore $d_1(z) + d_2(z)$ must decrease (by concavity of V'').

If $d_2 > 0$ it must also be that replacing from the first order condition,

$$\left[\frac{1}{1-\delta} - d_1(z) - c_1(\Delta, z)\right] \frac{\partial \Delta}{\partial d_2} + \Delta - \delta\mu = 0$$

and using the fact that the Δ that maximizes profits for the innovator satisfies $d_2(z) = c_1(\Delta, z)$ we can replace it

$$\left[\frac{1}{1-\delta} - [d_1(z) + d_2(z)]\right] \frac{\partial \Delta}{\partial d_2} + \Delta - \delta\mu = 0. \tag{8}$$

We take the derivative with respect to z and we obtain

$$-\left[\frac{\partial d_1}{\partial z} + \frac{\partial d_2}{\partial z}\right] \frac{\partial \Delta}{\partial d_2} + \left[\frac{1}{1-\delta} - [d_1(z) + d_2(z)]\right] \frac{\partial^2 \Delta}{\partial^2 d_2} \frac{\partial d_2}{\partial z} + \frac{\partial^2 \Delta}{\partial z \partial d_2} + \frac{\partial \Delta}{\partial z} + \frac{\partial \Delta}{\partial d_2} \frac{\partial d_2}{\partial z}.$$

If $\frac{\partial^2 \Delta}{\partial^2 d_2} \geq 0$ this is a contradiction, since the first order condition could not hold with equality for more than one z. ∎

This condition is for example satisfied by the cost function introduced in *Example 1* when $\alpha < 2$.

6.2 Collusion and Reporting of Quality

Whether or not an innovator knows the inherited q upon which new innovations are built, there could be information about that q in the profits earned. If the innovator had, for instance, sole rights to the entire quality spectrum, the realization of demand as a function of price would generate information about q. Even if q is known to all the innovators, though, it is not possible to extract that information in a way that is immune to simple collusion based on bribery. This is in contrast to Crémer and Scotchmer (1997), where private

31

information is elicited at no cost due to the correlation across innovators. Here the sequential arrival leads to problems of collusion.

Suppose q is observable to the innovator, and call $(\widehat{z}_t, \widehat{q}_t)$ the report that innovator t makes about his type and the previous state of the art. We denote $\widehat{z}^t = (\widehat{z}_1, \widehat{z}_2, ..., \widehat{z}_t)$ and $\widehat{q}^t = (\widehat{q}_1, \widehat{q}_2, ..., \widehat{q}_t)$ the history of all reports made by all previous agents up to t.

Just as in the previous section, we focus on revelation mechanisms. Agents declare a message $m_i = (q_i, z_i)$ corresponding to their type and the state of knowledge that they inherited. The two instruments that the planner might use, $d_t(q^t, z^t)$ and $F_t(q^t, z^t)$ can depend on those messages.

Consider the following strategy space. Firm i with type z_i and inherited state q_i chooses $\left(\widehat{z}_i, \widehat{q}_i, \left\{B_i, q_{i+1}^B\right\}\right)$, where $\widehat{z}_i$ and $\widehat{q}_i$ are the reports made to the principal, and $\left\{B_i, q_{i+1}^B\right\}$ is a bribe: It transfers an amount B_i to the next innovator if he agrees to report q_{i+1}^B to the planner.

Profits can be rewritten in this case as,

$$W_i(z_i) = \max_{\widehat{z}, \widehat{q}, q_{i+1}^B, B} \left\{ \max_{\Delta_i} d\left(m_1, ..., m_i\right) \Delta_i - F_i\left(m_1, ..., m_t\right) - c\left(\Delta, z_i\right) - \delta B_i \right\} + B_{i-1}.$$

Denote the term in brackets $U_i\left(z_i, \widehat{z}, \widehat{q}, \widehat{q}_{i+1}^B, B_i\right)$. Since $c\left(\Delta, z\right)$ is decreasing in z_i, U_i is increasing in z_i. We next define a bribe-proof mechanism to be one where a single innovator cannot offer a bribe to another innovator and encourage deceit. Note that bribe-proofness is a very weak form of collusion proofness: it only requires that the mechanism be immune to collusion between two innovators, for instance the outgoing leader and the new leader.

