Rethinking the Prospect Theory: A Neo-Demsetzian View

John F. Duffy*

Traditionally, the grant of intellectual property rights in innovations has been justified on the grounds that the rights provide a reward for the sizeable investments needed to create the intellectual property disclosed in the patent document. Because such rewards exist, firms will have an incentive to generate the valuable intellectual property protected by the patent. Implicitly or explicitly, such "reward" or "incentive" theories embrace *backward-looking* justifications for awarding rights: The patent serves to protect the investments made to create the invention and, of course, those investments occur *prior to* patenting.

In his 1977 article *The Nature and Function of the Patent System*,[1] Edmund Kitch theorized that the patent system serves an important, and previously unrecognized, "prospect" function. Kitch observed that the patent system often confers "prospect" patents—broad patents issued in the very early stages of technical development—and that such patents have "a scope that reaches well beyond what the reward function would require."[2] These patents were nonetheless socially beneficial because they put their owner "in a position to coordinate the search for technological and market

---

*Professor of Law, William & Mary School of Law. Please send comments to jfduff@wm.edu. The author thanks Mark Grady, Richard Hynes, Edmund Kitch, Clarisa Long, John McGinnis, Alan Meese, Robert Merges, Anne Sprightley Ryan and Stewart Sterk for comments and assistance on earlier drafts of this article.

[1] Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J. L. & ECON. 265, 266 (1977).

[2] *Id.* at 267. Subsequent commentators, even those skeptical of Kitch's theory, have also viewed the prospect theory as an attempt to justify broad patent rights. *See* ROBERT P. MERGES & RICHARD R. NELSON, MARKET STRUCTURE AND TECHNICAL ADVANCE: THE ROLE OF PATENT SCOPE DECISIONS, IN ANTITRUST, INNOVATION AND COMPETITIVENESS (Thomas M. Jorde & David J. Teece eds., 1992) (noting that Kitch's "concept of the patent system would argue for the granting an initial patent of broad scope to enable the pioneering inventor or firm to plan, undertake, or orchestrate future developments"); *id.* ("Kitch argues that granting broad claims leads to good social results. We are not sure.").

enhancement of the patent,"[3] and that coordination "increases the efficiency with which investment in innovation can be managed."[4] Kitch's justification for the patent system was thus *forward-looking*. The function of the patent system was to encourage efficient investment in, and management of, the technological prospect *after* the property right has been granted.

This basic exposition identifies the difference between the traditional, backward-oriented theory and the forward-looking prospect theory but does provide a reason to prefer the forward-looking justification. Indeed, at first blush, the traditional theory may seem to be a better basis for structuring a patent system. Under both the traditional view and Kitch's view, patent systems exist to generate technical knowledge that, without patent protection, would be easy to appropriate. Why then shouldn't society wait and grant the patent only when relevant knowledge is fully developed?

For Kitch, the reason to prefer earlier rather than later patenting had been previously discovered by Yoram Barzel, who demonstrated that innovation presents a classic common resource or "common pool" problem: Because the right to innovate is a common right (it is not under exclusive control of any one firm), competition among firms will lead to inefficient races-to-invent that can dissipate any social surplus associated with an invention.[5] Barzel had suggested that the surplus could be preserved if the government assigned or auctioned off exclusive claims to develop technological opportunities at a very early time — i.e., before any resources were expended on

---

[3] Kitch, 20 J. L. Econ. at 276.

[4] *Id.*

[5] Yoram Barzel, *Optimal Timing of Innovations*, 50 REV. ECON. & STAT. 348, 348-349 (1968). The inefficiency arises because competition will push firms to innovate when the total discounted returns from the innovation first exceed the total costs of the innovation, not when profits would be maximized. Id. at 348, and at 354.

developing the technology.[6] The goal of Barzel's auctioned claim system was to eliminate the common pool problem before the inefficient racing for patents occurred. Kitch formulated his prospect theory "in response to Barzel[]"[7] and claimed that "[w]hat Barzel did not realize is that a patent system can be such a claim system and, indeed, that it is a more sensible system than an auction system would be."[8] For Kitch, the prospect features of the patenting system (again, the granting of broad rights early in time) help to solve the common pool problem noted by Barzel "by awarding exclusive and publicly recorded ownership of a [technological] prospect shortly after its discovery."[9]

Kitch's prospect theory has become a standard part of the law and economics literature on patent law,[10] and yet it has remained "highly controversial."[11] While some commentators such as Mark Grady and Jay Alexander have found the prospect theory to be "bold" and "inspired,"[12] others have been quite critical. In a reply to Kitch's article, Donald McFetridge and Douglas Smith argued that "prospect features [in a patent system] fail to assist market participants in their attempt to

---

[6] *Id.* at 352 n.11.

[7]Id. at 265.

[8] Kitch, *supra* note 1, at 265-66.

[9] *Id.* at 266.

[10] *See*, *e.g.*., RICHARD A. POSNER, ECONOMIC ANANLYSIS OF LAW §3.3 (3d ed. 1986); Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 1045 (1997) (describing Kitch's theory as "one of the most significant efforts to integrate intellectual property with property rights theory").

[11] A. Samuel Oddi, *Un-Unified Economic Theories of Patents - The Not-Quite-Holy Grail*, 71 NOTRE DAME L. REV. 267, 269 (1996).

