RAND Journal of Economics
Vol. 40, No. 4, Winter 2009
pp. 611–635

Sequential innovation, patents, and imitation

James Bessen*

and

Eric Maskin**

We argue that when innovation is "sequential" (so that each successive invention builds in an essential way on its predecessors) and "complementary" (so that each potential innovator takes a different research line), patent protection is not as useful for encouraging innovation as in a static setting. Indeed, society and even inventors themselves may be better off without such protection. Furthermore, an inventor's prospective profit may actually be enhanced by competition and imitation. Our sequential model of innovation appears to explain evidence from a natural experiment in the software industry.

1. Introduction

■ The standard economic rationale for patents is to protect inventors from imitation and thereby give them the incentive to incur the cost of innovation. Conventional wisdom holds that, unless would-be competitors are restrained from imitating an invention, the inventor may not reap enough profit to cover that cost. Thus, even if the social benefit of invention exceeds the cost, the potential innovator without patent protection may decide against innovating altogether.[1]

Yet interestingly, some of the most innovative industries of the last forty years—software, computers, and semiconductors—have historically had weak patent protection and have experienced rapid imitation of their products.[2] Defenders of patents may counter that, had stronger

*Boston University School of Law and Research on Innovation; jbessen@researchoninnovation.org.
**Institute for Advanced Study and Princeton University; maskin@ias.edu.
We thank Stefan Behringer, Lee Branstetter, Daniel Chen, Iain Cockburn, Partha Dasgupta, Nancy Gallini, Alfonso Gambardella, Bronwyn Hall, Adam Jaffe, Lawrence Lessig, David Levine, Suzanne Scotchmer, Jean Tirole, Michael Whinston, two referees, the editor Joseph Harrington, and participants at many seminars and conferences for helpful comments. We gratefully acknowledge research support from the NSF.

[1] This is not the only justification for patents. Indeed, we will emphasize a different, although related, rationale in Section 2. But, together with the spillover benefit that derives from the patent system's disclosure requirements, it constitutes the traditional justification.

[2] Software was routinely excluded from patent protection in the United States until a series of court decisions in the mid-1980s and 1990s. Semiconductor and computer patent enforcement was quite uneven until the organization of the Federal Circuit Court in 1982. Both areas contend with substantial problems of prior art (Aharonian, 1992), and some experts argue that up to 90% of semiconductor patents are not truly novel and therefore invalid (Taylor and Silberston, 1973). These problems make consistent enforcement difficult and patents in these fields are much more likely to be litigated (Bessen and Meurer, 2008). Surveys of managers in semiconductors and computers typically report that patents only weakly protect innovation. Levin et al. (1987) found that patents were rated weak at protecting the returns

intellectual property rights been available, these industries would have been even more dynamic. But we will argue that there is reason to think otherwise.

In fact, the software industry in the United States was subjected to a revealing natural experiment in the 1980s and 1990s. Through a sequence of court decisions, patent protection for computer programs was significantly strengthened. Evidence suggests that, far from unleashing a flurry of new innovative activity, the firms that acquired most of these patents actually reduced their R&D spending relative to sales (Bessen and Hunt, 2004).[3]

We maintain, furthermore, that there is nothing paradoxical about this outcome. For industries like software or computers, theory suggests that imitation may *promote* innovation and that strong patents (long-lived patents of broad scope) might actually *inhibit* it. Society and even the innovating firms themselves could well be served if intellectual property protection were more limited in such industries. Moreover, these firms might genuinely *welcome* competition and the prospect of being imitated.[4]

This is, we argue, because these are industries in which innovation is both *sequential* and *complementary*. By "sequential," we mean that each successive invention builds on the preceding one, in the way that the Lotus 1-2-3 spreadsheet built on VisiCalc, and Microsoft's Excel built on Lotus. And by "complementary," we mean that each potential innovator takes a different research line and thereby enhances the overall probability that a particular goal is reached within a given time. Undoubtedly, the many different approaches taken to voice-recognition software hastened the availability of commercially viable packages.

Imitation of a discovery can be socially desirable in a world of sequential and complementary innovation because it helps the imitator develop further inventions. And because the imitator may have valuable ideas not available to the original discoverer, the overall pace of innovation may thereby be enhanced. In fact, in a sequential setting, the inventor himself could be better off if others imitate and compete against him. Although imitation reduces the profit from his current discovery, it raises the probability of follow-on innovations, which improve his future profit. Of course, some form of protection would be essential for promoting innovation, even in a sequential setting, if there were no cost to entry and no limit to how quickly imitation could take place: in that case, imitators could immediately stream in whenever a new discovery was made, driving the inventor's revenues to zero. Throughout the article, however, we assume that entry requires investment in specialized capital, human or otherwise. Alternatively, we could assume that even if setup costs do not deter would-be imitators from entering, entry does not occur instantly, and so the original innovator has at least a temporary first-mover advantage. The ability of firms to generate positive (although possibly reduced) revenues when imitated accords well with empirical evidence for high-technology industries.[5]

to innovation, far behind the protection gained from lead time and learning-curve advantages. Patents in electronics industries were estimated to increase initiation costs by only 7% (Mansfield, Schwartz, and Wagner, 1981) or 7–15% (Levin et al., 1987). Taylor and Silberston (1973) found that little R&D was undertaken to exploit patent rights. As one might expect, diffusion and imitation are rampant in these industries. Tilton (1971) estimated the time from initial discovery to commercial imitation in Japanese semiconductors to be just over one year in the 1960s.

[3] As Sakakibara and Branstetter (2001) show, a similar phenomenon occurred in Japan. Starting in the late 1980s, the Japanese patent system was significantly strengthened. However, Sakakibara and Branstetter argue that there was no concomitant increase in R&D or innovation.

[4] Here are some examples in which firms have appeared to encourage imitation: When IBM announced its first personal computer in 1981, Apple Computer, then the industry leader, responded with full-page newspaper ads headed, "Welcome, IBM. Seriously." Adobe put Postscript and PDF format in the public domain, inviting other firms to be direct competitors for some Adobe products. Cisco (and other companies) regularly contribute patented technology to industry standards bodies, allowing any entrant to produce competing products. Finally, IBM and several other firms have recently donated a number of patents for free use by open source developers. The stated reason for this donation was to build the overall "ecosystem." See also Keely (2005).

[5] For example, consider that (1) the software industry is highly segmented (see Mowery, 1996), suggesting that specialized knowledge prevents a firm that is successful in one segment to move to another and (2) survey data from the electronics and computer industries (see Levin et al., 1987) indicate that "lead-time advantage" and "moving down the learning curve quickly" provide more effective protection than patents.

But whether or not an inventor without patent protection himself gains from being imitated, he is more likely (as we will show in one of our main results, Proposition 6) to be able to cover his cost of innovation in a sequential than a static environment, provided that it is socially desirable for him to incur this cost. This conclusion weakens the justification for intellectual property protection, such as patents, in sequential settings. Indeed, as we establish in Proposition 7, patents may actually reduce welfare: by blocking imitation, they may interfere with further innovation. Of course, patent defenders have a counterargument to this criticism: if a patent threatens to impede valuable follow-on innovative activity, the patent holder should have the incentive to grant *licenses* to those conducting the activity (thereby allowing innovation to occur). After all, if the follow-on R&D is worthwhile, she could share in its value by a suitably chosen licensing fee/royalty, thereby increasing her own profit (or so the argument goes).

A serious problem with this counterargument, however, is that it ignores the likely asymmetry between potential innovators in information about future profits. There is a large literature on patent licensing, but to our knowledge it has not addressed this asymmetry.[6] In our setting, if a patent holder is not as well-informed about a rival's potential future profits as the rival is himself, she may have difficulty setting a mutually profitable license fee, and so, as Proposition 7 shows, licensing may fail,[7] thereby jeopardizing subsequent innovation (of course, informational asymmetries about *current* profits, which we rule out for convenience, would only aggravate this problem).

In short, when innovation is sequential and complementary, standard conclusions about patents and imitation may get turned on their heads. Imitation becomes a *spur* to innovation, whereas strong patents become an *impediment*.

Sequential innovation has also been studied by Scotchmer (1991, 1996), Scotchmer and Green (1990), Green and Scotchmer (1996), and Chang (1995) for the case of a single follow-on innovation. Hunt (2004), O'Donoghue (1998), and O'Donoghue, Scotchmer, and Thisse (1998) study a single invention with an infinite sequence of quality improvements. We depart from this literature primarily in our model of competition. In our analysis, different firms' products at any given stage differ from one another.[8] That is, imitators do not produce direct "knockoffs," but rather differentiated products. This sort of differentiation is widely observed and is, of course, the subject of its own literature. But our main point here is that the different R&D paths behind these products permit innovative complementarities. Imitation then increases the "biodiversity" of the technology (see footnote 4), improving prospects for future innovation.

We proceed as follows. In Section 2, we introduce a static (nonsequential) model that, we claim, underlies the traditional justification for patents. We emphasize the point that, besides helping an inventor to cover his costs, an important role of patents is to encourage innovative activity on the part of others who would otherwise be inclined merely to imitate. Analytically, we show that (i) without patents, the equilibrium level of innovative activity is less than or equal to the optimum, and (ii) with patents, the level is greater than or equal to the optimum (Proposition 1). Despite the potential welfare ambiguity this result suggests, we argue that, on balance, patents are better than no patents: provided that the upper tail of the distribution of innovation values is sufficiently thick (which, we argue, is the empirically relevant case), expected welfare with patent protection exceeds that without it (Proposition 2). Not surprisingly, inventors themselves

[6] Some articles on patent licensing consider, as we do, the issue of licensing to one's own competitiors, including Katz and Shapiro (1985, 1986), Gallini (1984), and Gallini and Winter (1985). In these articles, however, the social loss from failure to license tends to derive from higher costs (because of decreasing returns to scale in monopoly production) and high consumer prices, rather than from reduced innovation.

[7] Although this phenomenon does not appear to have been analyzed in the existing patents literature, a closely related phenomenon has been examined in the literature on common pool resources such as oil reservoirs. As Wiggins and Libecap (1985) discuss, oil well owners can realize larger total profits if they contract to jointly manage production. But contract negotiations typically involve asymmetric information about future profits. The empirical evidence shows that contracting over oil production typically fails, despite industry-specific regulation designed to encourage it. In one case, only 12 out of 3,000 oil fields were completely covered by joint production agreements.

[8] This feature also figures prominently in the models in Dasgupta and Maskin (1987) and Tandon (1983).

are also better off with patent protection (Proposition 3). We also note that, in this static model, competition unambiguously diminishes the payoff of a prospective inventor (Proposition 4).

In Section 3, we modify the model to accommodate a potentially infinite sequence of inventions, each building on its predecessor. Because R&D now serves to raise the probability not only of the current invention but of future ones too, the equilibrium level of R&D will generally be higher than in the static model. However we show that the equilibrium level of innovative activity when there is no patent protection is still generally less than the optimum (Proposition 5). Even so, we establish that the *gap* between the equilibrium level and the optimal level is smaller than in the static model. Thus, equilibrium without patents is more nearly optimal with sequential than with static innovation (Proposition 6), implying that the case for intellectual property protection is correspondingly weaker. Indeed, under the same hypotheses for which we derived the *opposite* conclusion in the static model, the levels of social welfare and innovation when there is patent protection are actually *lower* on average than when there is not, (Proposition 7). Finally, we establish that, under somewhat more stringent assumptions, inventors *themselves* benefit from the absence of patent protection (Proposition 8) and may actually gain from being imitated, whether or not there is patent protection (Proposition 9), again in contrast to the static model. Most proofs are relegated to Appendix B.