Definition 1 *A mechanism is bribe-proof if there does not exist a $\left\{B_i, q_{i+1}^B\right\}$ such that $B_i > 0$, $q_{i+1}^B \neq q_{i+1}$ that maximizes U_i for inventor i and for all z_{i+1}*

$$W_{i+1}(z_{i+1}) < \max_{\widehat{q}_{i+2}^B, B_{i+1}} U_{i+1}(z_{i+1}, \widehat{z}_{i+1}, \widehat{q}_{i+1}, \widehat{q}_{i+2}^B, B_{i+1}) + B_i.$$

That is, the profits from accepting the bribe are lower than those from telling the truth. Here we assume that bribes can only be offered to a firm with the next innovation that is implemented.

In the next proposition we show that there is no incentive compatible and bribe-proof mechanism that uses any report on q.

Proposition 8 *For any incentive compatible and bribe-proof mechanism $\{d_t(q^t, z^t), F_t(q^t, z^t)\}$ there exists another incentive compatible and bribe-proof mechanism $\{d_t(z^t), F_t(z^t)\}$ that achieves the same allocation.*

Proof. Assume towards a contradiction that there is an incentive compatible and bribe-proof mechanism where U_i depends on the vector of reports q^t for some t.

If $U_i\left(z_i, \widehat{z}, \widehat{q}, \widehat{q}^B_{i+1}, B_i\right)$ is decreasing in $\widehat{q}_i$, a firm with type z and state q can always lie and declare $(\widehat{q}, \widehat{z})$, for example, in such a way that $\Delta(\widehat{z}) > \Delta_i$ and $\widehat{q} + \Delta(\widehat{z}) = q + \Delta_i$. Such a manipulation will always increase profits, since $U_i\left(z_i, \widehat{z}, \widehat{q}, \widehat{q}^B_{i+1}, B_i\right)$ is decreasing in z_i. On the other hand, if U_i is decreasing in $\widehat{q}_i$, firms will declare $\widehat{q}_i = \infty$. Hence, U_i is independent of $\widehat{q}_i$, and profits are the result of $W_i(z_i) = \max_{\widehat{z}, q^B_{i+1}, B_i} U_i\left(z, \widehat{z}, \widehat{q}^B_{i+1}, B_i\right)$.

Innovator i has a cost $\varepsilon > 0$ of bribing innovator $i + 1$, since from the previous step $i+1$'s profits do not depend on that report. However, $\widehat{q}_i$ affects the profits for innovator i, and the bribe will be chosen to maximize these profits. ∎

It is important to notice that here we are not solving for what innovator $i + 1$ will do, or who will be bribed later. In fact, the intuition is simple. Innovator $i + 1$ obtains an invention, investing some costs $c(\Delta_{i+1}, z_{i+1})$ but q is not the one that the previous innovator reported. Since the profits for the innovator remain unchanged regardless of the report, any bribe will be accepted.

6.3 Constant Breadth (Proof of Proposition 2)

For a given incumbent z, the optimal policy amounts to determining the fate of that innovator for each report z'_t made by the potential innovator t periods later. It is easy to see that the optimal policy must be monotone in z'_t; i.e. better reports should be more likely to be allowed to supercede z. Therefore, the optimal policy is determined by a sequence of thresholds $\underline{z}'_t(z)$, where a new leader is chosen if a report of at least $\underline{z}'_t(z)$ is made t periods after the report z. The planner must commit to a certain discounted duration of the monopoly power d, that can be written recursively as,

$$d_t = 1 + \delta \Phi \left(\underline{z}'_t \right) d_{t+1},$$

and $d = d_1$. Therefore, the social planner's problem can be written in the following way,

$$
\begin{aligned}
V\left(q, d_t\right) \;=\; & \max_{\underline{z}'_t, d'(z)} \int_{\underline{z}'_t} \left[q + \Delta\left(z, d'\left(z\right)\right) - c\left(\Delta\left(z, d'\left(z\right)\right), z\right) + V\left(q, d'\left(z\right)\right) \right] \phi\left(z\right) dz + \\
& + \delta \Phi\left(\underline{z}'_t\right) V\left(q, \frac{d_t - 1}{1 - \delta \Phi\left(\underline{z}'_t\right)} \right).
\end{aligned}
$$