[12] Mark F. Grady & Jay I. Alexander, *Patent Law and Rent Dissipation*, 78 VA. L. REV. 305, 314, 315 (1992).

economize on the common property resource" and concluded that "[t]he prospect approach is not a useful framework within which to assess the merits of the patent system."[13] A later review by Roger Beck argued that Kitch's claims about the patent system were "without foundation,"[14] and Robert Merges and Richard Nelson have challenged Kitch's "view that coordinated development is better than rivalrous," contending that "[i]n principle it could be, but in practice it generally is not."[15] Perhaps the most harsh commentator is Frederic Scherer, who has derided the prospect theory as "little influenced by any concern for reality."[16]

Two problems have dogged the prospect theory. The first arises from a basic limitation on the rights of a patent holder. Even after the grant of a patent on a particular technology, the patent holder does not have the exclusive right to make patentable improvements to the technology.[17] Thus, for example, the first holder of a broad patent on all laser technology cannot prevent another inventor from obtaining a patent on an improved laser.[18] In other words, the right to continue

---

[13] Donald McFetridge and Douglas Smith, *Patents, Prospects, and Economic Surplus: A Comment*, 23 J. L. & ECON. 197, 203 (1983). *See also* Pankaj Tandon, *Rivalry and the Excess Allocation of Resources to Research*, 14 BELL. J. ECON. 152 n. 1 (1983) (noting that the prospect theory "merely shifts [the common pool problem] to a more primitive level").

[14] Roger L. Beck, *The Prospect Theory of the Patent System and Unproductive Competition*, 5 RES. L. & ECON. 193, 194 (1983).

[15] Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 872 (1990).

[16] FREDERIC M. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 447 n.30 (2d ed 1980).

[17] Interestingly, patent law differs on this point from copyright law, which does confer on the initial rightholder the right to development improvements or, in the language of copyright, "derivative works" based on the first work. See 17 U.S.C. § 106.

[18] The holder of the first or pioneer patent and the improvement patent holder will have overlapping exclusive rights; each will be able to exclude the other from using the improved laser.

innovating remains a common right *even within the claims of the granted patent*. This problem has been recognized by Roger Beck, Mark Lemley and others,[19] and it casts doubt on whether, as Kitch claimed, a prospect patent holder has sufficient legal rights to solve the common pool problem by coordinating and controlling further investment in innovation.

As I have noted in another paper,[20] this first problem may not be as large as some critics of the prospect theory have suggested, but it does still present a puzzle. While a prospect patent holder does not have sufficient legal rights to eliminate the common pool problem existing within the claims of its patent, it can maintain control over its patented technology by accelerating second-generation innovations (i.e., improvements to the prospect technology) to a time before uncoordinated rivalry would have produced those innovations. In short, the prospect patent holder can maintain control over the development of its technology by continuing to innovate quickly. Thus, even after the grant of a broad prospect patent, the legal structure of the patent system seems to encourage patent races of the very sort that, under the Barzelian model, could lead to inefficient acceleration of innovation. That structure seems puzzlingly inconsistent with the prospect theory. If the patent system grants early patents to foreclose inefficient patent racing, why then does it continue to allow racing within the claims of those early patents?

The second problem also concerns a common pool problem. It was first noted by McFetridge and Smith in a reply to Kitch's article,[21] and it is perhaps a more fundamental objection

---

[19] *See* Beck *supra* note 14, at 197-98 & 205; Lemley, *supra* note 10, at 1047 (noting that patent law does not give the inventor exclusive control over improvements to the invention).

[20] *See* John F. Duffy, *A Prospect Patent's Power to Coordinate Further Innovation* (2001).

[21] *See* Donald McFetridge and Douglas Smith, *Patents, Prospects, and Economic Surplus: A Comment*, 23 J. L. & ECON. 197, 203 (1983).

to the prospect theory. McFetridge and Smith accepted that the early grant of patent rights may lead to more efficient commercial development of a technology. But, they noted, the increase in the efficiency of the commercial exploitation will serve only to increase the value of the prospect right and, since the right to seek the prospect patent right remains a common right, the increased value of the prospect right will intensify the racing for that right. Thus, the granting of broad "prospect" patents in the early stages of technical development will not eliminate, but merely shift back in time, the rent-dissipating races noted by Barzel.

This paper reexamines the prospect theory and concludes that, while the "prospect features" of patent law (i.e., the rules permitting broad patents to be issued in the early stages of technical development) do serve a socially useful function, it is not the precise function identified by Kitch. McFetridge and Smith are correct in arguing that the prospect features of patent law cannot prevent rent-dissipating patent races and cannot provide an answer to the common pool problem identified by Barzel. However, rents associated with a patent grant can be dissipated through acceleration or duplication of efforts. The prospect features of the patent system—particularly the preference for the early grant of patent rights—trade duplication for acceleration. They determine not *whether* rents will be dissipated, but *how* they will be dissipated.

Patent rights can be awarded early or late in development. In both cases, rents implicit in the patent grant will be dissipated by rivalry to obtain the right. With a late grant of patent rights, duplication of research projects becomes more costly, and the risk of costly duplication causes firms to delay undertaking projects, or even to forego them entirely. With an early grant of patent rights, duplication occurs less frequently and, when it does occur, is less costly because it is discovered sooner. Because rents are not dissipated by duplication, firms will dissipate rents by undertaking projects earlier. But unlike duplicative efforts, which are likely to be a pure waste of social resources, acceleration of research and patenting is likely to have positive externalities if, as has always been the case, patent rights are limited in scope and term. A race to claim such partial property rights is a race to increase the social surplus associated with the innovation.

Under this view, the patent system does not attempt to limit rivalry, but to direct it toward a specific goal — reducing the social cost of providing innovation to the public. The system is analogous to Harold Demsetz's proposal for regulating so-called natural monopoly industries (industries characterized by continuously declining average cost).[22] For such industries, Demsetz proposed holding an auction of exclusive franchises in which the auction winner would be the firm providing the best combination of price and quality provided to customers. The auction would thereby channel rivalry for the exclusive right toward maximizing the consumer surplus in the industry.