2. The static model

■ We consider an industry consisting of two (*ex ante* symmetric) firms.[9] Each firm chooses whether or not to undertake R&D[10] to discover and develop an invention with (social) value v,[11] where v is known publicly and drawn *ex ante* from distribution with c.d.f. $F(v)$. We suppose that a firm's cost C of R&D is a random variable: with probability q, $C = c$, and with probability $1 - q$, $C = 0$.[12] A firm learns the realization of C before it decides whether to undertake R&D, but this information is *private*. Realizations are (statistically) independent across the two firms.

If a single firm undertakes R&D, the probability of successful innovation is p_1.[13] If both firms do R&D, the probability that at least one of them will succeed is p_2. We model the idea of complementarity—that having different firms pursue the same technological goal raises the probability that someone will succeed—by supposing that $p_2 > p_1$; each firm's probability of success is only p_1,[14] but because their research strategies are not perfectly positively correlated, the overall probability of success is higher. Of course, we must also have $p_2 < 2p_1$, that is,

$$p_1 < p_2 < 2p_1, \tag{1}$$

because, at best, the two firms' research strategies will be perfectly negatively correlated (in which case, $p_2 = 2p_1$).

We first consider the *socially efficient* R&D decisions for the two firms, that is, the decisions that a planner maximizing social welfare would direct them to take. Actually, because we

[9] Limiting the model to two firms in particular is a matter only of expositional convenience; all our results extend to any other finite number of firms. However, by assuming the number of firms is finite, we are implicitly supposing that there is a limit on how many firms are able to imitate a given invention, implying that an innovator's revenue need not be driven to zero in equilibrium. As we note above (and in footnote 5), we could alternatively assume that imitation takes time. Either way, it is important for our argument that, even without patent protection, the inventor be able to obtain *some* revenue from an invention.

[10] Throughout this article, a firm's R&D decision is a *binary* choice: to do it or not to do it. But all our results generalize to the case where the firm can vary the *intensity* of its R&D effort.

[11] There is no additional social value that accrues if both firms discover the innovation.

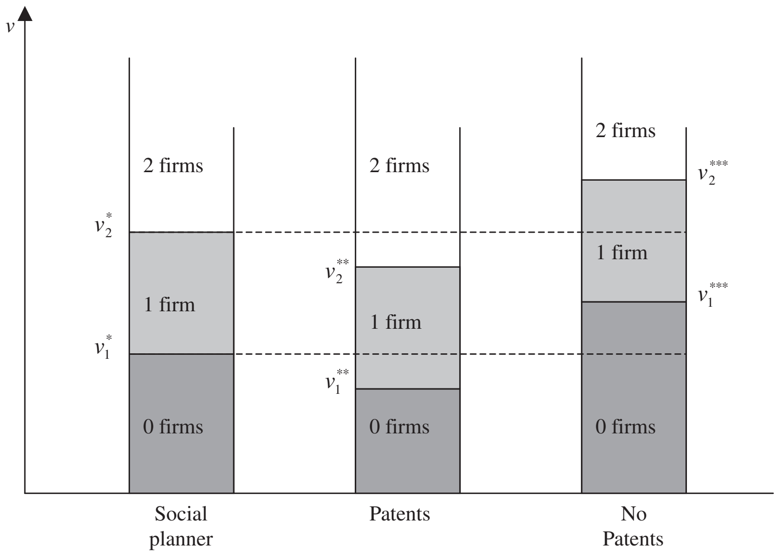
[12] We have chosen the smaller realization to be zero merely for analytic convenience; all our results hold with a positive lower cost.

[13] Our framework in this section is static, but if it were viewed as the reduced form of a dynamic setting, then p_1 could alternatively be interpreted as the discount factor corresponding to the time lag to innovation.

[14] That is, we rule out externalities in which one firm's R&D raises the other's chance of success. Of course, such spillovers are interesting and, in practice, important. But here they would serve only to strengthen our findings.

FIGURE 1

STATIC MODEL—R&D THRESHOLDS FOR HIGH-COST FIRMS



suppose that R&D costs are private information, the notion of social efficiency is not completely unambiguous. One possibility entails the planner first having the firms report their costs to him and then issuing them with R&D directives based on these costs. Another—more constrained—concept posits that the planner is unable to collect cost information, in which case the best he can do is to give each firm a *conditional* directive, for example, "Do R&D if your cost $C = 0$ but not if $C = c$."[15] As we will see, the latter notion makes comparisons with market equilibrium easier, and so we will adopt it henceforth.

Clearly, the planner will direct each firm to undertake R&D if its cost is zero. However, despite the complementarity in different R&D lines, the value of v may not be high enough to warrant the two firms each undertaking R&D if their costs are both c. Thus, in that case, the planner will want to treat the firms *asymmetrically* (even though they are inherently identical). That is, it will designate one of them, say firm 1, as the "aggressive" firm and have it undertake R&D with $C = c$ as soon as v exceeds some threshold v_1^*. Firm 2, however, will be directed not to undertake R&D with $C = c$ unless v is bigger than a higher threshold $v_2^*(> v_1^*)$.[16] See Figure 1 for a diagram of the various thresholds discussed in this section.

To calculate v_1^*, note that if firm 2's cost is c (which occurs with probability q)—so that, if v is only slightly bigger than v_1^*, firm 2 will not undertake R&D—then the gross expected social value of R&D by firm 1 is $p_1 v$, whereas if firm 2's cost is 0 (which occurs with probability $1 - q$)—implying that firm 2 *will* do R&D—the gross expected marginal contribution of firm 1's R&D is $(p_2 - p_1)v$. Hence, the expected net value of firm 1's R&D is $(qp_1 + (1 - q)(p_2 - p_1))v - c$, and so

[15] An alternative interpretation of this notion of constrained efficiency is that there is a *separate* planner for each firm and that the planners cannot communicate or coordinate with each other. Under this interpretation, their welfare-maximizing directives will constitute a "social Nash optimum" (see Grossman, 1977).

[16] If we adopt the two-planner interpretation (see footnote 15), there are three social Nash optima (SNOs): one in which firm 1 is aggressive, a second in which firm 2 is aggressive, and a third (but less efficient) SNO in which the two firms are treated symmetrically: for a range of v, each firm *randomizes* between doing and not doing R&D. Note that asymmetric treatment improves but does not fully solve the planner's coordination problem. For example, if $v \in (v_1^*, v_2^*)$, $C_1 = 0$, and $C_2 = c$, firm 1 will perform R&D and firm 2 will not, do so according to the planner's directives. But it would be more efficient if firm 2 also performed R&D.

$$(qp_1 + (1-q)(p_2 - p_1))\, v_1^* - c = 0,$$

that is,

$$v_1^* = \frac{c}{qp_1 + (1-q)(p_2 - p_1)}. \tag{2}$$

Similarly, we have

$$(p_2 - p_1)\, v_2^* - c = 0,$$

that is,

$$v_2^* = \frac{c}{p_2 - p_1}. \tag{3}$$

Turning from this normative analysis, we next examine the nature of equilibrium when the invention in question can be patented. We suppose that a firm with a patent can capture the *full* social benefit v of the invention.[17] If both firms undertake R&D, each has a probability $\frac{1}{2}p_2$ of getting the patent.[18]

Corresponding to the three possible social Nash optima of the efficiency analysis (see footnote 16), there are three possible equilibria when the invention is patentable: (i) one in which firm 1 is aggressive and firm 2 is passive, (ii) the mirror image, in which the firms' roles are reversed, and (iii) a symmetric equilibrium in which, for a range of values of v, the firms both randomize between doing and not doing R&D. For comparison with the planner's problem, we will focus on (i), which is strictly more efficient than (iii)—of course, we could just as easily have concentrated on (ii)

In equilibrium (i), each firm will undertake R&D if its cost is zero—it has nothing to lose by doing so. If v is not too big, then, from firm 1's point of view, the probability that the other firm does R&D is $1 - q$ and the probability that it does not is q. Hence, firm 1 will undertake R&D with $C = c$ if its expected revenue $(qp_1 + \frac{1}{2}(1-q)p_2)v$ exceeds its cost c, that is, if $v > v_1^{**}$, where

$$\left(qp_1 + \frac{1}{2}(1-q)\, p_2\right) v_1^{**} - c = 0,$$

or

$$v_1^{**} = \frac{c}{qp_1 + \frac{1}{2}(1-q)\, p_2}. \tag{4}$$

As for firm 2, it will not undertake R&D with $C = c$ unless v is sufficiently high for it to make a profit even when firm 1 does R&D too. That is, v must exceed v_2^{**}, where

$$\frac{1}{2} p_2 v_2^{**} - c = 0,$$

or

$$v_2^{**} = \frac{2c}{p_2}. \tag{5}$$

[17] This, of course, is a strong assumption. However, the incentive failures and monopoly inefficiencies that arise when it is not imposed are already well understood. The assumption is a simple way to abstract from these familiar distortions. It also accords with our approach of making suppositions favorable to patenting in order to draw stronger conclusions about patents' failures in Section 3.

[18] The total probability of discovery is p_2, and each firm has a one-half chance of making it first. This gets at the idea that patents have *breadth* (so that a patent holder can block the implementation of other firms' discoveries that are similar, but not identical, to its own). That is, only one firm can get a patent.

Finally, we investigate the nature of equilibrium when there is no patent protection. We assume that, without patents, if either firm is successful in making the discovery, the other can imitate costlessly[19] and that competition then drives each firm's gross revenue down to a fraction $s(0 < s \leq \frac{1}{2})$ of the total value v.[20]

Once again, there are three possible equilibria, and, as before, we will concentrate on the one in which firm 1 is aggressive and firm 2 is passive. In this equilibrium, either firm will undertake R&D if its cost is zero. Firm 1 will undertake R&D with $C = c$ if $v > v_1^{***}$, where

$$(qsp_1 + (1-q)\,s\,(p_2 - p_1))\,v_1^{***} - c = 0,$$

that is,

$$v_1^{***} = \frac{c}{qsp_1 + (1-q)\,s\,(p_2 - p_1)}. \tag{6}$$

Comparing (2) with (6), we see that $v_1^* < v_1^{***}$. This inequality corresponds to the classic incentive failure that the patent system is meant to address. When $v_1^* < v < v_1^{***}$, firm 1 cannot make a profit on its R&D investment without protection from imitation, despite the fact that such investment would be socially beneficial. A patent solves this problem by proscribing imitation. From (1), $\frac{1}{2}p_2 > p_2 - p_1$, and so from (2) and (4), $v_1^{**} < v_1^*$. Hence, with the prospect of patent protection, firm 1 will be willing to undertake R&D investment, provided this is socially worthwhile.

But even in a setting where $v > v_1^{***}$—so that R&D is profitable despite imitation—patents may well serve a useful purpose. This is because they can encourage *several* firms to go after the same innovation, which may be beneficial because of complementarity. In the absence of patent protection, firm 2 will earn expected profit

$$sp_2v - c, \tag{7}$$

if it decides to undertake R&D like firm 1. If, instead, it sits back and waits to imitate firm 1's invention, it can expect profit

$$sp_1v. \tag{8}$$

Hence, in equilibrium, it will invest in R&D only if (7) exceeds (8), that is, if $v > v_2^{***}$, where

$$s\,(p_2 - p_1)\,v_2^{***} - c = 0,$$

or

$$v_2^{***} = \frac{c}{s\,(p_2 - p_1)}. \tag{9}$$

But $v_2^* < v_2^{***}$, and so if v lies between v_2^* and v_2^{***}, we again have an incentive failure: although firm 1 will undertake R&D, firm 2 will merely imitate, despite the net social benefit from its investing too.