As stated in Proposition 2, it turns out that using a constant threshold is the optimal policy, and so, it is without loss of generality that in the rest of the paper we use a threshold as a measure of the duration of monopoly power.

Proof to Proposition 2

Proof. Define $S\left(q, z\right) = q + \Delta\left(z, d'\left(z\right)\right) - c\left(\Delta\left(z, d'\left(z\right)\right), z\right)$, increasing in z and q. Consider all the sequences $\left\{ \underline{z}'_{t+1}, \underline{z}'_{t+2}, ... \right\}$ that result in a duration d. We prove by contradiction in two steps that a constant threshold $\underline{z}'$ is optimal.

First suppose that the optimal plan consists of different thresholds for some periods and after that they become constant, i.e. $\left\{\underline{z}'_{t+1}, \underline{z}'_{t+2}, ..., \underline{z}'_{t+s}, \widehat{\underline{z}}, \widehat{\underline{z}}, \widehat{\underline{z}}, ...\right\}$ for some s. Take the problem in period $t+s-1$.. In this case, $\underline{z}'_{t+s} \neq \underline{z}'_{t+s+1}$ must solve the following problem,

$$V(q, d_{t+s-1}) = \max_{\underline{z}'} \int_{\underline{z}'} S(z)\phi(z)\,dz + \delta\Phi(\underline{z}')\frac{\int_{\widehat{\underline{z}}} S(z)\phi(z)\,dz}{1 - \delta\Phi\left(\widehat{\underline{z}}\right)}$$

$$\text{s.t.} \quad \frac{1}{1 - \delta\Phi\left(\widehat{\underline{z}}\right)} = d_{t+s} = \frac{d_{t+s-1} - 1}{\delta\Phi(\underline{z}')}.$$

Therefore, it must be that

$$-S(\underline{z}'_{t+s}) + \delta\phi(\underline{z}'_{t+s})\frac{\int_{\widehat{\underline{z}}} S(z)\phi(z)\,dz}{1 - \delta\Phi\left(\widehat{\underline{z}}\right)} +$$

$$\delta\Phi(\underline{z}'_{t+s})\left[\frac{-S\left(\widehat{\underline{z}}\right)}{1 - \delta\Phi\left(\widehat{\underline{z}}\right)} + \frac{\int_{\widehat{\underline{z}}} S(z)\phi(z)\,d\delta\phi\left(\widehat{\underline{z}}\right)}{\left[1 - \delta\Phi\left(\widehat{\underline{z}}\right)\right]^2}\right]\frac{\partial\widehat{\underline{z}}}{\partial\underline{z}'_{t+s}} = 0$$

where from the constraint we obtain $\frac{\partial\widehat{\underline{z}}}{\partial\underline{z}'} = -\frac{\phi\left(\underline{z}'_{t+s}\right)\left[1 - \delta\Phi\left(\widehat{\underline{z}}\right)\right]}{\delta\phi\left(\widehat{\underline{z}}\right)\Phi(\underline{z}'_{t+s})}$. Replacing,

$$\phi\left(\underline{z}'_{t+s}\right)\left[S\left(\widehat{\underline{z}}\right) - S\left(\underline{z}'_{t+s}\right)\right] = 0$$

and so, the maximum is attained at $\underline{z}'_{t+s-1} = \widehat{\underline{z}}$, contradicting our premise.