The patent system fosters competitions similar to Demsetzian auctions. Because competitors can push patenting back to a time well before commercialization of the innovation (when the stream of rents begins), patent races can approximate auctions for patent rights, with the winner being the competitor willing to provide the innovation to the public for the least rents — or, equivalently, to dedicate the innovation to the public domain at the earliest time. Policies that permit patenting of embryonic research results — i.e., that allow patenting prior to the bulk of the investment needed

---

[22] *See* Harold Demsetz, *Why Regulate Utilities?*, 11 J. L. & ECON. 55, 63 (1968) ("A franchise system that awarded the franchise to that company which seemed to offer the best price-quality package would be one that allowed market competition between bidding rivals to determine that package."); see also Richard A. Posner, *The Appropriate Scope of Regulation in the Cable Television Industry*, 3 Bell J. Econ. & Man. Sci. 98, 115 (1972) (proposing that local governments auction cable television franchises, with the winner being the firm "offer[ing] the most attractive price-quality package from a subscriber standpoint").

to bring the innovation to market — increase the efficiency of the competition by ensuring that the predominant private cost of earlier patenting is the earlier expiration of the patent right (which has a private but not a social cost), not the premature expenditure of resources on innovation or the duplication of innovative efforts (both of which are private costs *and* social costs).

The patent system's relentless quest to accelerate innovation does raise the possibility noted by Barzel that investment in innovation could be undertaken inefficiently early. Yet that inefficiency is certain to occur only where patent terms are infinite and the patentee is able to capture the entire social value of the innovation. If the actual length of patent terms is taken into account, then even the model used by Barzel shows that inefficient acceleration of innovation is not likely to be a significant problem.[23] Moreover, where the patent system provides incentives sufficient to trigger premature investment in innovation, the prospect features of the patent system insure that accelerating innovation will dissipate private rents much faster than social rents.

This view of the patent system also provides new insight into the first problem noted in the prospect theory. Under the prospect theory, patent rights should be sufficient in scope to permit a patentee to coordinate further development of the asset. But the partial property rights actually conferred by a patent do not include an exclusive right to innovate within the patent's claims. The paradoxical result is that a patent holder will be able to coordinate investment only by continuing to accelerate innovation. This paradox is, however, consistent with the overall goal of the patent system, which is not to curb rivalry but merely to channel it into the acceleration of innovative efforts, which confers the innovation to the public sooner.

---

[23] *See* Part III, subsection B, *infra*, discussing the issue of patent term length. For a complete discussion, see John F. Duffy, *A Minimum Optimal Patent Term* (2003) (available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=354282).

This paper begins where Kitch began, with an examination of the American mineral claim system. For "expositional clarity" and for purposes of "policy argument," Kitch's 1977 article developed a "detailed institutional analogy between patents and mineral claims as they developed in the American West."[24] The analogy, Kitch claimed, "works in considerable detail."[25] Yet the McFetridge and Smith critique—that prospect patents cannot stop the rent-dissipating races identified by Barzel—is foreshadowed by the very history of the mineral claim system. Far from solving the common pool problem associated with unowned resources, the mineral claim system actually triggered rent-dissipating races to capture property —"gold rushes"—of the very sort that Barzel thought would also infect technological advance. Indeed, gold rushes were invited by the exceptionally narrow scope of mineral claims. This history it presents obvious difficulties for a theory that proposes broad patents as a solution to rent-dissipating races to invent.

Because of its weaknesses, the mineral claim analogy is a better introduction to a critique of the prospect theory than to a defense of it, and this introduction is provided in Part I of the paper. Part II demonstrates that a prospect patent system, while unable to prevent pre-patent races-to-invent and the associated private rent dissipation, can exchange duplication for acceleration of technological development and thereby preserve social surplus. Part III outlines some policy implications of this view of the patent system.

[24] *Id.* at 267.

[25] *Id.* at 271.

I. The Mineral Claim Analogy.

Both the patent system and the mineral claim system developed for public lands in the American West are "rule-of-capture" systems for allocating initial property rights. Such systems are, of course, ubiquitous: They govern not only patents and mining claims, but also salvage rights in maritime law, game and fishing rights, and even ownership to home-run balls hit by famous baseball players.[26] It is not surprising that the patent system bears some resemblance to other rule-of-capture systems.

Kitch identified several features of the mineral claim system that are duplicated in the patent system. For example, priority in both systems is awarded on the basis of actual discovery; a "near miss" loses in both systems, without regard to the quality of the efforts by the almost-discoverer.[27] Both systems require the claimant to identify his claim with specificity and to distinguish it from the public domain.[28] Both systems have mechanisms (albeit different ones) designed to eliminate claims

---

[26] See Dean Lueck, *The Rule of First Possession and the Design of the Law*, 38 J. L. & ECON. 393, 417-18 (1995) (detailing property systems that rely on the rule of capture or rule of first possession); Carol M. Rose, *Possession as the Origin of Property*, 52 U. Chi. L. Rev. 73 (1985) ("analogies to the capture of wild animals show up time and again when courts have to deal on a nonstatutory basis with some 'fugitive' resource that is being reduced to property for the first time, such as oil, gas, groundwater, or space on the spectrum of radio frequencies"). Phil Rogers, *McGwire: "I've Amazed Myself*," CHICAGO TRIBUNE, Sept. 28, 1998, p. 1 (noting that the 70th home-run ball hit by Mark McGwire now belonged to Phil Ozersky because it "landed in [his] hands"); Michael Grunwald, *He's Having a Ball, for $3 Million*, WASHINGTON POST, February 9, 1999, p. A1 (reporting that Ozersky sold the ball to Todd McFarlane, a comic book mogul, for $3 million).

[27] Kitch, *supra* note 1, at 273. This feature of the mining law, however, changed in the earlier twentieth century when the Supreme Court recognized the doctrine of *pedis possessio*, which confers a "foothold" to the first person who begins exploratory efforts. *See* JOHN D. LESHY, THE MINING LAW: A STUDY IN PERPETUAL MOTION 100-01 (1987); *Union Oil v. Smith*, 249 U.S. 337 (1919). In effect, the doctrine follows the rule advocated by the dissent in the famous case *Pierson v. Post*, 3 Cai. R. 175 (N.Y. Sup. Ct. 1805)—that the chase is sufficient to vest priority in property rights. No similar doctrine has ever been recognized in patent law.

