Rewarding Sequential Innovators: Prizes, Patents and Buyouts*

Gerard Llobet[†] Hugo Hopenhayn[‡] Matthew Mitchell[§]

December 1999

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, contrary to the existing literature, 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, a higher price ceiling is obtained. Any subsequent innovator must pay this price and purchase its own buyout fee contract.

*Address for correspondence: Gerard Llobet, Department of Economics, Harkness Hall, University of Rochester, Rochester, NY 14627. We thank Fernando Alvarez for helpful comments.

†University of Rochester

‡Universidad Torcuato Di Tella and University of Rochester

§University of Minnesota

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 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. 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. Although not the focus of her paper, this is true in the single innovation model used by Scotchmer (1999). We show that, in a slightly different model, a mechanism similar to the one used to regulate a monopolist with unknown cost (Baron and Myerson (1982)) can support the efficient choice of research without any patent rights being granted. In both Scotchmer (1999) and the model here, the optimal reward for an innovator with unknown costs is identical to the optimal regulation of a monopolist with unknown costs.

When value is known, 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.

We characterize the optimal reward system in a cumulative innovation context with incomplete innovation 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]

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

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

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. 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 buyout fees 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 buyout 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 (1998) 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 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 results 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 A Single Innovation

In order to set the stage for the model with cumulative innovations, it is useful to start by addressing how the mechanism designer reacts to a single innovator in isolation. 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 at according to a discount factor δ.

This section sets the stage for the cumulative case by studying a single innovator. Consider a single firm arising with an idea, which can be turned into an improved product if research is undertaken. 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

[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.

Any of these interpretations makes natural to assume that inventions cannot be bought and put in the public domain, because this kind of knowledge is non-transferable.

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 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, the product is sold for $p = \Delta$, the consumers are left with no surplus, and the firm makes Δ units of profits per period.

Suppose that there is a government regulator which can observe q for each good, but cannot tell what the costs of innovation for the good were. That is, $z \in R_+$ is private information to the innovator. The question is, what can the regulator do to encourage innovation in this case?

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

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

The patent office can reward the innovator with a fee $F(z)$ (which may be negative, 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 referred to as a prize system. Denote $\Delta(z)$ the improvement requested for a type-z firm.[3] Here, the fees will be negative (prizes), but in future sections they may not be. Let $R(T) = \sum_{i=1}^{T} \delta^{i-1}$. The patent office's problem of facing this adverse selection is

$$\max_{\Delta(z);F(z);T(z)} \int \left[\frac{\Delta(z)}{1-\delta} - c(\Delta(z);z) \right] \psi(z)dz,$$

$$s.t.$$

$$z = \arg\max_{\hat{z}} R(T(\hat{z}))\Delta(\hat{z}) - F(\hat{z}) - c(\Delta(\hat{z});z) \quad \text{for all } z,$$
$$R(T(z))\Delta(z) - F(z) - c(\Delta(z);z) \geq 0 \quad \text{for all } z.$$

The first constraint, incentive compatibility, ensures that each firm truthfully reveals its type z. The second, individual rationality, guarantees that each agent prefers the reward to remaining anonymous. The following depicts the optimal prize system $F(z)$ which support the optimal level of research. Let $\underline{z}^*$ denote the smallest project to be implemented, i.e. the one such that $\Delta^*(z) = 0$ for all $z < \underline{z}$.

Proposition 1 When Δ is observed, the socially optimal level of research, i.e. $\Delta(z) = \Delta^*(z)$ and $\underline{z} = \underline{z}^*$, is achieved with $T(z) = 0$ and

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

The nature of the optimal mechanism 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.

[3] Implicitly, the reward agency is offering punishments to firms which report z but develop $\Delta \neq \Delta(z)$.

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 $F(z)$). We have also disregarded the possibility that funds are costly for the government to obtain. If both funds were costly and monopoly generated static markup distortions, it is reasonable to think the latter would be more important, and therefore the government would choose the prize method rather than granting monopoly power to the innovators. In fact, if the government has access to a consumption tax, it can always raise funds with the distortion which generates profits for the monopolist. It is not hard to imagine that the government can raise funds at least as efficiently as with a consumption tax; in that case the optimal choice is to use the prize scheme when possible.