Now suppose that instead there is a plan that gives strictly bigger social welfare than one with a constant threshold and it consists on infinitely different protections. That is, there is no time $\widehat{s}$ such that for all $s > \widehat{s}$, $\underline{z}'_{t+s} = \underline{z}'_{t+s+1}$. By continuity and discounting, there must exist another path that has constant protection after a number of periods $s > 1$ and that it also gives higher welfare than a constant threshold. However, by the previous step, this is a contradiction. Therefore, given a duration d the optimal plan has the form,

$$d = \frac{1}{1 - \delta\Phi\left(\underline{z}'\right)}.$$

It can be shown that the value function $V(q, d)$ is strictly concave in d and so, allowing for random protections will not improve upon it. ∎

7 Appendix B: Other Proofs

Proof to Proposition 3:

Proof. For our result it is enough to show that V is strictly supermodular in z and $\underline{z}'$. In particular, this condition means that the derivative of the first order condition

$$\left[\frac{1}{1-\delta} - c_1(\Delta^p(z,\underline{z}'),z)\right]\frac{\partial\Delta^p}{\partial\underline{z}'} + \delta\frac{\partial V}{\partial\underline{z}'}(q + \Delta^p(z,\underline{z}'),\underline{z}') \qquad (9)$$

with respect to z is strictly positive. Since $\frac{\partial V}{\partial\underline{z}'\partial q} = 0$, we obtain that,

$$-c_{11}(\Delta^p(z,\underline{z}'),z)\frac{\partial\Delta^p}{\partial\underline{z}'}\frac{\partial\Delta^p}{\partial z} - c_{12}(\Delta^p(z,\underline{z}'),z)\frac{\partial\Delta^p}{\partial\underline{z}'} + \left[\frac{1}{1-\delta} - c_1(\Delta^p(z,\underline{z}'),z)\right]\frac{\partial^2\Delta^p}{\partial\underline{z}'\partial z}$$

$$= \left[\frac{1}{1-\delta} - c_1(\Delta^p(z,\underline{z}'),z)\right]\frac{\partial^2\Delta^p}{\partial\underline{z}'\partial z}.$$

Notice that the first term is strictly positive, since the optimal choice of Δ^P satisfies

$$\frac{1}{1-\delta} - c_1(\Delta^p(z,\underline{z}'),z) > \frac{1}{1-\delta\Phi(\underline{z}')} - c_1(\Delta^p(z,\underline{z}'),z) = 0.$$

Therefore, the derivative will be strictly positive if and only if $\frac{\partial^2\Delta^p}{\partial\underline{z}'\partial z} > 0$. Function V represents a contraction mapping. Hence, usual dynamic programming arguments ensure that this function is weakly supermodular. The proof of strict supermodularity is obtained by contradiction. ∎

Proof to Proposition 4:

Proof. That $\Delta^p(z,\cdot) \leq \Delta^*(z)$ is obvious from the remarks in the text. For the second part, we take a recursive argument. Assuming that $\underline{z}' \geq \underline{z}^*$ we show that $\underline{z} > \underline{z}^*$.

First notice that from equation (6) we can compute

$$\frac{\partial V}{\partial \underline{z}}(q, z) = \frac{-\phi(\underline{z}(z))}{1 - \delta\Phi(\underline{z}(z))}\{\Delta^p(\underline{z}, z') - c(\Delta^p(\underline{z}, z'), \underline{z})+$$
$$\delta\left[V(q + \Delta^p(\underline{z}, z'), z') - V(q, \underline{z})\right]\}.$$

Since for all z, $\underline{z}$ is the maximizer of this function, it must be that for all x

$$\frac{\Delta^p(z, \underline{z})}{1 - \delta} - c(\Delta^p(z, \underline{z}), \underline{z}) + \delta V(q, \underline{z}) \geq \frac{\Delta^p(z, x)}{1 - \delta} - c(\Delta^p(z, x), x) + \delta V(q, x),$$

and together with the fact that $\frac{\partial V}{\partial q} = \frac{1}{1-\delta}$ we obtain that

$$\frac{\partial V}{\partial \underline{z}}(q, \underline{z}) \geq \frac{-\phi(\underline{z}(z))}{1 - \delta\Phi(\underline{z}(z))}\left[\frac{\Delta^p(\underline{z}, \underline{z})}{1 - \delta} - c(\Delta^p(\underline{z}, \underline{z}), \underline{z})\right].$$