Here again patents come to the rescue. With the prospect of patent protection, firm 2 will undertake R&D provided that $v > v_2^{**}$. So, from (1), (2), and (5), $v_2^{**} < v_2^*$, implying that it will undertake R&D if such investment is socially desirable.

[19] In reality, even imitations that are complete knockoffs may involve substantial expenses, but our assumption gets at the idea that such expenses will typically be dwarfed by the innovating firm's R&D costs. Of course, some inventions are so difficult to reverse-engineer that trade secrecy adequately protects against imitation. But such inventions are not likely to be patented anyway, even if they could be, because of the patent system's disclosure requirements. To study potential shortcomings of the patent system, our focus in this article is on innovations that inventors would choose to patent if offered the opportunity.

[20] By assuming symmetry here, we simplify the computations a bit but, perhaps more importantly, we are *strengthening* the case for patents (if instead the innovating firm got the lion's share of the profit from the discovery, then safeguarding intellectual property would not matter as much). This will bolster our argument in Section 3, where we point out why patent protection may be socially undesirable.

Patents, therefore, accomplish more than merely protecting inventors from imitation; they encourage would-be imitators to invest in innovation *themselves*. Indeed, they create a risk of *over*investment in R&D: notice that v_1^{**} is strictly less than v_1^*, and v_2^{**} is strictly less than v_2^*.[21] Overinvestment can come about because when a firm decides to undertake R&D, it increases the probability that the discovery will be made, but also *diminishes* the other firm's chances of getting a patent. Because it doesn't take this negative externality into account, it is overly inclined to undertake R&D.

We summarize these results with the following proposition (see Figure 1 for a graphical summary):

Proposition 1. In the static model, the equilibrium level of R&D investment without patents is less than or equal to the social optimum. By contrast, the equilibrium level of R&D investment with patent protection is greater than or equal to the social optimum.

Observe that the possible overinvestment in R&D induced by patents could, in principle, be avoided if there were no complementarities of research across firms. Specifically, one could imagine awarding a firm an "*ex ante* patent," for example, the right to research and develop a vaccine against a particular disease.[22] Such protection would, of course, serve to prevent additional firms from attempting to develop the invention in question. But this would be efficient, provided that the firm with the patent had the greatest chance for success (which could be ensured, for example, by awarding the patent through an auction) and that the other firms would not enhance the probability or speed of development, that is, provided that they conferred no complementarity.

But even with the possibility of overinvestment, there is an important sense in which a regime with patents may be superior to one without them—if patents serve to encourage R&D projects with large returns, then the benefits from these projects can more than offset the welfare losses from overinvestment in more marginal projects. That is, despite potential welfare ambiguities, the standard economic doctrine that patents are a "good thing" *does* follow once we suppose the probability of high returns is not too much lower than that of low returns (indeed, this is more than just a theoretical hypothesis; see the empirical discussion in footnote 23).

To make this claim precise, imagine that the social (gross) value of innovation v is drawn from a distribution with twice-differentiable c.d.f. $F(v)$ and that, for some $k > 0$ and $\bar{v} > 0$, the following condition holds:

$$\textit{Upper Tail Condition}: \ \frac{d^2 F(v)}{dv^2} \geq -k \quad \textit{for all } v \in \left[\frac{c}{p_1}, \bar{v}\right].$$

For $\bar{v}$ sufficiently big and k sufficiently small, this condition ensures that the upper tail of the distribution does not fall off too quickly (the lower bound $\frac{c}{p_1}$ is chosen low enough so that the Condition applies in all the Propositions below where it is invoked). The Pareto and lognormal distributions, which are commonly found to fit distributions of returns to inventions and patent values in empirical research,[23] meet this requirement for appropriately chosen parameters.[24] (In Section 3, we shall offer another reason for invoking the Upper Tail Condition.) We can now state a formal justification for the standard economic doctrine favoring the patent system:

[21] The possibility that patents can give rise to excessive spending on R&D is well known from the patent-race literature; see Dasgupta and Stiglitz (1980) and Loury (1979).

[22] Wright (1983) and Shavell and Ypersele (2001) explore similar schemes.

[23] Early survey evidence suggested that the distribution of returns on patented inventions was highly skewed (Sanders, Rossman, and Harris, 1958). Using patent renewal data from Europe, Pakes and Schankerman (1984) found that the values of low-value patents could be fit with a Pareto distribution. More recent research has assessed the value of inventions in the upper tail by a variety of means and concluded that the distribution is fit well with a Pareto distribution function or a truncated lognormal distribution (only the upper tail of the lognormal distribution is observed); see Scherer and Harhoff, 2000 and Silverberg and Verspagen, 2004.

[24] This is established in Appendix A.

Proposition 2. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then expected net social welfare in the static model is higher with patent protection than without it.

Proof. See Appendix B.

The proof of Proposition 2 is somewhat involved, but the rough idea behind it is straightforward. To simplify, concentrate on the comparison between one and two firms. Patents will lead to *overinvestment*—two firms will invest when one would be more efficient—if v satisfies

$$(*) \qquad \frac{2c}{p_2} < v < \frac{c}{p_2 - p_1},$$

where $(*)$ comes from the inequality $v_2^{**} < v_2^{*}$. Similarly, there will be *underinvestment* (one firm investing rather than two) without patents if v satisfies

$$(**) \qquad \frac{c}{p_2 - p_1} < v < \frac{c}{s\,(p_2 - p_1)},$$

where $(**)$ derives from the inequality $v_2^{*} < v_2^{***}$. But the width of the interval in $(*)$ is strictly less than $\frac{c}{p_2 - p_1}$, whereas the width of the interval in $(**)$ is more than $\frac{c}{p_2 - p_1}$. Hence, provided that the probability density of v in the latter interval is not too much smaller that that in the former (which is ensured by the Upper Tail Condition), the gains from patents outweigh the losses.

Notice that patent licensing brings no advantage to a patent holder in this static model. Without licensing, the patent holder obtains a payoff of v. Were it instead to license the other firm, the firms could at best split a total of v. Thus, even if the patent holder set a license fee equal to the other firm's share of proceeds, it would still end up with at most v.[25]

Similarly, whether or not patent protection is available, a firm does not benefit from competition in this model:

Proposition 3. In the static model, a firm undertaking R&D is (weakly) worse off if it has a competitor.

Proof. See Appendix B.

Finally, just as patents are desirable for society in this static model, they are —even more clearly—good for the firms themselves:

Proposition 4. In the static model, a firm undertaking R&D is better off if there is patent protection than if there is no such protection.

Proof. See Appendix B.

Besides rationalizing the patent system, this simple static model captures the basic results of patent-race models such as Loury (1979) and Dasgupta and Stiglitz (1980). It also illustrates aspects of static models involving spillover complementarities, such as Spence (1984), who emphasizes the socially redundant R&D that can occur under patents. Our conclusions require reassessment, however, once we introduce sequentiality.

3. Sequential model

■ Let us now enrich the model to accommodate *sequential* innovation. Formally, consider an infinite sequence of potential inventions (indexed by $t = 1, 2, \ldots$), each of which has social

[25] This finding might change if the firms developed complementary innovations that could advantageously be cross-licensed; see Fershtman and Kamien (1992).

value v[26] drawn from c. d. f. $F(v)$.[27] To avoid the complications that arise when a new invention renders old discoveries obsolete, we suppose that v constitutes *incremental* value (i.e., an innovation can be thought of as an *improvement* that enhances the value of the initial invention).[28]

Complementarity between firms arises naturally in this sequential setting when some, but not all, of the technical information required for innovation diffuses freely or at low cost. Why doesn't the first inventor always make the subsequent discoveries itself? The usual answer is that the second firm possesses specialized information, such as expertise in a particular technology (see, for instance, Scotchmer, 1991, p. 31).[29] If all such information were freely available, the first inventor would indeed most likely make the subsequent innovations as well—it would have information about its own invention before other firms, and so it would be in the best position to make improvements. On the other hand, if the first inventor were able to keep all technical information about the innovation secret, then other firms could find making improvements extremely difficult, and so again we would expect the first inventor to continue alone. Of course, in reality, neither extreme generally holds. In a typical scenario, the commercial success of an innovative product reveals partial information that is useful to would-be subsequent innovators (perhaps because it facilitates reverse-engineering), who then apply their own particular expertise. For example, Lotus's success with a spreadsheet that included an integrated graphics feature revealed the large commercial importance of such a feature. Even though Lotus's source code remained secret, competitors who had already developed spreadsheet products of their own were then able to add integrated graphics without much additional cost.

Consistent with this view, we assume as in the static model that, in a setting without patents, firms can costlessly imitate each sequential innovation and that firms incurring the investment cost have an equal probability of developing the next innovation (so that the current invention's discoverer has no real advantage). However, we suppose that, absent licenses, a patent on an invention is sufficiently broad to block the next innovation in the sequence.[30] It is sometimes argued that, through the disclosure requirement, patents promote diffusion of technical information, and our assumptions admittedly neglect this effect. Still, both empirical evidence and theoretical argument call this potential advantage of patents into question (see Machlup and Penrose, 1950; Bessen, 2005).[31]

---

[26] Here v is the *direct* social value of an invention. In addition, there will be an *indirect* or *option* value deriving from the fact that the invention makes subsequent innovations possible.

[27] We are assuming that all innovations have the same value, that is, that v is drawn once and for all from F. This gets at the point that some innovative sequences are very fruitful (high v) and others not as beneficial (low v), that is, that there may be a great deal of correlation between the importance of successive innovations. But our findings would not change qualitatively if instead we supposed that there were independent drawings from F for each successive invention. Indeed, the only change to the formal argument below is to replace the continuation values W_1 and W_2—which, as the argument stands, depend on the once-and-for-all value of v—with their *expected* values (where the expectations are taken with respect to the cdf F).

[28] If instead new inventions *replaced* old ones, an innovation's social value (or a firm's profit from the innovation) could no longer be represented by a single parameter v but rather would become a *sum* of flow benefits that begin with discovery and end with the innovation's replacement. Because the replacement date would itself be endogenous, the analysis of replacement is rather more complex than that of improvement (see Hunt, 2004), O'Donoghue (1998), and O'Donoghue, Scotchmer, and Thisse (1998) for models of replacement in sequential innovation). Although for simplicity, we opt to model innovation as improvement, our major conclusions would not change if we invoked replacement instead (see footnote 39). Furthermore, our assumption of improvement may also be more consistent than replacement with technologically differentiated products.

[29] Note that this is the same kind of specialized information that gives rise to innovative complementarity in the static model. The presence of such specialized information is consistent with Sutton's finding (1998) that R&D-intensive industries that are not highly concentrated are associated with greater heterogeneity.

[30] Under patent law, an invention that builds on a patented invention infringes that patent, even if the second invention is patentable in its own right (Lemley, 1997).

[31] In brief, firms have no motivation to patent inventions that can be maintained as secrets, and so they will patent only inventions that would otherwise diffuse. Indeed, survey evidence finds that firms do not typically use patent disclosures as a valuable source of technical information. But in any case, the addition of a reverse-engineering cost to our model (which would weaken the promotion of innovation in the absence of patents) would not overturn our qualitative

Formally, there are, as before, two firms. For each invention t, a firm's cost of R&D is either c (with probability q) or zero (with probability $1-q$). Costs are independent across firms and inventions. For any t, if just one firm invests in R&D, then, following the static model, the probability that innovation $t+1$ is discovered conditional on the current invention t having already been discovered is p_1 (if invention t has not yet been discovered, then there is *no* chance that innovation $t+1$ will be developed). The corresponding conditional probability if both firms undertake R&D is p_2. The probabilities p_1 and p_2 should be thought of as the *cumulative* probabilities of developing innovation $t+1$. That is, firms may well try to develop it several times, and the p_is aggregate the probabilities of the multiple attempts.