[28] Kitch, *supra* note 1, at 273-74.

that have proven worthless.[29] And in both systems, claims are freely alienable.

Yet the analogy between patents and mineral claims fails on several points, two of which are important for purposes of this paper. First, and most obviously, patents are limited to a specific term of years, while mineral claims extend for as long as the claimant continues mining. Patent term limits have an overarching importance for the nature and function of the patent system, and their role is not adequately explained by the prospect theory.

A second difference is that mineral claims have always been quite narrow in scope, while patents can be quite broad, particularly the "prospect" patents that Kitch sought to explain. Kitch observed that both the mineral claim and patent systems impose limits on the scope of a claim:

> The mineral claim system restricts the area that can be claimed through rules that specify maximum boundaries in relation to the location of the mineralization. In the patent system, the applicant must limit his claims to his invention.

This similarity was not detailed any further, and the statement that the patent applicant "must limit his claims to his invention" seems remarkably brief given that, a few pages earlier, the article judged "very misleading" the "hornbook" rule that "the inventor may not claim more than he has invented."[30] This brevity is all the more striking because claim scope is crucial to the prospect theory. The first of the reasons given by Kitch for the importance of his theory is that the prospect theory helped to explain "the scope given to patent claims, a scope that reaches well beyond what the reward function would require."[31]

---

[29] *Id.* at 274. Mining law requires claimants to make certain annual expenditures in working their claims in order to retain them; patent law requires periodic maintenance fees.

[30] *Id.* at 268.

[31] *Id.*at 267. Subsequent commentators, even those skeptical of Kitch's theory, have also viewed the prospect theory as an attempt to justifybroad patent rights. *See* ROBERT P. MERGES & RICHARD R. NELSON,

Mineral claims are, in fact, exceptionally narrow in scope. Placer claims—claims to mineral deposits "not in place, that is not fixed in rock, but which are in a loose state"[32]—are generally limited under federal law to 20 acres each, with an exception that allowed an associated group to hold up to 8 such claims or 160 acres.[33] Lode claims—i.e., claims on minerals still fixed in rock—are limited to a surface area of 1500 feet by 600 feet, which is also about 20 acres of land.[34] The narrow scope of mining claims, and the balkanizing effect of these limits, are well illustrated by Figure 1, which shows the mineral claims existing early in the twentieth century from a quarter section (one quarter square mile) of land near Cripple Creek, Colorado.

MARKET STRUCTURE AND TECHNICAL ADVANCE: THE ROLE OF PATENT SCOPE DECISIONS, IN ANTITRUST, INNOVATION AND COMPETITIVENESS (Thomas M. Jorde & David J. Teece eds., 1992) (noting that Kitch's "concept of the patent system would argue for the granting an initial patent of broad scope to enable the pioneering inventor or firm to plan, undertake, or orchestrate future developments"); *id.* ("Kitch argues that granting broad claims leads to good social results. We are not sure.").

[32] United States v. Iron Silver Mining Co., 128 U.S. 673, 679 (1888). *See* 1 American Law of Mining § 32.02[3], at 32-11 (Cheryl Outerbridge ed. 1996) ("The classic placer deposit is a collection of heavy metals, such as gold or ilmenite, which, over a great period of time, has been removed from its original lode deposit, mechanically moved by the forces of erosion, and collected in a new deposit, perhaps in the gravel or sand bars of a stream bed or former stream bed.").

[33] 30 U.S.C. §§ 35, 36. These limits date back to the Mining Act of 1872. See Leshy, *supra* note 22, at 170 (1987).

[34] 30 U.S.C. § 23; see *also* A.H. RICKETTS, AMERICAN MINING LAW § 785, 442 (3d ed. 1931).

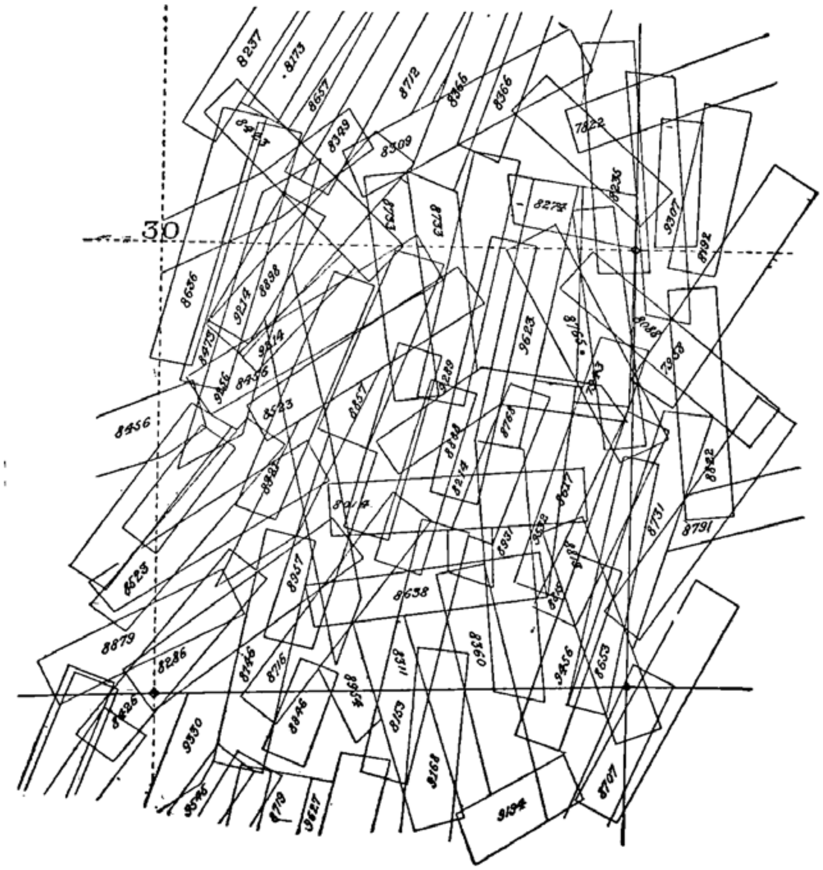

Figure 1

Mineral Claims Near Cripple Creek, Colorado ca. 1900
(Scale: Center Square is One-Quarter Square Mile)