If $\Delta^p(z, \cdot) < \Delta^*(z)$, the FOC implies that $\frac{\partial V}{\partial \underline{z}} < 0$ and, $\frac{\Delta^p(\underline{z}, \underline{z})}{1-\delta} - c(\Delta^p(\underline{z}, \underline{z}), \underline{z}) > 0$, so that

$$\frac{\Delta^*(\underline{z})}{1 - \delta} - c(\Delta^*(\underline{z}), \underline{z}) \geq \frac{\Delta^p(\underline{z}, \underline{z})}{1 - \delta} - c(\Delta^p(\underline{z}, \underline{z}), \underline{z}) > 0.$$

Because $\underline{z}^*$ is characterized by $\frac{\Delta^*(\underline{z}^*)}{1-\delta} - c(\Delta^*(\underline{z}^*), \underline{z}^*) = 0$, and this is a strictly increasing function, $\underline{z} > \underline{z}^*$.

If $\Delta^p(z, \cdot) = \Delta^*(z)$, it must be that $\underline{z} = \infty > \underline{z}^*$ ∎

Proof to Lemma 1:

Proof. Suppose that the following innovations are implementable. Using the Revelation Principle this means in particular that for the worse innovator $\underline{z}'_\tau$,

$$\underline{z}'_\tau \in \arg\max_{\widehat{z}} \frac{\Delta^p}{1 - \delta\Phi(\underline{z}''_\tau(\widehat{z}))} + \frac{1 - \Phi(\underline{z}''_\tau(\widehat{z}))}{1 - \delta\Phi(\underline{z}''_\tau(\widehat{z}))}\delta\tau''(\widehat{z}) - c(\Delta^p, \underline{z}'_\tau) - \tau' - \sigma''.$$

Using the Implicit Function Theorem and the fact that $\tau'(z)$ and $\sigma(z)$ are independent of z, we obtain that for any contract $\frac{\partial \underline{z}'_\tau}{\partial \sigma'} = 0$ and $\frac{\partial \underline{z}'_\tau}{\partial \tau'} > 0$.

For the current inventor with quality z, we need to check the Spence-Mirrlees condition,

$$\frac{\partial W^\tau}{\partial \tau' \partial z}(z, \tau', \tau) = \frac{\delta \phi(\underline{z}'_\tau)}{1 - \delta \Phi(\underline{z}'_\tau)} \frac{\partial \Delta^p}{\partial \underline{z}'_\tau} \frac{\partial \underline{z}'_\tau}{\partial \tau'} > 0.$$

And this immediately implies using the previous result, that in order for a mechanism to be implementable, $\frac{\partial \tau'}{\partial z} \geq 0$.

From the envelope condition,

$$\frac{\partial W^\tau}{\partial z} = -\frac{\partial c}{\partial z}(\Delta^p, z),$$

and by the definition of $W^\tau(\underline{z}^0_\tau, \tau, 0) = 0$, this means that $W^\tau(\underline{z}^0_\tau, \tau', \tau) = -\tau$. Integrating with respect to z and using this boundary condition we obtain the expression for $\sigma'(z)$ which does not depend on τ. ∎

Proof to Proposition 5:

Proof. Two contracts $\{\underline{z}'(z), F(z)\}$ and $\{\tau'(z), \sigma'(z)\}$ will be equivalent if they guarantee the same profits to innovators of all types z.