Just as in the static model, a planner maximizing efficiency will treat the firms asymmetrically. As before, let us assume that firm 1 is the more aggressive of the two. Then, for efficiency, the planner will (i) direct each firm to undertake R&D for a given innovation if its cost for that period is zero; and (ii) direct firm 1 to undertake R&D if its cost is c and $v > v_1^\circ$, where

$$q\left(p_1 v_1^\circ - c + p_1 W_1^\circ\right) + (1-q)\left((p_2 - p_1)\, v_1^\circ - c + (p_2 - p_1)\, W_1^\circ\right) = 0 \tag{10}$$

and

$$\begin{aligned} W_1^\circ &= q^2\left(p_1 v_1^\circ - c + p_1 W_1^\circ\right) + q\,(1-q)\left(p_2 v_1^\circ - c + p_2 W_1^\circ\right) \\ &\quad + (1-q)\,q\left(p_1 v_1^\circ + p_1 W_1^\circ\right) + (1-q)^2\left(p_2 v_1^\circ + p_2 W_1^\circ\right) \\ &= \frac{(q p_1 + (1-q)\, p_2)\, v_1^\circ - qc}{1 - q p_1 - (1-q)\, p_2}, \end{aligned} \tag{11}$$

and W_1° is the expected long-run social payoff when the value of each innovation is v_1° and both firms invest if their costs are zero but, of the two, only firm 1 invests if its cost is c.

Equation (10) incorporates the idea that R&D makes possible not only the next invention but also innovations after that: if, for example, firm 1 does R&D and firm 2 does not, then there is a probability p_1 that the next invention (worth v_1°) will be discovered and thus a probability p_1 that the subsequent sequence of innovations (whose expected social value is W_1°, if each innovation is worth $v = v_1^\circ$) have a chance of being discovered. To understand equation (11), note that if only firm 1 invests when $C = c$, then the terms on the right-hand side led by q^2, $q(1-q)$, $(1-q)q$, and $(1-q)^2$ correspond, respectively, to the events $(C_1 = c, C_2 = c)$, $(C_1 = c, C_2 = 0)$, $(C_1 = 0, C_2 = c)$, and $(C_1 = 0, C_2 = 0)$.

From (10) and (11), we obtain (see Figure 2 for the dynamic model thresholds)

$$v_1^\circ = \frac{c\left(1 - 2q\,(1-q)\, p_1 - (1-q)^2\, p_2\right)}{q p_1 + (1-q)(p_2 - p_1)}. \tag{12}$$

Finally, the planner will (iii) direct firm 2 to undertake R&D if its cost is c and $v > v_2^\circ$, where

$$(p_2 - p_1)\, v_2^\circ - c + (p_2 - p_1)\, W_2^\circ = 0$$

$$W_2^\circ = p_2 v_2^\circ - 2qc + p_2 W_2^\circ = \frac{p_2 v_2^\circ - 2qc}{1 - p_2},$$

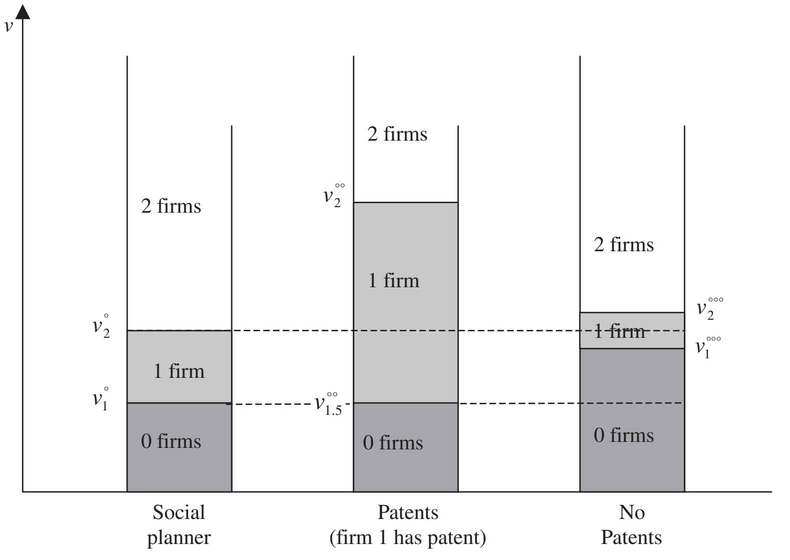
and W_2° is the expected long-run social payoff when the value of an innovation is v_2° and both firms always invest in R&D. Thus,

$$v_2^\circ = \frac{c\,(1 - p_2 + 2q\,(p_2 - p_1))}{p_2 - p_1}. \tag{13}$$

results, provided that that cost were not too large. Of course, there are many inventions for which the reverse-engineering costs are high. But those are precisely the inventions we would not expect to see patented anyway, and so they fall outside the scope of an article attempting to assess the effect of patents.

FIGURE 2

DYNAMIC MODEL—R&D THRESHOLDS FOR HIGH-COST FIRMS



Next, we look at behavior in the dynamic model with *no patent protection*, where we continue to focus on the equilibrium in which firm 1 is aggressive and firm 2 passive. As in the static model, if just one firm undertakes R&D then the other firm gets share s of the gross expected profit simply by imitating any invention arising from the investment (without necessarily conducting R&D itself). Clearly, a firm will undertake R&D if its cost is zero. Firm 1 will undertake R&D with a cost of c if $v > v_1^{\circ\circ\circ}$, where

$$\begin{aligned} q\left(sp_1 v_1^{\circ\circ\circ} - c + p_1 W_1^{\circ\circ\circ}\right) + (1-q)\left(sp_2 v_1^{\circ\circ\circ} - c + p_2 W_1^{\circ\circ\circ}\right) \\ = (1-q)\left(sp_1 v_1^{\circ\circ\circ} + p_1 W_1^{\circ\circ\circ}\right) \end{aligned}$$

$$\begin{aligned} W_1^{\circ\circ\circ} &= 2q(1-q)\left(sp_1 v_1^{\circ\circ\circ} + p_1 W_1^{\circ\circ\circ}\right) + (1-q)^2\left(sp_2 v_1^{\circ\circ\circ} + p_2 W_1^{\circ\circ\circ}\right) \\ &= \frac{\left(2q(1-q)sp_1 + (1-q)^2 sp_2\right) v_1^{\circ\circ\circ}}{1 - 2q(1-q)p_1 - p_2(1-q)^2}, \end{aligned}$$

and $W_1^{\circ\circ\circ}$ is firm 1's expected long-run payoff when the value of an innovation is $v_1^{\circ\circ\circ}$ and each firm invests in R&D only when its cost is zero; thus,

$$v_1^{\circ\circ\circ} = \frac{c\left(1 - 2q(1-q)p_1 - (1-q)^2 p_2\right)}{s\left(qp_1 + (1-q)(p_2 - p_1)\right)}. \tag{14}$$

Firm 2 will also undertake R&D with cost c if $v > v_2^{\circ\circ\circ}$, where

$$sp_2 v_2^{\circ\circ\circ} - c + p_2 W_2^{\circ\circ\circ} = sp_1 v_2^{\circ\circ\circ} + p_1 W_2^{\circ\circ\circ}$$

$$W_2^{\circ\circ\circ} = sp_2 v_2^{\circ\circ\circ} - qc + p_2 W_2^{\circ\circ\circ} = \frac{sp_2 v_2^{\circ\circ\circ} - qc}{1 - p_2},$$

and $W_2^{\circ\circ\circ}$ is firm 2's expected long-run payoff when the value of each innovation is $v_2^{\circ\circ\circ}$ and both firms always invest in R&D[32]; thus,

$$v_2^{\circ\circ\circ} = \frac{c\,(1 - p_2 + q\,(p_2 - p_1))}{s\,(p_2 - p_1)}. \tag{15}$$

From (12) and (14), we have $v_1^{\circ\circ\circ} > v_1^{\circ}$. From (13) with (15) and because $s < \frac{1}{2}$, we have $v_2^{\circ\circ\circ} > v_2^{\circ}$. Hence, we see that, as in the static model, there is too little R&D in equilibrium relative to efficiency:

Proposition 5. In the sequential model, the equilibrium level of R&D investment in a regime without patents is less than or equal to the social optimum.

Although equilibrium without patents remains inefficient in the dynamic model, there is an important sense in which the inefficiency is smaller than that in the static model. To begin with, notice, from (6) and (14), that $v_1^{\circ\circ\circ} < v_1^{***}$ and, from (9) and (15), that $v_2^{\circ\circ\circ} < v_2^{***}$. That is, the expected equilibrium levels of R&D in the dynamic model are higher than those in the static model (as we would anticipate since, in the sequential setting, investing in R&D raises the probability not only of the next innovation but of subsequent innovation).

Still, the fact that there is more R&D in the dynamic model does not by itself settle the matter that the dynamic equilibrium is more efficient. After all, efficiency *also* entails a higher expected level of R&D in the sequential than the static model: from (2) and (12), $v_1^{\circ} < v_1^{*}$, and, from (3) and (13) $v_2^{\circ} < v_2^{*}$. Nevertheless, under the same hypothesis invoked to show that patents are more efficient than no patents in the static model (Proposition 2), we can show that equilibrium without patents is more nearly efficient in the sequential than the static model:

Proposition 6. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then, for $i = 1, 2$, the likelihood that firm i behaves inefficiently in the sequential model without patents is lower than that in the static model without patents.

Proof. See Appendix B.

To get a feel for why Proposition 6 holds, notice that the probability that firm 1's behavior is inefficient in the static-model equilibrium without patents is the probability that $v \in [v_1^{*}, v_1^{***}]$, whereas the corresponding probability in the dynamic-model equilibrium without patents is the probability that $v \in [v_1^{\circ}, v_1^{\circ\circ\circ}]$. But the interval $[v_1^{\circ}, v_1^{\circ\circ\circ}]$ is smaller than $[v_1^{*}, v_1^{***}]$, and the former also lies below the latter. Hence if the density of $F(v)$ does not drop off too rapidly as v increases, the probability that $v \in [v_1^{\circ}, v_1^{\circ\circ\circ}]$ is smaller than the probability that $v \in [v_1^{*}, v_1^{***}]$. The argument is similar—although slightly more complicated—for firm 2.

Equilibrium with patents is more complicated in the dynamic than the static model. To begin with, when the model is dynamic, we must, distinguish between the R&D behavior of the two firms *before* a patent is obtained on the first invention and their behavior *after* this patent is obtained (in the static model, by contrast, there is obviously no R&D after the patent is obtained). Furthermore, we have to consider the levels at which patent holders will set license fees, an issue that also does not arise in the static model.

We suppose that a license gives a non-patent holder the possibility of developing the next innovation without infringing the patent. Moreover we assume that the license is written in such a way that there is no dissipation of profit for the *current* innovation (because the license limits the total quantity sold). Finally, we suppose that the property rights to the next innovation accrue to the original patent holder (so that the licensing arrangement must award a non-patent holder

[32] At $v = v_2^{\circ\circ\circ}$, firm 2 is indifferent between investing and not investing when its cost is c. Hence, we could alternatively (and equivalently) have defined $W_2^{\circ\circ\circ}$ as 2's expected long-run payoff when firm 1 always invests and firm 2 invests only if its cost is zero.

a sufficient share of the profit from that innovation if that firm is to be willing to incur the R&D cost).