Not only do these narrow claims bear little resemblance to the broad patent rights that Kitch sought to explain, but their history undermines any useful analogy for patent prospect rights.[35] The

[35] Professors Merges and Nelson have also noted that the mining law tries to "prevent hoarding and speculation," although they focused on the requirement that a mining claim must be actively worked to maintain the rights of the claimant. Merges & Nelson, *supra* note15, at 875. While Merges and Nelson state that a claimant is required "to work a claim actively *before* property rights will vest, *id*. at 875 (emphasis added; citing 30 U.S.C. § 28), that is not quite right. A property right to "exclusive possession and enjoyment" vests upon location of a mineral claim, 30 U.S.C. § 28, and continues so long as the claimant or locator works the claim. A *patent* for the land may not issue before a claim has been actively worked, *see* 30 U.S.C. § 29 (requiring $500 of labor expended or improvements made prior to the issuance of a patent), but a patent does not significantly expand the property rights of a locator. *See* CURTIS H. LINDLEY, LINDLEY

size limitation on mining claims "were not established with the idea that an entire deposit would be embraced with a single claim;" rather, they were designed "to prevent the discoverer of a mineral deposit from gaining exclusive rights to exploit the entire deposit."[36] As Supreme Court Justice Stephen Field, himself a former "49er,"explained in 1881: "The extent of which each [prospector] might locate, that is, appropriate to himself, was limited so that all might, in the homely and expressive language of the day, have an equal chance in the struggle for the wealth there buried in the earth."[37] The size limitations were so restrictive that the claims were usually not broad enough for effective economic exploitation. Again, in Justice Field's words, "[e]very one, at all familiar

---

ON MINES: A TREATISE ON THE AMERICAN LAW RELATING TO MINES AND MINERAL LANDS § 539, 1202 (3d. ed. 1914) ("[a]lthough a locator may obtain a patent, this patent adds but little to his security"); Ricketts, *supra* note 29, § 914, at 487 ("[a] patent is not essential to the use and enjoyment of a mining claim as it confers no greater mining rights than those obtained by valid location, and adds but little to the security of a party in continuous possession"). As noted above, the active work requirement of mining law is analogous to the maintenance fees required under patent law; it is not the primary means that the mining law traditionally used to limit the scope of claims.

[36] Leshy*, supra* note 22, at 170.

[37] *Smelting Co. v. Kemp*, 104 U.S. 636, 650 (1881). *See also* PAUL KENS, JUSTICE STEPHEN FIELD: SHAPING LIBERTY FROM THE GOLD RUSH TO THE GILDED AGE (1997) (detailing Field's relocation to California in 1849 and his experience in mining districts). Public choice considerations may also help to explain the narrow scope of mining claims. Prospectors in the middle part of the nineteenth century were in a precarious legal position. They were trespassing upon, and extracting the wealth from, publicly owned lands, and they had no authorization from the government. Their rights, if that is a proper expression for them, were based on

> [t]he early announcement of the doctrine by the courts in the mining states that controversies between the occupants of the public mineral lands were to be determined by the law of possession, and that persons claiming and in possession of mining claims on these lands were, as between themselves and all other persons, except the United States, owners of the same.

Lindley, *supra* note 30, § 536, at 1196. Of course, the courts *in the mining states* had every economic reason to "announce[]" this doctrine, since it essentially allowed residents of the mining states to expropriate wealth from the owner of the lands, the United States as a whole. Within the mining states, political support for miners' rights would be increased to the extent that the residents believed that those right benefitted not just a few, but the many—offering all an "equal chance in the struggle for the wealth buried in" the public lands.

with our mineral regions, knows that the great majority of claims . . . can be worked advantageously only by a combination among the miners, or a consolidation of their claims."[38] The narrow scope of mining claims "guarded the mines from being monopolized" and were a product of a "common aversion of the frontier democracy to monopoly."[39]

The size limitations on mining claims would not be an important check on prospecting, however, if a single prospector could (1) file multiple claims, each of which was (2) based on a discovery of only slight evidence of mineralization. For purposes of this paper, I will assume that individuals can file multiple claims, provided that the requisite level of mineralization has been discovered within each claim. In passing, however, I must note that, while this assumption seems correct under modern law,[40] it is almost certainly wrong as a matter of history: Decisional law in the nineteenth and early twentieth centuries interpreted the claim size limits in federal mining law to restrict the number of claims that a single person (including a single corporation) could file based on a discovery in a single area.[41]

---

[38] *Smelting Co.*, 104 U.S. at 654.

[39] Leshy, *supra* note 22, at 170.

[40] "There is no limit on the number of claims that a person may locate, and therefore a person is not disqualified from locating a mining claim by reason of the extent of his holdings." Lindley, *supra* note 30, § 31.01, at 31-2 – 31-3. *Cf.* Merges & Nelson, *supra* note 15, at 875 (assuming that "there are no statutory limits on the number of [mineral] claims an individual can make").

[41] Under the case law in nineteenth and early twentieth centuries, a single entity could aggregate multiple claims only by purchasing the claims from other claimants or "locators." As the Supreme Court of California explained in 1890:

> The policy and object of [the size limitations in federal law] are to limit the quantity of placer mineral land which may be located by one person to 20 acres; and, although one person may obtain a patent for more than 20 acres, he can do so only by representing to the government that he is a purchaser of the excess from one or more *bona fide* locators, whose locations were made in conformity with the above statutory limitation as to quantity.