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. As a result, the regulator faces an additional constraint due to moral hazard:

$$\delta(z) = \arg\max_{\hat{\delta}} R(T)\hat{\delta} - c(\hat{\delta}; z) - F(z).$$

Because the monopolist can extract all the surplus from the consumers, the efficient outcome can be attained, but only through the granting of monopoly power.

Proposition 2 When δ is unobserved, the socially optimal level of research, i.e. $\delta(z) = \delta^*(z)$, is only achieved with $T(z) = \infty$.

Because of the inelastic demand, even unknown ¢ may not prevent the government from implementing the e¢cient level of innovation. However, that result can only be attained through the granting of patent rights, as opposed to the earlier case with known ¢, 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 ¢ 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 di¤erences in the next section is that the optimal policy will be di¤erent and will be unable to attain the e¢cient level of research.

3 Multiple Innovations

Suppose that each period a new …rm arrives with a new idea z, allowing an improvement over the current quality in the amount ¢ with cost c(¢; z) as before. If ¢ is observable, the government could simply employ the prize system of the last section in sequence to each innovator. The conclusions of that section would remain unaltered. As such, we focus our attention on the case where ¢ is unobserved, so that there is both moral hazard and adverse selection. The optimal patent for a single innovator involved a patent. O¤ering a patent, however, is quite di¤erent 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 …rst innovator might preclude some future improvements. The patent, in order to be economically meaningful, may have to preclude small improvements, else the original innovator's ability to pro…t from research may be short lived.

Given that property rights might change as innovations arrive, a patent will be de…ned by its "breadth" $\underline{z}_t^0(z)$. That is, a …rm reporting that it is of type z receives sole rights until a report of $\underline{z}_t^0(z)$ is made by another innovator t periods after the patent is issued. For instance, the current US statute can be thought of as calling for a common breadth until the patent reaches a statutory time limit, at which point breadth is reduced to zero, and so any improvement is allowed to be produced.[4]

The set of possible instruments we allow, then, is very large.[5] Fortunately, we find a relatively simple characterization of the optimal mechanism. From the innovator's perspective, all that matters is the expected duration d of the patent, where duration is defined by

$$d = \left(1 + \sum_{i=t}^{\infty} \delta^{i+1-t} \left(\prod_{j=t}^{i} \psi(\underline{z}_j^0)\right)\right)$$

The innovator's expected revenue from sales of the product are $\Delta \times d$, the product of Δ, the per period profits of producing the innovation, and d, the expected duration. 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 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]

We have two problems to solve. The first is, given d, how should breadth $\underline{z}_t^0(z)$ be assigned? Once that is resolved, it remains to be decided how to allocate duration for different types z. 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

[4] Of course, the other element is that after the end of the patent's statutory life, imitators may enter. Imitators are ignored here, though, since there are no static costs of monopoly in this model

[5] It turns out that the value function we derive for the patentor is concave; therefore, it is without loss of generality that we do not explicitly consider the possibility of randomization over rewards.

[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.

additional trade-o¤ that larger duration means excluding more future ideas.

It turns out that it is never optimal to o¤er changing breadth over time, and therefore we will look at in…nitely lived patents of constant breadth $\underline{z}^0(z)$.

Proposition 3 Given a duration d, it is optimal to grant a constant breadth, in…nitely lived patent $\underline{z}^0(z)$.

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 …xed 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 o¤ered, less breadth needs be o¤ered 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 pro…ts 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 di¤erent 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.

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 idea required for a noninfringing innovation. As we will show next, there is a strictly increasing relationship between the size of the invention achieved, ¢, and the parameter z. Breadth, then, can be thought of in the usual sense: a higher $\underline{z}^0$ means that it will take a larger improvement for a subsequent product to be produced.

In order to …gure out the optimal $\underline{z}^0(z)$, and hence the optimal duration, …rst an innovator's response, in terms of innovation, to a given breadth $\underline{z}'$ must be calculated. A given firm chooses Δ to solve

$$\max_{\Delta} \frac{\Delta}{1 - \delta\Phi(\underline{z}')} - c(\Delta; z) - F(z).$$

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

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

where 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,

$$\frac{\partial \Delta^p}{\partial z} = \frac{c_{12}(\Delta^p; z)}{-c_{11}(\Delta^p; z)} > 0, \tag{2}$$

$$\frac{\partial \Delta^p}{\partial \underline{z}} = \frac{\frac{-\delta\phi(\underline{z}')}{(1-\delta\Phi(\underline{z}'))^2}}{-c_{11}(\Delta^p; z)} > 0. \tag{3}$$

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.