As we have done all along, we shall focus on the equilibrium in which firm 1 is aggressive and firm 2 is passive.[33] If inventions are protected by patents, each firm will invest if its cost is zero and neither firm yet has a patent. A firm will also invest if its cost is zero and it has a patent. If neither firm yet has a patent, firm 1 will invest if its cost is c and $v > v_{.5}^{\circ\circ}$, where

$$qp_1\left(v_{.5}^{\circ\circ} + W_{.5}^{\circ\circ}\right) + \frac{1}{2}(1-q)\,p_2\left(v_{.5}^{\circ\circ} + W_{.5}^{\circ\circ}\right) - c = 0 \tag{16}$$

$$\begin{aligned} W_{.5}^{\circ\circ} &= 2q(1-q)p_1\left(v_{.5}^{\circ\circ} + W_{.5}^{\circ\circ}\right) + (1-q)^2 p_2\left(v_{.5}^{\circ\circ} + W_{.5}^{\circ\circ}\right) \\ &= \frac{\left(2q(1-q)p_1 + (1-q)^2 p_2\right) v_{.5}^{\circ\circ}}{1 - 2q(1-q)p_1 - (1-q)^2 p_2}, \end{aligned} \tag{17}$$

and $W_{.5}^{\circ\circ}$ is the expected long-run payoff of a firm that holds a patent when the value of an innovation is $v_{.5}^{\circ\circ}$, if it conducts R&D only when its cost for the next innovation is zero and licenses the other firm only when that firm's cost is zero, that is, it sets the license fee so high that only a low-cost firm will accept. More specifically, when only a low-cost firm is licensed, the patent holder gets the entire joint profit for itself.[34] This outcome is reflected in formula (17): note that $2q(1-q)p_1 v_{.5}^{\circ\circ} + (1-q)^2 p_2 v_{.5}^{\circ\circ}$ is the total expected surplus generated when each firm does R&D if and only if its cost is zero. From (16) and (17), we have

$$v_{.5}^{\circ\circ} = \frac{\left(1 - 2q\,(1-q)\,p_1 - (1-q)^2\,p_2\right)c}{qp_1 + \frac{1}{2}(1-q)\,p_2}. \tag{18}$$

If neither firm yet has a patent, firm 2 will also invest in R&D when its cost is c, if $v > v_{.75}^{\circ\circ}$, where

$$\frac{1}{2}p_2\left(v_{.75}^{\circ\circ} + W_{.75}^{\circ\circ}\right) - c = 0,$$

$$W_{.75}^{\circ\circ} = \frac{\left(2q\,(1-q)\,p_1 + (1-q)^2\,p_2\right) v_{.75}^{\circ\circ}}{1 - 2q\,(1-q)\,p_1 - (1-q)^2\,p_2},$$

and $W_{.75}^{\circ\circ}$ is the expected long-run payoff of a patent holder, when the value of an innovation is $v_{.75}^{\circ\circ}$, if it conducts R&D only when its cost is zero and licenses the other firm only when that firm's cost is zero. Hence,

$$v_{.75}^{\circ\circ} = \frac{\left(1 - 2q\,(1-q)\,p_1 - (1-q)^2\,p_2\right)c}{\frac{1}{2}p_2}. \tag{19}$$

A firm with a patent will invest in R&D with a cost of c (as opposed to just licensing) if $v > v_{1.5}^{\circ\circ}$, where

[33] This asymmetry pertains only to the situation *before* one of the firm obtains a patent; *after* that point, the only asymmetry is between the patent holder and the other firm.

[34] Implicit in this conclusion is the assumption that the patent holder has all the bargaining power in setting the license fee. That is, it can make a take-it-or-leave-it offer to the other firm. But if we assumed instead that the other firm shares in the bargaining power, none of our qualitative conclusions would change.

$$q\, p_1 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right) + (1-q)p_2 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right) - c = (1-q)p_1 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right)$$

$$\begin{aligned} W_{1.5}^{\circ\circ} &= q^2 \left(p_1 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right) - c\right) + q\,(1-q) \left(p_2 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right) - c\right) \\ &\quad + (1-q)\, q p_1 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right) + (1-q)^2\, p_2 \left(v_{1.5}^{\circ\circ} + W_{1.5}^{\circ\circ}\right) \\ &= \frac{(qp_1 + (1-q)\,p_2)\, v_{1.5}^{\circ\circ} - qc}{1 - qp_1 - (1-q)\,p_2}, \end{aligned}$$

and $W_{1.5}^{\circ\circ}$ is the expected long-run payoff of a patent holder when the value of each innovation is $v_{1.5}^{\circ\circ}$, if it invests when its cost is c and licenses the other firm only when that firm's cost is zero. Hence,

$$v_{1.5}^{\circ\circ} = \frac{\left(1 - 2q\,(1-q)\,p_1 - (1-q)^2\, p_2\right) c}{qp_1 + (1-q)(p_2 - p_1)}. \tag{20}$$

Notice, from (19) and (20), that we have presumed that the v-threshold at which firm 2 with cost c does R&D when neither firm has a patent is less than that at which a firm with a patent does R&D when its cost is c. However, it is readily verified that (19) is indeed less than (20) provided that q is sufficiently small—and the latter is a hypothesis of the propositions we are coming to.

A firm with a patent will license the other firm (and perform R&D itself) even if that other firm's cost is c provided that $v > v_2^{\circ\circ}$, where

$$p_2 \left(v_2^{\circ\circ} + W_2^{\circ\circ}\right) - c = qp_1 \left(v_2^{\circ\circ} + W_2^{\circ\circ}\right) + (1-q)p_2 \left(v_2^{\circ\circ} + W_2^{\circ\circ}\right), \tag{21}$$

$$W_2^{\circ\circ} = p_2 \left(v_2^{\circ\circ} + W_2^{\circ\circ}\right) - (1+q)\,c = \frac{p_2 v_2^{\circ\circ} - (1+q)\,c}{1 - p_2}, \tag{22}$$

and $W_2^{\circ\circ}$ is the expected long-run payoff of a patent holder, when the value of each innovation is $v_2^{\circ\circ}$, if it always invests in R&D and always licenses the other firm. By licensing the other firm even in the event that it has a high cost of R&D, the patent holder raises the probability of discovery from p_1 to p_2 in that event, but must reduce its license fee by c (and, because costs are private information, it must do so even when the other firm's cost is low[35]).[36] From (21) and (22), we have

$$v_2^{\circ\circ} = \frac{\left(1 - (q+q^2)p_1 - (1-q-q^2)p_2\right) c}{q\,(p_2 - p_1)}. \tag{23}$$

Once again, we have presumed the ranking of threshold values: implicit in (20) and (23) is the presumption that $v_{1.5}^{\circ\circ} < v_2^{\circ\circ}$. That this is indeed the case is easily shown as long as q is sufficiently small, which the following result assumes:

[35] Thus, in this case, the other firm will earn a rent of c if its cost is low. Moreover, to ensure that the firm will actually undertake R&D when its cost is high, the patent holder must cede it some of the profit if the discovery is made (which the patent holder can take back in the form of a higher license fee).

[36] We have been assuming that a firm wishing to build on a patented invention must first obtain a license from the patent holder (see footnote 33). But let us imagine that the firm instead goes ahead and attempts to develop the next innovation without a license. If taking this next step entails first marketing some imitation of the patented item, then the firm can expect to be sued for patent infringement and so presumably will not proceed in this way. But suppose that it can potentially move to the next generation without direct market experience in the current generation. In that case, if it is successful, it can apply for a license *ex post* (see Scotchmer, 1996 and 2005, for treatments of *ex post* licensing). Notice, however, that the patent holder will then set a license fee that appropriates all of the firm's profit from its invention. Furthermore, in contrast to *ex ante* licensing, the patent holder will not reduce this fee by c, even if that was the firm's R&D cost, because this expenditure has already been *sunk*. Thus, a firm with R&D cost c will do worse by waiting for *ex post* licensing (the analysis would be a bit more involved if the firm had some bargaining power in determining the license fee, but as footnote 34 points out, our qualitative conclusions would remain the same.)

Proposition 7. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then there exists $\bar{q} > 0$ such that expected net social welfare in the dynamic model is higher in equilibrium without patent protection than in equilibrium with such protection provided that $q < \bar{q}$.

Proof. See Appendix B.

Remark. Note that this result holds even for small (but positive) values of s, that is, even with substantial dissipation of rents in the no-patent case. Moreover, one can show that it does not depend on the assumption that the development cost for a low-cost type is zero; it holds for a positive lower cost as well. Finally, notice that we have considered private information only regarding development costs. But because rivals are presumably using different technologies, private information about production costs is also likely. This would lead to a broader range of circumstances in which patents would generate lower social welfare.

To get a feel for why Proposition 7 holds, let us suppose that higher values of v are much more likely than lower values (which is considerably stronger than the actual hypothesis). If there is no patent protection, then, as derived in (15), firm 2 will undertake R&D with cost c as long as

$$v > \frac{c\,(1 - p_2 + q\,(p_2 - p_1))}{s\,(p_2 - p_1)} \left(= v_2^{\circ\circ\circ}\right). \tag{$\circ$}$$

In a regime with patent protection, by contrast, once one firm has a patent, it will refrain from setting the license fee low enough so that the other firm will undertake R&D with cost c, if the patent holder's expected benefit from the additional R&D, $q(p_2 - p_1)(v + W)$, is less than the loss in fee revenue c from setting the lower fee, that is, if

$$v < \frac{\left(1 - \left(q + q^2\right) p_1 - \left(1 - q - q^2\right) p_2\right) c}{q\,(p_2 - p_1)} \left(= v_2^{\circ\circ}\right). \tag{$\circ\circ$}$$

But if q is small, then the right-hand side of ($\circ\circ$)exceeds the right-hand side of ($\circ$). Thus, for $v \in [v_2^{\circ\circ\circ}, v_2^{\circ\circ}]$, society gets more R&D investment when there is no patent protection than when such protection is available.[37] And since there is underinvestment anyway in the absence of patent protection, this additional R&D is socially beneficial, that is, for this interval, "no patents" are better than "patents." Now, of course, for lower values of v, the comparison can go the other way. But if higher values are sufficiently more likely than lower values, we can conclude that having no patent protection is better *on average*.[38]

This conclusion remains even if we take into account the overall effect of patents on R&D. Indeed, suppose we focus purely on the incentives of firms to undertake R&D *before* a patent has been obtained. Then, we see from (12) and (18), that $v_{.5}^{\circ\circ} < v_1^{\circ}$ and, from (13) and (19) that $v_{.75}^{\circ\circ} < v_2^{\circ}$ if q is sufficiently small. That is, firms will tend to *overinvest*, as in the static model (see Proposition 1). But this overinvestment result pertains only to the first innovation, whereas the underinvestment logic applies to all subsequent innovations. Hence, under the very condition that makes the case for patents compelling in the static model, overinvestment is more than counterbalanced in the dynamic model by the constraint on R&D imposed by the patent holder's unwillingness to license a high-cost firm.

If $s \approx \frac{1}{2}$—so that there is little profit dissipation—then most of the net social welfare generated without patents accrues to the firms themselves. Hence, in that case, we can conclude that *ex ante* the firms themselves will prefer that innovations not be protected by patents:

[37] If, by contrast, q is near 1, then the patent holder will set a relatively low license fee, and, in fact, licensing will occur at nearly the socially efficient rate. Thus, in this case, patents will dominate the absence of patent protection from a social standpoint.