Even if the law does not restrict the number of claims that a single person could file, it has always required the prospector to discover a valuable mineral deposit "within the limits of the claim."[42] Thus, a prospector who discovers evidence of a potentially large deposit can protect the whole find by filing multiple claims only if he can prove the required level of mineralization within each claim. Kitch asserted that "[t]he claimant for the mineral claim need not show that the mineralization is of commercial significance," but rather that "[t]he mineralization showing required was of surface mineralization which could be found without extensive excavation."[43] Patent law was similar on this score, Kitch claimed, because "the patent applicant need not show that his invention has commercial significance."[44]

Kitch was accurate in his description of federal patent law, which truly does make no inquiry

---

Mitchell v. Cline, 84 Cal. 409, 24 P. 164, 165 (1890). Thus the courts would void schemes where an individual would, "by the use of the names of his friends, relatives, or employés as dummies, locate for his own benefit a greater area of mining ground than allowed by law." Cook v. Klonos, 164 F. 529, 538 (9th Cir. 1908). *See also id.* (noting that the size restrictions of federal law were "intended to prevent the primary location and accumulation of large tracts of land by a few persons, and to encourage the exploration of the mineral resources of the public land by actual bona fide locators" and that these limitations could not be avoided by using "dummy locators"). *See also* Nome & Sinook Co. v. Synder, 187 F. 385, 389 (9th Cir. 1911) (voiding placer claim located by a partnership where one partner's interest would exceed 20 acres); Gird v. California Oil Co., 60 F. 531, 545 (C.C. S.D. Cal. 1894) (limiting a corporation's interest to 20 acres of land where three individuals were working in the employment of a single corporation at the time of locating a placer claim.). Later cases limited this early case law by stressing that the size limits applied only to a single "*discovery*." Houck v. Jose, 72 F. Supp. 6, 9 (S. D. Cal. 1947) (emphasis in original), *aff'd* 171 F.2w 211 (9th Cir. 1948). Thus, under this later case law, a single entity could make multiple claims if the requisite level of discovery was achieved within each claim. This modern rule, however, cannot be reconciled with the early case law addressing the use of "dummy" locators by individuals to avoid the federal size limitations *See* Leshy, *supra* note 22, at 172-73 (noting that, when the decision law changed to permit multiple claims, "[t]he net result was rather silly: the use of dummy locators remained fraudulent, but such fraud was unnecessary in the first place because the Law was held to place no limit on the number of claims one could locate"); Lindley, *supra* note 30, § 450 at 1062-63 (noting that individuals frequently resort to the use of "dummies" to obtain more ground than permitted and that "this is a fraud upon the government").

[42] 30 U.S.C. § 23; *see also id.* § 35 (applying same rules for placer claims).

[43] Kitch, *supra* note 1, at 271-73.

[44] *Id.* at 273.

into the commercial value of the invention sought to be patented. All that is required is that the invention be "useful"—a requirement that is satisfied if the invention achieves any beneficial result, without regard to whether the prior art could achieve the same result in a less expensive way.[45] Indeed, outside of a very narrow range of cases, the requirement of utility is usually presumed by the Patent and Trademark Office without any bother of proof.[46] Simply put, the patent law has no aversion to awarding a commercially worthless patent. But not so in mining law.

While a prospector does not have to prove with *certainty* that a mineral discovery will lead to commercial success, he is required to prove a mineral discovery with—in the words of the treatise cited by Kitch—"a reasonable prospect of success," which was defined as evidence sufficient that "a person of ordinary prudence would be justified in further expenditure of his labor and means "[47]

---

[45] 1 DONALD S. CHISUM, CISUM ON PATENTS: A TREATISE ON THE LAW OF PATENTABILITY, VALIDITY AND INFRINGEMENT § 4.01, at 4-2 (2000) ("an invention need not be superior to existing products or processes"). Justice Story provides the classic statement of the patent utility test:

> All that the law requires is, that the invention should not be frivolous or injurious to the well-being, good policy, or sound morals of society. . . . But if the invention steers wide of these objections, whether it be more or less useful is a circumstance very material to the interests of the patentee, but of no important to the public.

Lowell v. Lewis, 3 F. Cas. 1018, 1019 (C.C. Mass. 1817).

[46] *See* Chisum *supra* note 40, § 4.04[1], at 4-25 (noting that "[t]he burden of proving operability and utility shifts to the applicant only if there is a reasonable doubt as to the truth of the applicant's assertions"). In those narrow classes of cases where utility is an issue, commercial value is also not considered. Thus, for example, even where a chemical composition may have a commercial value because it is valuable object of study and further research, a patent on the composition may not be granted until some present utility for the composition is shown. Brenner v. Manson, 383 U.S. 519 (1966). Yet an economically worthless utility such as curing laboratory mice of a man-made disease does satisfy the requirement. *In re* Brana, 51 F.3d 1560 (Fed. Cir. 1995).

[47] Lindley, *supra* note 30, § 336, at 772; *see also id.* at 773 (noting that "[t]here is a material difference between a discoverer being *willing* to spend his time and money in exploiting the ground and being *justified* in so doing," with the latter "a question for expert testimony and determination by a jury"). *See also* United States v. Coleman, 390 U.S. 599 (1968) (reaffirming the "prudent-man" test for determining the sufficiency of a mineral discovery and holding that "profitability is an important consideration in applying" the test.).