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 principal can observe neither the quality of the invention nor the research efficiency of the firm. The condition for truthful revelation given the two instruments, $\underline{z}'$ and F, is for the report $\mathbf{b}$ to be the one that solves

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

Checking the Spence-Mirrlees condition we obtain

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

The following theorem characterizes implementable choices.

Proposition 4 (Guesnerie and Laffont, 1984) A piecewise C^1 decision function $\underline{z}'$ is implementable only if $\frac{\partial^2 W}{\partial \underline{z}'\partial z}\frac{\partial \underline{z}'}{\partial z} \geq 0$ whenever $\underline{z}'$ is differentiable at z.

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:

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

Because $\frac{\partial W}{\partial \underline{z}'} = \frac{\delta\phi(\underline{z}(\mathbf{b}_i))\Delta^p(z;\underline{z}_i)}{1-\delta\Phi(\underline{z}(\mathbf{b}_i))} > 0$ 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. Fortunately, the problem has a relatively simple recursive structure. The dynamic problem of the principal is

$$\begin{aligned} V(q;\underline{z}) &= \Phi(\underline{z})[q + \delta V(q;\underline{z})] + \\ &\max_{\underline{z}'(z)} \int_{\underline{z}} [(q+\Delta^p(z;\underline{z}')) - c(\Delta^p(z;\underline{z}');z) + \delta V^p(q+\Delta^p(z;\underline{z}');\underline{z}')]\phi(z)dz \end{aligned} \quad (4)$$

subject to the constraints

$$(IR)\ W(z;\underline{z}'(z)) = 0$$

$$(IC)\ \underline{z}'(z) \text{ increasing}$$

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 λ of acquiring funds, as in Laffont and Tirole (1986). None of the results presented are adversely effected by the inclusion of such a cost.

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 5 *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,

[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.

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.

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

This illustrates the cost of the patent office's lack of knowledge of ϕ in the cumulative research context. Another way to show the cost of the incomplete information is to define the growth rate conditional to the previous innovation as,

$$g(z) = \int_{\underline{z}(z)} \phi(x; \underline{z}'(x))\psi(x)dx$$

And the unconditional growth rate (given the stationarity of the policy function in both cases) as,

$$\overline{g} = \int \left[\int_{\underline{z}(z)} \phi(x; \underline{z}'(x))\psi(x)dx \right] \psi(z)dz$$

We define g^* and g^p as the growth rates when actions are observable (the optimum) and hidden (where some projects are precluded) respectively. It is immediate from Proposition 6 that the growth rate will be lower when there is private information.

Corollary 7 For all z, $\overline{g}^* = g^*(z)$ and $\overline{g}^* > g^p(z)$. Moreover, $\overline{g}^* > \overline{g}^p$.

The fact that we can compare the two kinds of economies allow us to study which is the value of information in this model. Such an estimation can be related to whether we want the patent system to be just a registration mechanism or an exhaustive way to review and study innovations.

4 Decentralization through Buyouts

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 three instruments $\{\Delta; \underline{z}; F\}$ to reward innovators. We now consider different mechanisms 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. In implementing these breadths, we will focus on buyout fees. 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 buyout 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 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. 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 implement, but potentially restrictive in what they can implement. We show that the optimal patent menu from the prior section can, in fact, be implemented 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

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

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 $\underline{z}_{\tau}(\hat{\Delta})$, 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 buyout-based contract can implement the mechanism described in the prior section.

Lemma 8 The buyout-based 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:

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

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

Proposition 9 Any incentive compatible contract $\{\underline{z}'(z);F(z)\}$ can be achieved using a buyout-based contract $\{\tau'(z);\sigma'(z)\}$. That is, for all z $\underline{z}'(z) = \underline{z}^0_{\tau}(z)$. 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 innovation will have to pay more in order to produce.

5 Summary

The patent statue involves a single, somewhat vague de…nition 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, we show here how patent breadth can be implemented through a system of buyout fees, replacing much of the burden now placed on the courts. This is especially important if one wishes to o¤er di¤ering patent breadth to di¤erent inventions. It is not surprising that in light of the heterogeneity of inventions, the optimal reward policy may reward di¤erent innovations with di¤erent breadth.