[38] Boldrin and Levine (2004) show that the inefficiencies of patents are aggravated if a firm has to obtain licenses from *different* patent holders in order to innovate.

Proposition 8. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then each firm's *ex ante* expected profit in the dynamic model is higher in equilibrium *without* patent protection than in equilibrium *with* protection, provided that q is sufficiently small and s is near enough $\frac{1}{2}$.

Proof. See Appendix B.

Remark. The conclusion of Proposition 8 depends critically on neither firm having a patent *ex ante*. It is evident that once a firm obtains such protection, it will definitely prefer to keep it and exercise monopoly power in licensing, even though this may reduce innovation and therefore total profit. But this observation prompts the question why, *before* any discovery has been made, the firms do not enter an agreement to ensure that licensing always occurs regardless of which of them ends up getting the patent on the first discovery. Such an arrangement would be desirable both for the firms and society. It would, however, require the two firms to know of one another's existence before the industry has even begun, which could be a very tall order indeed.

Finally, we turn to a fourth important difference between the static and sequential models: whether or not an innovating firm *itself* benefits from competition and being imitated. In Proposition 3, we showed that a firm undertaking R&D clearly loses from competition and imitation in the static model. By contrast, in the sequential model we have:

Proposition 9. Assume that the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small and that

$$p_2 > \frac{2p_1}{1+p_1}. \tag{24}$$

If s is near enough $\frac{1}{2}$, then in the sequential model a firm gains from having a competitor and being imitated, whether or not there is patent protection.

Proof. See Appendix B.

Remark 1. Proposition 9 is a formal justification for the dictum that "competition expands the market" and explains Apple's welcoming greeting to IBM (see footnote 4).

Remark 2. Condition (24) holds if, for example, $p_1 = p$ and $p_2 = 1 - (1 - p)^2$, that is, if the two firms' chances of success are statistically independent.

Remark 3. Propositions 6–9 suggest another reason beyond empirical realism for invoking the Upper Tail Condition.[39] As we noted in Section 2, the welfare comparison between the patent and no-patent regimes is ambiguous in the absence of any assumption about the distribution of returns: the absence of patents leads to underinvestment in R&D, but, in the static model, patents induce overinvestment. Because something like the Upper Tail Condition is needed to generate the standard conclusion that, on balance, patents are desirable in a static setting, it is of interest to see that this same condition invoked in a sequential setting leads to quite different results: the no-patent regime is now closer to efficiency than in the static model; patents may generate less innovation than in the absence of patents; and imitation may be welcomed by inventors themselves.

Having a competitor may be advantageous to a would-be inventor because, for v big enough (which, given the Upper Tail Condition, is sufficiently likely), this other firm will undertake R&D too and thereby raise the probability of discovery from p_1 to p_2, which improves the inventor's

[39] We have established these propositions under the hypothesis that new inventions enhance rather than replace old inventions, but the qualitative contrast between the static and dynamic models—on which Proposition 6 turns—and the unwillingness of a patent holder to license high-cost competitors—on which Propositions 7–9 turn—do not depend on this distinction. Hence, the propositions can be shown to hold for replacement.

future profit. Of course, there is also the drawback that the competitor obtains a share of this profit. But if s is not too small, this latter effect is outweighed by the former.

4. Conclusion

■ Intellectual property appears to be an area in which results that seem secure in a static model may be overturned in a sequential setting. The prospect of being imitated inhibits inventors in a static world; in a dynamic world, imitators can provide benefit to both the original inventor and to society more generally. Patents may be desirable to encourage innovation in a static world, but they are less important in a sequential setting, where they may actually inhibit complementary innovation.

The static-sequential distinction is more than just a theoretical nicety. Indeed, it may help resolve a puzzle emanating from the U.S. natural experiment in software patents. Strikingly, the firms that obtained the most software patents (largely firms in the computer and electronics hardware industries) actually *reduced* their R&D spending relative to sales after patent protection was strengthened (Bessen and Hunt, 2004). This behavior is difficult to reconcile with the static model, in which the prospect of patents should *encourage* R&D, but is quite consistent with the sequential model and specifically Proposition 7.

Thus we would suggest a cautionary note about intellectual property protection. The reflexive view that "stronger is better" could well be too extreme; rather, a balanced approach seems called for. The ideal patent policy limits "knockoff" imitation, but allows developers who make similar, but potentially valuable complementary contributions. In this sense, copyright protection for software programs (which has gone through its own evolution over the last decade) may have achieved a better balance than patent protection. In particular, industry participants complain that software patents have been too broad (and patented discoveries too obvious), leading to holdup problems (USPTO 1994, Oz 1998). Systems that limit patent breadth, such as in the Japanese system before the late 1980s, may offer a better balance.[40]

Appendix A

■ **Proof that the Pareto distribution satisfies the Upper Tail Condition.** The Pareto distribution is

$$F(v) = 1 - \left(\frac{v_0}{v}\right)^{\alpha}, \quad v_0 \le v,\ 0 < \alpha,$$

so that

$$\frac{d^2F(v)}{dv^2} = -\alpha(1+\alpha)\frac{v_0^{\alpha}}{v^{2+\alpha}}.$$

Since $\lim_{\alpha\to 0}\frac{d^2F(v)}{dv^2} = 0 \ge -k$, we see that, given k and $\bar{v}$, the Pareto distribution satisfies the Upper Tail Condition for k and $\bar{v}$ for all v in the defined domain, provided that α is small enough and $v_0 = \frac{c}{p_1}$.

□ **Proof that the lognormal distribution satisfies the Upper Tail Condition for values of v above the median.** For the lognormal distribution with parameters μ and σ,

$$\frac{dF}{dv}(v) = \frac{Exp[-(\ln(v)-\mu)^2/2\sigma^2]}{\sigma v\sqrt{2\pi}}$$

and the median value of v is e^{μ}. Thus,

$$\frac{d^2F(v)}{dv^2} = -\frac{Exp[(\ln(v)-\mu)^2/2\sigma^2]\left(\ln(v)-\mu+\sigma^2\right)}{\sigma^3 v^2\sqrt{2\pi}}.$$

[40] In a review of the literature, Gallini and Scotchmer (2001) conclude, "Thus, with some caution, we can extract from the literature a case for broad (and short) patents. Broad patents can serve the public interest by preventing duplication of R&D costs, facilitating the development of second generation products, and protecting early innovators who lay a foundation for later innovators. However, these benefits disappear if licensing fails." Our model establishes that broad patents may be especially harmful if licensing fails and that there is good reason to expect such a failure.

It is then straightforward to show that

$$\left.\frac{d^2F(v)}{dv^2}\right|_{v=e^{\mu}} = -\frac{e^{-2\mu}}{\sigma\sqrt{2\pi}} \leq \frac{d^2F(v)}{dv^2}, \quad \forall v \in [e^{\mu}, \infty].$$

Since, for any $k > 0$, $\lim_{\sigma\to\infty} \cdot \frac{d^2F(v)}{dv^2}|_{v=e^{\mu}} = 0 \geq -k$, the lognormal distribution satisfies the Upper Tail Condition for all v above the median, given a large enough value of σ.

Appendix B

■ Proofs of the Propositions

Proposition 2. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then expected net social welfare in the static model is higher with patent protection than without it.

Proof. The expected difference in welfare between having patents and not having patents as this relates to firm 1's participation is:

$$\int_{v_1^{**}}^{v_1^{*}} \left[q(1-q)((p_2-p_1)v-c)+q^2(p_1v-c)\right]dF(v)+\int_{v_1^{*}}^{v_1^{***}} \left[q(1-q)((p_2-p_1)v-c)+q^2(p_1v-c)\right]dF(v), \tag{B1}$$

where the first integral in (B1) is negative because of overinvestment under patent protection (the fact that $v_1^{**} < v_1^{*}$) and the second integral is positive because of underinvestment without patent protection (the fact that $v_1^{*} < v_1^{***}$). Summing the two integrals, we must show that for k small enough and $\bar{v}$ big enough,

$$\int_{v_1^{**}}^{v_1^{***}} [av-c]\,dF(v) \tag{B2}$$

is positive, where $a = qp_1 + (1-q)(p_2-p_1)$. We can rewrite (B2) as

$$\left(F\left(v_1^{***}\right)-F\left(v_1^{**}\right)\right)\int_{v_1^{**}}^{v_1^{***}} (av-c)\,dF_1(v) \tag{B3}$$

where $F_1(v) = \frac{F(v)-F(v_1^{**})}{F(v_1^{***})-F(v_1^{**})}$, so that

$$F_1\left(v_1^{**}\right)=0 \quad \text{and} \quad F_1\left(v_1^{***}\right)=1. \tag{B4}$$

To show that (B2) is positive, it suffices, from (B3), to show that

$$\int_{v_1^{**}}^{v_1^{***}} (av-c)\,dF_1(v) > 0 \tag{B5}$$

After integration by parts, the left-hand side of (B4) can be written as

$$av_1^{***} - c - a\int_{v_1^{**}}^{v_1^{***}} F_1(v)\,dv. \tag{B6}$$

From the hypothesis of Proposition 2, we can choose k sufficiently small and $\bar{v}$ sufficiently big that

$$F_1(v) \leq \left(\frac{v}{v_1^{***}-v_1^{**}} - \frac{v_1^{**}}{v_1^{***}-v_1^{**}}\right) + \frac{v_1^{**}}{2\left(v_1^{***}-v_1^{**}\right)} \text{ for all } v \in \left[v_1^{**}, v_1^{***}\right] \tag{B7}$$

From (B7), (B6) exceeds

$$av_1^{***} - c - \frac{1}{v_1^{***}-v_1^{**}}\left[\frac{a}{2}\left(\left(v_1^{***}\right)^2-\left(v_1^{**}\right)^2\right)\right] + \frac{av_1^{**}}{2} = \frac{av_1^{***}}{2} - c,$$

which is positive because $\frac{a}{2}v_1^{***} - c = \frac{c}{2s} - c > 0$.

The expected difference in welfare between having patents and not having patents as this relates to firm 2's participation is

$$q\int_{v_2^{**}}^{v_2^{*}} ((p_2-p_1)v-c)\,dF(v) + q\int_{v_2^{*}}^{v_2^{***}} ((p_2-p_1)v-c)\,dF(v), \tag{B8}$$

where the first integral in (B8) is negative because of overinvestment by firm 2 with patents (the fact that $v_2^{**} < v_2^{*}$) and the second is positive because of underinvestment by firm 2 without patents (the fact that $v_2^{*} < v_2^{***}$). Summing the two integrals and dividing by q, we must show that

$$\int_{v_2^{**}}^{v_2^{***}} ((p_2-p_1)v-c)\,dF(v) \tag{B9}$$

is positive for k sufficiently small and $\bar{v}$ sufficiently big. We can rewrite (B9) as

$$\left(F\left(v_2^{***}\right)-F\left(v_2^{**}\right)\right)\int_{v_2^{**}}^{v_2^{***}}\left(\left(p_2-p_1\right)v-c\right)dF_2\left(v\right) \tag{B10}$$

where $F_2(v)=\frac{F(v)-F\left(v_2^{**}\right)}{F\left(v_2^{***}\right)-F\left(v_2^{**}\right)}$, so that

$$F_2\left(v_2^{**}\right)=0 \quad \text{and} \quad F_2\left(v_2^{***}\right)=1. \tag{B11}$$

To show that (B9) is positive, it suffices, from (B10), to show that

$$\int_{v_2^{**}}^{v_2^{***}}\left(\left(p_2-p_1\right)v-c\right)dF_2\left(v\right)>0. \tag{B12}$$

After integration by parts, the left-hand side of (B12) can be written as

$$\left(p_2-p_1\right)v_2^{***}-c-\left(p_2-p_1\right)\int_{v_2^{**}}^{v_2^{***}}F_2\left(v\right)dv. \tag{B13}$$

From hypothesis, we can choose k sufficiently small and $\bar{v}$ sufficiently big so that

$$F_2\left(v\right)\leq\left(\frac{v}{v_2^{***}-v_2^{**}}-\frac{v_2^{**}}{v_2^{***}-v_2^{**}}\right)+\frac{v_2^{**}}{2\left(v_2^{**}-v_2^{***}\right)} \tag{B14}$$

From (B14), (B13) exceeds

$$\begin{aligned}&\left(p_2-p_1\right)v_2^{***}-c-\left(p_2-p_1\right)\left(\frac{\left(v_2^{***}\right)^2-\left(v_2^{**}\right)^2}{2\left(v_2^{***}-v_2^{**}\right)}-\frac{v_2^{**}}{2}\right)\\&=\frac{\left(p_2-p_1\right)v_2^{***}}{2}-c\\&=\frac{c}{2s}-c,\end{aligned}$$

which is positive because $\frac{c}{2s}-c>0$. *Q.E.D*

Proposition 3. In the static model, a firm undertaking R&D is (weakly) worse off if it has a competitor.