Moreover, even the requirement of making a mineral discovery is itself a test pegged to commercial value because, within mining law, the term "mineral" is used in "its commercial sense" to refer only to those minerals "having sufficient value separate from its situs as part of the earth to be mined, quarried, or dug for its own sake."[48] Thus, the requirement that prospectors discover a mineral deposit, not mere indications of valuable deposits or the presence of noncommercial minerals such as quartz, prevents prospectors from accumulating rights to large tracts of potentially valuable land.[49]

The size limitations on mining claims, reinforced by the discovery requirements, foreclose any broad "prospect" mining rights of the sort envisioned by Kitch for patent law. This weakness in Kitch's analogy has even been acknowledged by champions of the prospect theory such as Grady and Alexander who, while praising the analogy between the patent and mineral claim systems as "inspired," note that "[t]he patent laws resemble the mineral claim system that might have evolved if giant finds were more common, or at least easier to anticipate."[50] But in fact, the mineral claim system in the American West evolved *even though* "giant finds"—such as the discovery of gold at Sutter's Mill in 1848—did occur, and *even though* similar finds were easy to anticipate. Indeed, it was the very anticipation of such finds that triggered gold rushes such as that experienced by California in 1849. The narrow claim structure of federal mining law evolved to serve other goals;

---

[48] Ricketts, *supra* note 29, at 24 (distinguishing between "mineral" in "its broadest and scientific meaning" and in "its commercial sense"); Leshy, *supra* note 22, at 132-32 (noting that the Interior Department interpreted the term "mineral" in the federal mining laws to refer to mineral deposits "found in quantity and quality to render the land sought to be patented more valuable on this account than for purposes of agriculture").

[49] See United States v. Coleman, 390 U.S. 599 (1968) (holding that a discovery of quartzite is not sufficient to support a claim). Lindley, *supra* note 30, § 336 at 771-72 (noting that "even indication[s]" of valuable mineralization, or the presence of quartz, do not satisfy the discovery requirement); *id.* § 437, at 1025 (noting that oil seepages on the surface do not constitute a valid discover warranting a claim because there is no "direct connection" between a commercial deposit and the surface indication).

[50] Grady & Alexander, *supra* note 12, at 315.

it was, in Justice Field's words, "so that all might, in the homely and expressive language of the day, have an equal chance in the struggle for the wealth there buried in the earth."[51]

By attempting to give all an "equal chance" in the struggle for the wealth buried in the earth, the mining law triggered precisely the type of rent-dissipating races to capture property rights — i.e., gold rushes — that Barzel claimed would also affect innovation. To the extent that the prospect theory attempts to respond to Barzel, it gains nothing from the mineral claim analogy.

II. Channeling Rivalrous Behavior: Partial Property Rights and a Quasi-Demsetzian Approach to the Innovation Industry.

In both the patent and mineral claim systems, the initial right to claim property is a common right and is therefore subject to competition among parties seeking the property. Such competition can dissipate all rents associated with the property right, and the American mineral claim system did, in fact, precipitate significant rent dissipation. The problem for the prospect theory here is not that the mineral claim analogy fails, but that it may work: Just as the mineral claim system triggered rent-dissipating "gold rushes," so too with prospect patents.

Kitch and other defenders of the prospect theory have suggested that the prospect features of patent law may avoid the rent-dissipating races that infected the mineral claim system. There is, however, no evidence to support this view, and much reason to doubt it. Subsection A reviews the basic rent-dissipation critique of the prospect theory, the attempted answers by Kitch and others to this critique, and the evidence suggesting those rejoinders are inadequate.

---

[51] Smelting Co. v. Kemp, 104 U.S. 636, 650 (1881). *See also* Durant v. Corbin, 94 Fed. 382, 383 (1899) ("The policy of the government in disposing of the mineral lands as well as other portions of the public domain is to make a general distribution among as large a number as possible of those who wish to acquire such land for their own use rather to favor a few individuals who might wish to acquire princely fortunes by securing large tracts of land.")

Nevertheless, as shown in subsection B, the prospect features of the patent system can still be useful, not because they stop rivalry, but because they channel it. The patent system can be usefully compared to Demsetz's suggestion that competition to gain an exclusive right to a natural monopoly can be harnessed to reduce deadweight loss.[52] A competition to patent earlier will resemble a Demsetzian auction because, by trying to be first to patent, competing inventors are also vying to diminish their rents by placing the patent in the public domain sooner.[53] The prospect features of patent law — i.e., the law's preference for early patenting — increase the efficiency of the competition by channeling rivalry away from duplication of efforts, which is likely to dissipate both private and social rents, and toward greater acceleration of efforts, which is more likely to dissipate private but not social rents.

Finally, subsection C demonstrates how this conception of the patent system accounts for the legal limitations on a patentee's power to control subsequent innovation within the claims of the patent.

[52] If price discrimination is not possible and elasticity of demand is positive, deadweight loss will not eliminated because the bidders will offer service at their average, not marginal costs. Requiring the consumers of intellectual property to bear average costs may, however, be optimal given real world constraints. *See generally* Ronald Coase, *The Marginal Cost Controversy*, 13 (n.s.) ECONOMICA 169 (1946).

[53] I assume here that the patent is awarded to the inventor seeking the earliest expiration date. See Part IV.A for further discussion.

A. The McFetridge & Smith Critique and the Ubiquity of Rivalry.

In response to Kitch's article, Donald McFetridge and Douglas Smith showed that a prospect patent system—one that awards a patent "relatively early in the process of technological development"[54]—might allow the patent holder to delay commercialization of an invention to a time later than if the exclusive rights were not awarded until commercialization. Eliminating an inefficient race to commercialize would generate a surplus and increase the value of the patent. Yet because the right to patent remains a common right, any increase in value of the patent would lead rival inventors to accelerate their efforts to win the patent until "[t]he gains from the lag in commercialization disappear in the rush to patent."[55] Thus, McFetridge and Smith saw the long lags between patenting and commercialization observed by Kitch as reflecting "the ubiquity of rivalry rather than the suppression of it."

McFetridge and Smith did not expressly mention the "gold rushes" triggered by the mineral claim system, but their critique is entirely consistent with that historical phenomenon. Strengthening the property rights available for capture merely increases the rivalry to obtain the rights—the greater the riches available, the more 49'ers who come West.