In this paper, the optimal policy is characterized, and it is found that in many cases such di¤erentiation is optimal. This is an important practical point: it may be both optimal as well as feasible to o¤er multiple patent breadths. In addition, the optimal policy has patents of in…nite statutory duration. The sort of patent policy described here would involve a more complicated set of patents o¤ered, 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 in…nite statutory duration. They expire, e¤ectively, only when something suitably better comes along. This results because it is always better to transfer pro…ts 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 di¤ers from the lessons of the one innovation case, which has been studied in slightly di¤erent 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 de…nition of a patent when further innovations will arrive. It may be more di¢cult to achieve e¢cient research outcomes in the cumulative case. The optimal policy might require di¤erentiation 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 bene…cial 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 e¢cient outcome. It is possible to imagine that a buyout 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 bene…t.

Important questions remain. An important issue is that of strategic behavior by investors. Here each innovator in the sequence is di¤erent. 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 di¢cult task of government.

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 di¤erentiation,’ Cambridge, Mass., M.I.T. Press.

Arrow, Kenneth J. (1969). ‘Classi…catory 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.

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

Green, Jerry and Suzanne Scotchmer (1990). ‘On the Division of Pro…t between Sequential Innovators’, Rand Journal of Economics 26: 20-33.

Guesnerie, Roger and Jean-Jacques La¤ont (1984). ‘A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm,’ Journal of Public Economics, 25: 328-370.

La¤ont, 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.

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.

6 Proofs

Proof to Proposition 1:

Proof. The first order condition for reporting the true value is

$$T' = \frac{\partial \Delta^*}{\partial \hat{z}} c_1 = -\frac{c_{12}}{c_{11}} c_1,$$

which is satisfied for $\hat{z} = z$, since $T' = \frac{\partial \Delta^*}{\partial z} c_1(\Delta^*(\hat{z}); \hat{z})$. Integrating and solving in the profit function for F we obtain the desired result. ■

Proof to Proposition 2:

Proof. Because Δ is not observable and this is a function of z, F only depends on z. The private choice of Δ is obtained according to,

$$R(T) - \frac{\partial c}{\partial \Delta}(\Delta; z) = 0,$$

so, a necessary condition to achieve the first best is $R(T) = \frac{1}{1-\delta}$, or $T = \infty$. That is sufficient is an obvious consequence from the lack of monopoly distortion. ■

Proof to Proposition 3

Proof. Consider all the sequences $\{\underline{z}_1; \underline{z}_2; \ldots\}$ that result in a duration d. Define d_t for $t = 1; \ldots; \infty$ as $d_t = 1 + \delta\Phi(\underline{z}_t)d_{t+1}$.where $d = d_1$.

Denote S as the social welfare from implementing an idea z, increasing in z. Hence, the value of offering a duration d solves,

$$V(d_1) = \max_{\{\underline{z}_1; \underline{z}_2; \ldots\}} \int_{\underline{z}_1} S(z)\phi(z)\,dz + \delta\Phi(\underline{z}_1)\left[\int_{\underline{z}_2} S(z)\phi(z)\,dz + \delta\Phi(\underline{z}_2)[\ldots]\right]$$

and can be written recursively for any t as,

$$\begin{aligned} V(d_t) &= \max_{\underline{z}_t} \int_{\underline{z}_t} S(z)\phi(z)\,dz + \delta\Phi(\underline{z}_t)V(d_{t+1}) \\ \text{s.t.} \quad d_t &= 1 + \delta\Phi(\underline{z}_t)d_{t+1}. \end{aligned}$$

Let's start considering one period deviations. Suppose that from period $t+1$, we choose a plan with a constant $\underline{z}$ that we will call $\underline{z}^*$, and we want to find which is the optimal $\underline{z}_t$. For the next period on, the social welfare, that we will denote as V^* can be formulated as

$$V^*(d_{t+1}) = \frac{\int_{\underline{z}^*} S(z)\phi(z)\,dz}{1-\delta\Phi(\underline{z}^*)},$$

since in every period we provide the same protection. Of course, we need to set $\underline{z}^*$ to solve $d_{t+1} = \frac{1}{1-\delta\Phi(\underline{z}^*)}$ so that the previous commitment is satisfied. Replacing the expression for V^* into V we obtain,

$$\begin{aligned} V(d_t) &= \max_{\underline{z}_t} \int_{\underline{z}_t} S(z)\phi(z)\,dz + \delta\Phi(\underline{z}_t)\frac{\int_{\underline{z}^*} S(z)\phi(z)\,dz}{1-\delta\Phi(\underline{z}^*)} \\ \text{s.t.} \quad \frac{1}{1-\delta\Phi(\underline{z}^*)} &= \frac{d_t-1}{\delta\Phi(\underline{z}_t)} \end{aligned}$$