Proof. When the firm has no competitor, its expected payoff is, depending on its cost,

$$p_1v \quad \text{or} \quad p_1v-c, \tag{B15}$$

If there is no patent protection and the firm faces a competitor, its payoff is

$$sp_1v \quad \text{or} \quad sp_1v-c \quad \text{when the other firm simply imitates}$$

or

$$sp_2v \quad \text{or} \quad sp_2v-c \quad \text{when the other firm also invests,}$$

all of which are less than their counterparts in (B15). If instead there is patent protection, then the firm's payoff is (B15) when the other firm does not invest and

$$\frac{1}{2}p_2v \quad \text{or} \quad \frac{1}{2}p_2v-c \quad \text{when the other firm invests,} \tag{B16}$$

which is each less than its counterpart in (B15). *Q.E.D.*

Proposition 4. In the static model, a firm undertaking R&D is better off if there is patent protection than if there is no such protection.

Proof. If there is patent protection, then a firm that undertakes R&D with cost c has payoff either

$$\left(qp_1+\frac{1}{2}\left(1-q\right)p_2\right)v-c \quad \text{or} \quad \frac{1}{2}p_2v-c,$$

depending on whether or not the other firm does too. If instead there is no protection, the payoffs are

$$\left(sqp_1+s\left(1-q\right)p_2\right)v-c \quad \text{or} \quad sp_2v-c.$$

But

$$(sqp_1 + s(1-q)p_2)v < sp_2v < \frac{1}{2}p_2v < \left(qp_1 + \frac{1}{2}(1-q)p_2\right)v,$$

and so the firm is better off with patent protection. *Q.E.D.*

Proposition 6. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then, for $i = 1, 2$, the likelihood that firm i behaves inefficiently in the sequential model without patents is lower than that in the static model without patents.

Proof. For $i = 1, 2$ the probability of inefficiency for firm i without patent protection in the static model is

$$F\left(v_i^{***}\right) - F\left(v_i^{*}\right), \tag{B17}$$

whereas that in the dynamic model is

$$F\left(v_i^{\circ\circ\circ}\right) - F\left(v_i^{\circ}\right). \tag{B18}$$

Now, from (2), (6), (12) and (14),

$$\left(v_1^{*}, v_1^{***}\right) = \frac{1}{1-2q(1-q)p_1-(1-q)^2p_2}\left(v_1^{\circ}, v_1^{\circ\circ\circ}\right) \tag{B19}$$

If $\bar{v}$ is sufficiently big and k sufficiently small, the Upper Tail Condition implies that $F(xv_1^{\circ\circ\circ}) - F(xv_1^{\circ})$ is increasing in x, for $x \in [1, \frac{1}{1-2q(1-q)p_1-(1-q)^2p_2}]$. Hence, we conclude from (B19) that

$$F\left(v_1^{***}\right) - F\left(v_1^{*}\right) > F\left(v_1^{\circ\circ\circ}\right) - F\left(v_1^{\circ}\right). \tag{B20}$$

Similarly, we have

$$\left(v_2^{*}, v_2^{***}\right) = \frac{1}{1-p_2+2q(p_2-p_1)}\left(v_2^{\circ}, \frac{c(1-p_2+2q(p_2-p_1))}{s(p_2-p_1)}\right),$$

And so, from the Upper Tail Condition, we obtain

$$F\left(v_2^{***}\right) - F\left(v_2^{*}\right) > F\left(\frac{c(1-p_2+2q(p_2-p_1))}{s(p_2-p_1)}\right) - F\left(v_2^{\circ}\right) > F\left(v_2^{\circ\circ\circ}\right) - F\left(v_2^{\circ}\right). \tag{B21}$$

Hence from (B20) and (B21), (B17) is bigger than (B18). *Q.E.D.*

Proposition 7. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, then there exists $\bar{q} > 0$ such that, expected net social welfare in the dynamic model is higher in equilibrium without patent protection than in equilibrium with such protection provided that $q < \bar{q}$.

Proof. For q near enough 0, we have, from (12)–(15), (18)–(20), and (23)

$$v_{.5}^{\circ\circ} < v_{.75}^{\circ\circ} < v_{1.5}^{\circ\circ} = v_1^{\circ} < v_2^{\circ} < v_1^{\circ\circ\circ} < v_2^{\circ\circ\circ} < v_2^{\circ\circ}.$$

Hence, the expected difference in welfare between having patents and not having them is

$$\begin{aligned}
&\int_{v_{.5}^{\circ\circ}}^{v_{.75}^{\circ\circ}} \left[\frac{(qp_1+(1-q)p_2)v}{1-2q(1-q)p_1-(1-q)^2p_2} - qc - \frac{\left(2q(1-q)p_1+(1-q)^2p_2\right)v}{1-2q(1-q)p_1-(1-q)^2p_2}\right] dF(v) \\
&+ \int_{v_{.75}^{\circ\circ}}^{v_{1.5}^{\circ\circ}} \left[\frac{p_2v}{1-2q(1-q)p_1-(1-q)^2p_2} - 2qc - \frac{\left(2q(1-q)p_1+(1-q)^2p_2\right)v}{1-2q(1-q)p_1-(1-q)^2p_2}\right] dF(v) \\
&+ \int_{v_{1.5}^{\circ\circ}}^{v_1^{\circ\circ\circ}} \left[\frac{p_2(v-qc)}{1-qp_1-(1-q)p_2} - 2qc - \frac{\left(2q(1-q)p_1+(1-q)^2p_2\right)v}{1-2q(1-q)p_1-(1-q)^2p_2}\right] dF(v) \\
&+ \int_{v_1^{\circ\circ\circ}}^{v_2^{\circ\circ\circ}} \left[\frac{p_2(v-qc)}{1-qp_1-(1-q)p_2} - 2qc - \frac{(qp_1+(1-q)p_2)v-qc}{1-qp_1-(1-q)p_2}\right] dF(v) \\
&+ \int_{v_2^{\circ\circ\circ}}^{v_2^{\circ\circ}} \left[\frac{p_2(v-qc)}{1-qp_1-(1-q)p_2} - 2qc - \frac{p_2v-2qc}{1-p_2}\right] dF(v).
\end{aligned} \tag{B22}$$

It will suffice to show that there exists $A > 0$ such that, for q sufficiently near 0, (B22) is less than $-A$.

As $q \to 0$, the integrands of the first four integrals of (B22) tend to zero. Furthermore, $v_{.5}^{\circ\circ}$, $v_{.75}^{\circ\circ}$, $v_{1.5}^{\circ\circ}$, $v_1^{\circ\circ\circ}$, and $v_2^{\circ\circ\circ}$ all tend to finite limits. Hence, the first four integrals all tend to zero as $q \to 0$. To understand the fifth integrand, note that the social payoff from having patents is

$$p_2v - 2qc + p_2W_{1.5}^{\circ\circ} = \frac{p_2(v-qc)}{1-qp_1-(1-q)p_2} - 2qc,$$

whereas the social payoff when there is no patent protection is

$$\frac{p_2 v - 2qc}{1 - p_2}.$$

Now, the fifth integral can be written as

$$\left(F\left(v_2^{\circ\circ}\right) - F\left(v_2^{\circ\circ\circ}\right)\right) \int_{v_2^{\circ\circ\circ}}^{v_2^{\circ\circ}} (a^\circ v + b)\, dF^*(v) + F\left(v_2^{\circ\circ\circ}\right) \int_{v_2^{\circ\circ\circ}}^{v_2^{\circ\circ}} (a^\circ v + b)\, dv,$$

where

$$F^*(v) = \frac{F(v) - F\left(v_2^{\circ\circ\circ}\right)}{F\left(v_2^{\circ\circ}\right)}$$

$$a^\circ = \frac{qp_2(p_1 - p_2)}{(1 - qp_1 - (1-q)p_2)(1 - p_2)}$$

and

$$b = \frac{q\left(p_2 - p_2^2 - 2qp_1p_2 + 2qp_2^2\right)c}{(1 - qp_1 - (1-q)p_2)(1 - p_2)}.$$

Hence, it suffices to show that there exist $D > 0$ and $E > 0$ such that

$$\int_{v_2^{\circ\circ\circ}}^{v_2^{\circ\circ}} (a^\circ v + b)\, dv < -D \tag{B23}$$

$$\int_{v_2^{\circ\circ\circ}}^{v_2^{\circ\circ}} (a^\circ v + b)\, dF^*(v) < -E \tag{B24}$$

for q near enough 0. The left-hand side of (B23) can be rewritten as

$$\frac{a^\circ}{2}\left[\left(v_2^{\circ\circ}\right)^2 - \left(v_2^{\circ\circ\circ}\right)^2\right] + b\left(v_2^{\circ\circ} - v_2^{\circ\circ\circ}\right). \tag{B25}$$

Because

$$\begin{aligned} v_2^{\circ\circ} &\approx \frac{(1 - p_2)c}{q(p_2 - p_1)}, \quad a^\circ \approx \frac{qp_2(p_1 - p_2)}{(1 - p_2)^2} \\ b &\approx \frac{q\left(p_2 - p_2^2\right)c}{(1 - p_2)^2} \quad \text{for} \quad q \approx 0, \end{aligned} \tag{B26}$$

the limit of (B25) as $q \to 0$ is $-\infty$, and so (B23) holds. The left-hand side of (B24) can be rewritten as

$$a^\circ v_2^{\circ\circ} + b - a^\circ \int_{v_2^{\circ\circ\circ}}^{v_2^{\circ\circ}} F^*(v)\, dv. \tag{B27}$$

For k small enough and $\bar{v}$ big enough, the Upper Tail Condition ensures that

$$F^*(v) \leq \left(\frac{v}{v_2^{\circ\circ} - v_2^{\circ\circ\circ}} - \frac{v_2^{\circ\circ\circ}}{v_2^{\circ\circ} - v_2^{\circ\circ\circ}}\right) + \frac{v_2^{\circ\circ\circ}}{2\left(v_2^{\circ\circ} - v_2^{\circ\circ\circ}\right)}. \tag{B28}$$

From (B28), (B27) is no greater than

$$a^\circ v_2^{\circ\circ} + b - \frac{a}{2}\left(v_2^{\circ\circ} + v_2^{\circ\circ\circ}\right) + \frac{a^\circ v_2^{\circ\circ\circ}}{2} = \frac{a^\circ}{2} v_2^{\circ\circ} + b, \tag{B29}$$

and from (B26), the limit of the right-hand side of (B29) as $q \to 0$ is

$$\frac{-p_2 c}{2(1 - p_2)},$$

and so (B24) holds. *Q.E.D.*

Proposition 8. If the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small, than each firm's *ex ante* expected profit in the dynamic model is higher in equilibrium *without* patent protection than in equilibrium *with* protection, provided that q is sufficiently small and s is near enough $\frac{1}{2}$.