Any incentive compatible contract $\{\underline{z}'(z), F(z)\}$ achieves profits,

$$W(z, \underline{z}) = \int_{\underline{z}}^{z} -\frac{\partial c}{\partial z}(\Delta^p(x, \underline{z}'(x)), x) dx,$$

while a compulsory licensing based contract results in profits,

$$W^\tau(z, \tau', \tau) = \int_{\underline{z}_\tau^0}^{z} -\frac{\partial c}{\partial z}(\Delta^p(x, \underline{z}'_\tau(x)), x) dx - \tau.$$

Hence, for both contracts to be equivalent it must be that $W(z, \underline{z}) = W^\tau(z, \tau', \tau)$, which results in

$$\tau = \tau(z) = \int_{\underline{z}_\tau^0}^{\underline{z}(z)} -\frac{\partial c}{\partial z}(\Delta^p(x, \underline{z}'(x)), x) dx,$$

for all z. It can be verified, using convexity of c with respect to z, that this function is increasing in z. This result implies that there is an implementable compulsory licensing scheme that leads to the same profits as any implementable contract $\{\underline{z}'(z), F(z)\}$. ∎

References

Allison, John R. and Mark A. Lemley (1998). 'Empirical Evidence on the Validity of Litigated Patents,' Mimeo.

Anderson, Simon P., Andre de Palma and Jacques-Francois Thisse (1992). 'Discrete choice theory of product differentiation,' *Cambridge, Mass., M.I.T. Press.*

Arrow, Kenneth J. (1969). 'Classificatory Notes on the Production and Transmission of Technological Knowledge,' *American Economic Review*, **59**: 29-35.

Baron, David P. and Roger B. Myerson (1982). 'Regulating a Monopolist with Unknown Costs,' *Econometrica*, **50**: 911-930.

Chiesa, Gabriella and Vincenzo Denicolo (1999). 'Patents, Prizes, and Optimal Innovation Policy.' *Mimeo.* University of Bologna.

Cornelli, Francesca and Mark Schankerman (1999). 'Patent Renewals and R&D Incentives,' *Rand Journal of Economics*, **30**: 197-213.

Crémer, Jacques, and Suzanne Scotchmer (1997). 'Optimal R&D Procurement: Why Patents?' *Mimeo.*

Gilbert, Richard and Carl Shapiro (1990). 'Optimal Patent Length and Breadth,' *Rand Journal of Economics*, **30**: 106-112.

Green, Jerry and Suzanne Scotchmer (1995). 'On the Division of Profit between Sequential Innovators', *Rand Journal of Economics* **26**: 20-33.

Laffont, Jean-Jacques and Jean Tirole (1986). 'Using Cost Observation to Control a Public Firm,' *Journal of Political Economy*, **94**: 614-641.

Lewis, Tracy R. and David E.M. Sappington (1988a). 'Regulating a Monopolist with Unknown Demand and Cost Functions' *Rand Journal of Economics*, **19**: 438-57.

Lewis, Tracy R. and David E.M. Sappington (1988b). 'Regulating a Monopolist with Unknown Demand,' *American Economic Review*, **78**: 986-98.

Llobet, Gerard (1999). 'The Economics of Patent Litigation,' *Mimeo.*

Nordhaus, William D. (1969). 'Invention, growth, and welfare; a theoretical treatment of technological change,' *Cambridge, Mass., M.I.T. Press.*

Machlup, Fritz, and Edith Penrose (1950). 'The Patent Controversy in the Nineteenth Century,' *Journal of Economic History*, **10**:1-29.

O'Donoghue, Ted, Suzanne Scotchmer and Jacques-Francois Thisse (1998). 'Patent Breadth, Patent Life and the Pace of Technological Progress,' *Journal of Economics and Management Strategy*, **7**: 1-32.

Shavell, Steven and Tanguy van Ypersele (1999) 'Rewards versus Intellectual Property Rights,' *Mimeo.*

Scotchmer, Suzanne (1999) 'On the Optimality of the Patent Renewal System,' *Rand Journal of Economics*, **30**: 181-96.

Scotchmer, Suzanne, and Jerry Green (1990). 'Novelty and Disclosure in Patent Law,' *Rand Journal of Economics*, **21**: 131-46.

Tandon, Pankaj (1982) 'Optimal Patents with Compulsory Licensing,' *Journal of Political Economy*, **90**: 470-486.

Wright, Brian D. (1983) 'The Economics of Invention Incentives: Patents, Prizes, and Research Contracts,' *American Economic Review*, **73**: 691-707.