The FOC for this problem is

$$-S(\underline{z}_t) + \delta\phi(\underline{z}_t)\frac{\int_{\underline{z}^*} S(z)\phi(z)\,dz}{1-\delta\Phi(\underline{z}^*)} + \delta\Phi(\underline{z}_t)\left[\frac{-S(\underline{z}^*)}{1-\delta\Phi(\underline{z}^*)} + \frac{\int_{\underline{z}^*} S(z)\phi(z)\,d\delta\phi(\underline{z}^*)}{[1-\delta\Phi(\underline{z}^*)]^2}\right]\frac{\partial\underline{z}^*}{\partial\underline{z}_t}$$

The expression for $\frac{\partial\underline{z}^*}{\partial\underline{z}_t}$ can be obtained from the constraint as

$$\frac{\partial\underline{z}^*}{\partial\underline{z}_t} = -\frac{\phi(z_t)\left[1-\delta\Phi(\underline{z}^*)\right]}{\delta\phi(\underline{z}^*)\Phi(\underline{z}_t)}.$$

Replacing, the FOC becomes

$$\phi(z_t)\left[S(\underline{z}^*) - S(\underline{z}_t)\right],$$

Obviously, this FOC is zero if and only if $\underline{z}_t = \underline{z}^*$. Trivially, if $\underline{z}_t > \underline{z}^*$, the condition is negative, and if $\underline{z}_t < \underline{z}^*$ the opposite is true. Because this is the unique fixed point of V, $\underline{z}_t = \underline{z}^*$ is the maximizer of V.

To show that multiple deviations do not improve upon a constant $\underline{z}_t$ we do it in two steps. In the first step, suppose that there is a sequence with a finite number of deviations that achieves a bigger V that a constant $\underline{z}$. Take the one that deviates for less periods,

$\hat{t}$, $\{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_{\hat{t}}, \hat{\underline{z}}, \hat{\underline{z}}, \hat{\underline{z}}, \ldots\}$ from a constant $\underline{z}$ that lead to a higher value V. Take period $\hat{t}$. Because one period deviations are not profitable, we can construct another sequence that deviates $\hat{t} - 1$ periods and achieves a bigger value V, leading to a contradiction.

In the second step, suppose that an infinite number of deviations achieves a strictly bigger value V. By continuity, there must be a sequence with a finite number of deviations that also achieves a bigger value, and this is not possible from the previous step. Therefore, given a duration d the optimal breadth has the form,

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

■

Proof to Proposition 5:

Proof. For our result 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

$$\left[\frac{1}{1-\delta} - c_1(\Delta^p(z;\underline{z}'); z)\right]\frac{\partial \Delta^p}{\partial \underline{z}'} + \delta\frac{\partial V}{\partial \underline{z}'}(q + \Delta^p(z;\underline{z}'); \underline{z}') = 0. \tag{5}$$

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 is 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 6:

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

First notice that from equation (4) we can compute

$$\begin{aligned} \frac{\partial V}{\partial \underline{z}}(q;\underline{z}) &= \frac{-\phi(\underline{z}(z))}{1-\delta\Phi(\underline{z}(z))}\{\Delta^p(\underline{z};\underline{z}') - c(\Delta^p(\underline{z};\underline{z}');\underline{z}) + \\ &\qquad \delta[V(q+\Delta^p(\underline{z};\underline{z}');\underline{z}') - V(q;\underline{z})]\}. \end{aligned}$$

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

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

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

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

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

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

Because $\underline{z}^*$ is characterized by $\frac{\Delta^*(\underline{z}^*)}{1-\delta} - c(\Delta^*(\underline{z}^*);\underline{z}^*) = 0$, and this is an 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 8:

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

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

Profits for any inventor are independent of q and so, 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 9:

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 buyout based contract results in profits,

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

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

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

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 buyout scheme that leads to the same profits than any implementable contract $\{\underline{z}'(z); F(z)\}$. ■