Proof. Following the argument in the proof of Proposition 7, we can express the difference in firm 1's payoff between equilibrium with and without patents, assuming q is sufficiently small, as

$$\begin{aligned}
&\int_{v_{.5}^{\circ\circ}}^{v_{1.5}^{\circ\circ}} \left[\frac{\left(qp_1 + (1-q)\frac{1}{2}p_2\right) v}{1 - 2q(1-q)p_1 - (1-q)^2 p_2} - qc - \frac{s\left(2q(1-q)p_1 + (1-q)^2 p_2\right) v}{1 - 2q(1-q)p_1 - (1-q)^2 p_2} \right] dF(v) \\
&+ \int_{v_{.5}^{\circ\circ}}^{v_{1.5}^{\circ\circ}} \left[\frac{\frac{1}{2}p_2}{1 - 2q(1-q)p_1 - (1-q)^2 p_2} - qc - \frac{s\left(2q(1-q)p_1 + (1-q)^2 p_2\right) v}{1 - 2q(1-q)p_1 - (1-q)^2 p_2} \right] dF(v) \\
&+ \int_{v_{.5}^{\circ\circ}}^{v_{1}^{\circ\circ\circ}} \left[\frac{\frac{1}{2}p_2(v - qc)}{1 - qp_1 - (1-q)p_2} - qc - \frac{s\left(2q(1-q)p_1 + (1-q)^2 p_2\right) v}{1 - 2q(1-q)p_1 - (1-q)^2 p_2} \right] dF(v) \qquad \text{(B30)} \\
&+ \int_{v_{1}^{\circ\circ\circ}}^{v_{2}^{\circ\circ\circ}} \left[\frac{\frac{1}{2}p_2(v - qc)}{1 - qp_1 - (1-q)p_2} - qc - \frac{s(qp_1 + (1-q)p_2)v - qc}{1 - qp_1 - (1-q)p_2} \right] dF(v) \\
&+ \int_{v_{2}^{\circ\circ\circ}}^{v_{2}^{\circ\circ}} \left[\frac{\frac{1}{2}p_2(v - qc)}{1 - qp_1 - (1-q)p_2} - qc - \frac{sp_2 v - qc}{1 - p_2} \right] dF(v).
\end{aligned}$$

As $q \to 0$, the integrands of the first four integrals of (B30) tend to zero and $v_{.5}^{\circ\circ}$, $v_{.75}^{\circ\circ}$, $v_{1.5}^{\circ\circ}$, $v_{1}^{\circ\circ\circ}$, and $v_{2}^{\circ\circ\circ}$ tend to finite limits, as in the proof of Proposition 7. Hence, as in that proof, the first four integrals all tend to zero as $q \to 0$. Furthermore, for s near $\frac{1}{2}$, the fifth integral is about half that of its counterpart in (B22). Hence, by the same logic as in the previous proof, (B30) is negative, given the hypotheses. *Q.E.D.*

Proposition 9. Assume that the Upper Tail Condition holds for $\bar{v}$ sufficiently big and k sufficiently small and that

$$p_2 > \frac{2p_1}{1 + p_1}. \qquad (24)$$

If s is near enough $\frac{1}{2}$, then in the sequential model a firm gains from having a competitor and being imitated, whether or not there is patent protection.

Proof. If a firm has no competitor, its payoff for v big enough is

$$\frac{p_1 v - qc}{1 - p_1}. \qquad \text{(B31)}$$

By contrast in two-firm equilibria with and without patents, a firm's payoffs, for v big enough, are, respectively,

$$\frac{\frac{1}{2}(p_2 v - 2qc)}{1 - p_2} \qquad \text{(B32)}$$

and

$$\frac{sp_2 v - qc}{1 - p_2}. \qquad \text{(B33)}$$

Now, because (24) holds, $\frac{sp_2}{1-p_2} > \frac{p_1 v}{1-p_2}$ for s near $\frac{1}{2}$. Hence, for v big enough, (B32) and (B33) are bigger than (B31). We conclude that for k small enough and $\bar{v}$ big enough, a firm's equilibrium payoff with competition is bigger than that without competition. *Q.E.D.*

References

AHARONIAN, G. *Patent News Service*, 1995.

BESSEN, J. "Patents and the Diffusion of Technical Information." *Economics Letters,* Vol. 86, (2005), pp. 121–128.

—— AND HUNT, R. "An Empirical Look at Software Patents." Federal Reserve Bank of Philadelphia Working Paper no. 03-17, 2004.

—— AND MEURER, M. *Patent Failure: How Judges, Bureaucrats and Lawyers Put Innovators at Risk*. Princeton, N.J.: Princeton University Press, 2008.

BOLDRIN, M. AND LEVINE, D. "The Economics of Ideas and Intellectual Property." *Proceedings of the National Academy of Sciences,* Vol. 102 (2004), pp. 1252–1256.

CHANG, H. "Patent Scope, Antitrust Policy and Cumulative Innovation." *RAND Journal of Economics,* Vol. 26 (1995), p. 34.

DASGUPTA, P. AND MASKIN, E. "The Simple Economics of Research Portfolios." *The Economic Journal,* Vol. 97 (1987), p. 581.

—— and Stiglitz, J. "Uncertainty, Industrial Structure, and the Speed of R&D." *Bell Journal of Economics,* Vol. 11 (1980), p. 1.

Fershtman, C. and Kamien, M. "Cross Licensing of Complementary Technologies." *International Journal of Industrial Organization,* Vol. 10 (1992), p. 329.

Gallini, N. "Deterrence through Market Sharing: A Strategic Incentive for Licensing." *American Economic Review,* Vol. 74 (1984), pp. 931–941.

—— and Scotchmer, S. "Intellectual Property: When Is It the Best Incentive System?" In A. Jaffe, J. Lerner and S. Stern, eds., *Innovation Policy and the Economy,* Volume 2. Cambridge, Mass.: MIT Press, 2001.

—— and Winter, R. "Licensing in the Theory of Innovation." *RAND Journal of Economics,* Vol. 16 (1985), pp. 237–252.

Green, J. and Scotchmer, S. "On the Division of Profit in Sequential Innovation." *RAND Journal of Economics,* Vol. 26 (1995), p. 20.

Grossman, S. "A Characterization of the Optimality of Equilibrium in Incomplete Markets." *Journal of Economic Theory,* Vol. 15 (1977), pp. 1–15.

Hunt, R. "Patentability, Industry Structure, and Innovation." *Journal of Industrial Economics,* Vol. 52 (2004), pp. 401–425.

Katz, M. and Shapiro, C. "On the Licensing of Innovation." *RAND Journal of Economics,* Vol. 16 (1985), pp. 504–520.

—— and ——. "How To License Intangible Property." *Quarterly Journal of Economics,* Vol. 101 (1986), pp. 567–589.

Keely, L. 2005. "Choosing to Exchange Research Ideas." Unpublished.

Lemley, M. "The Economics of Improvement in Intellectual Property Law." *Texas Law Review,* Vol. 75 (1997), pp. 989–1084.

Levin, R., Klevorick, A., Nelson, R., and Winter, S. "Appropriating the Returns from Industrial Research and Development." *Brookings Papers on Economic Activity,* Vol. 3 (1987), p. 783.

Loury, G. "Market Structure and Innovation." *Quarterly Journal of Economic,* Vol. 93 (1979), p. 395.

Machlup, F. and Penrose, E. "The Patent Controversy in the Nineteenth Century." *The Journal of Economic History,* Vol. 10 (1950), pp. 1–29.

Mansfield, E., Schwartz, M., and Wagner, S. "Imitation Costs and Patents: An Empirical Study." *Economic Journal,* Vol. 91 (1981), p. 907.

Mowery, D. "Introduction." In D. Mowery, ed., *The International Computer Software Industry: A Comparative Study of Industry Evolution and Structure*. Oxford: Oxford University Press, 1996.

O'Donoghue, T. "A Patentability Requirement for Sequential Innovation." *RAND Journal of Economics,* Vol. 29 (1998), pp. 654–679.

——, Scotchmer, S., and Thisse, J. "Patent Breadth, Patent Life and the Pace of Technological Progress." *Journal of Economics and Management Strategy,* Vol. 7 (1998), pp. 1–32.

Oz, E. "Acceptable Protection of Software Intellectual Property: A Survey of Software Developers and Lawyers." *Information & Management,* Vol. 34 (1998), pp. 161–173.

Pakes, A., and Schankerman, M. "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources." In Z. Griliches, ed., *R&D, Patents and Productivity*. Cambridge, Mass.: NBER, 1984.

Sakakibara, M. and Branstetter, L. "Do Stronger Patents Induce More Innovation? Evidence from the 1998 Japanese Patent Law Reforms." *Rand Journal of Economics,* Vol. 32 (2001), pp. 77–100.

Sanders, B., Rossman, J., and Harris, L. "The Economic Impact of Patents." *Patent, Trademark and Copyright Journal of Research and Education,* Vol. 2 (1958), p. 355.

Scherer, F. and Harhoff, D. "Technology Policy for a World of Skew-Distributed Outcomes." *Research Policy,* Vol. 29 (2000), pp. 559–566.

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

——. "Protecting Early Innovators: Should Second-Generation Products be Patentable?" *RAND Journal of Economics,* Vol. 27 (1996), p. 322.

——. *Innovation and Incentives*. Cambridge, Mass.: MIT Press, 2005.

—— and Green, J. "Novelty and Disclosure in Patent Law." *RAND Journal of Economics,* Vol. 21 (1990), pp. 131–146.

Shavell, S. and Ypersele, T. "Rewards vs. Intellectual Property Rights." *Journal of Law and Economics,* Vol. 44 (2001), pp. 525–548.

Silverberg, G. and Verspagen, B. "The Size Distribution of Innovations Revisited: An Application of Extreme Value Statistics to Citation and Value Measures of Patent Significance." MERIT-Infonomics Research Memorandum no. 2004-021, 2004.

Spence, M. "Cost Reduction, Competition, and Industry Performance." *Econometrica,* Vol. 52 (1984), p. 101.

Sutton, J. *Technology and Market Structure*. Cambridge, Mass.: MIT Press, 1998.

Tandon, P. "Rivalry and the Excessive Allocation of Resources to Research." *Bell Journal of Economics,* Vol. 14 (1983), pp. 152–165.

Taylor, C. T. and Silberston, Z. A. *The Economic Impact of the Patent System: A Study of the British Experience*. Cambridge, UK: Cambridge University, p. 294, 1973.

TILTON, J. *International Diffusion of Technology: The Case of Semiconductors*. Washington, D.C.: Brookings Institution, 1971.

USPTO. "Hearings on Software Patent Protection." *United States Patent and Trademark Office*, January–February 1994.

WIGGINS, S. AND LIBECAP, G. "Oil Field Unitization: Contractual Failure in the Presence of Imperfect Information." *American Economic Review,* Vol. 75 (1985), pp. 368–385.

WRIGHT, B. "The Economics of Invention Incentives: Patents, Prizes, and Research Contracts." *American Economic Review,* Vol. 73 (1983), pp. 691–707.