Kitch responded to McFetridge and Smith by arguing that transaction costs might be lower at the early stages of innovation "because the number of firms with the necessary comparative advantage to exploit the inventive possibility will be small and the uncertainties attached to each possibility make it easier to agree upon a division of activities."[56] If transaction costs were minimal, firms could reach agreements to divide the surplus rather than dissipate it through competition. In other words, the firms cartelize the early stages of innovation.

That response to McFetridge and Smith is not supported by any empirical evidence of formal

---

[54] McFetridge & Smith, *supra* note 13, at 197.

[55] *Id.* at 201.

[56] Edmund W. Kitch, *Patents, Prospects, and Economic Surplus: A Reply*, 23 J. L. & Econ. 205, 206 (1983).

or informal agreements between firms. But there is a more fundamental problem: If agreements to cartelize innovation did preserve rents, the parties to the agreements — the small number of firms with "necessary comparative advantage to exploit the inventive possibility" — would earn higher than market rates of return on capital. Agreements among those firms could not prevent entry. Other firms could invest to obtain the necessary technical experience to exploit future inventive possibilities. Capital would flow into that sector of the economy until above market rates of return were eliminated—i.e., until all rents associated with patent grants were dissipated. In an economy with well-function capital markets, Kitch's mechanism for curbing rivalry is not plausible.

A slightly different response to McFetridge and Smith has been offered by Mark Grady and Jay Alexander, who make their point by returning to the mineral claim analogy. Acknowledging that mineral claims were exceptionally narrow (see Part I.A, supra), Grady and Alexander argue that the inaccuracy of the mineral analogy actually helps the prospect theory: By giving "giant finders giant rights," the patent system may be able to stop rent dissipating races-to-invent, provided that patenting is pushed back to a point so early that rivalry is not a problem—i.e., to a point where, in their words, "a few Forty-Niners are inherently swifter than most."[57] Thus, Grady and Alexander, like Kitch, view early patenting as a solution to rivalry, although they rely on the heterogeneity of competitors rather than cartelization agreements as the solution to rivalry.

Even Grady and Alexander, however, express doubts about whether early patenting really answers the McFetridge and Smith critique. As they note, any reduction in rent dissipation after a

[57] Grady & Alexander, *supra* note 12, at 316, 317. The Grady & Alexander view is also endorsed by Dean Lueck, who proposes a general theory under which "heterogeneity [of competitors] reduces and possibly eliminates the dissipation of wealth" associated with races to claim to property rights. Lueck, *supra* note 21, at 399; *see also id.* at 418-19 (accepting the Grady & Alexander view).

patent grant might be canceled out by "the losses from accelerated pioneering invention."[58] Those doubts are warranted. The evidence strongly suggests that uncoordinated rivalry persists even at early stages of research and development.

Actual cases demonstrate the ubiquity of rivalry to invent. Consider for example the invention of the telephone, which Grady and Alexander cite as a case where the patent system gave a "giant finder" a "giant right."[59] The history of rivalry in seeking to invent the telephone is plain. Both the concept and (roughly) the principle of the telephone was set forth at least as early as 1854, twenty-one years prior to Bell's invention, by the Frenchman Charles Bourseul, who suggested that the voice vibrations could be imparted to a "moveable disk," then converted into electrical impulses, transported by wire, and reproduced at the other end by another disk.[60] Bourseul well understood the implications of this possibility:

> It need not be said that the numerous applications of the highest importance will immediately arise from the transmission of speech by electricity. Any one who is not deaf and dumb may use this mode of transmission, which would require no apparatus except an electric battery, two vibrating disks, and a wire. In many cases, as, for example, in large establishments, orders might be transmitted in this way . . . . However this may be, it is certain that in a more or less distant future speech will be transmitted by electricity.[61]

As the Supreme Court would note: "It had long been believed that if the vibrations of air caused by the voice in speaking could be reproduced at a distance by means of electricity, the speech itself

---

[58] Grady & Alexander, *supra* note 12, at 317.

[59] Nor was Bell's right as "giant" as might be thought. Within five years, Thomas Edison invented a telephone system that did not infringe the Bell patent. Edison's device was invented to force Bell's English interests to consolidate with Edison's. When that goal was accomplish, the best parts of the Edison and Bell systems were combined. Matthew Josephson, *Edison* 149-155 (1959).

[60] *See* The Telephone Cases, 126 U.S. 1, 31-32 (1888) (statement of the case) (setting forth translation of article).

[61] *Id.* at 32-33 (quoting translated article).

would be reproduced and will be understood. How to do it was the question."[62]

At least five inventors—quite possibly many more[63]—were actively seeking to answer that question prior to Bell's success. Those inventors included the German Phillip Reis, who first produced a device capable of conveying sounds with electricity and coined the word "telephone" ("Telefon," in German) fourteen years before Bell's invention;[64] Thomas Edison, who began investigating telephone technology in 1875 and filed a patent caveat one month before Bell;[65] Elisha Grey, who filed a caveat for a telephone patent on the same day that Bell filed his patent application;[66] and Daniel Drawbaugh, who lost his claim to priority over Bell by a single vote in the Supreme Court.[67] In fact, the legal proceedings to determine priority of invention were extensive. "Numerous interferences [were] declared and considered at the Patent Office,"[68] and in the courts,

---

[62] *Id.* at 532.

[63] The papers filed in The Telephone Cases listed inventions by more than fifteen individuals that were claimed to have anticipated at least a part of Graham's work. *See id.*, at 24-5 n.1 (Court's statement of the case).

[64] *Id.* at 66 (Court's statement of the case) (quoting a translation of a 1863 German article, which notes that in 1861 Reis successfully "constructed an apparatus to which he gives the name *Telephone*, and which enables one to reproduce tones, with the aid of electricity, at any given distance"). Reis's telephone accurately reproduced tones, but not speech.

[65] Josephson, *supra* note 54, at 140; The Telephone Cases, 126 U.S. at 551 (noting that the Patent Office rejected a claim for priority by Edison).

[66] The Telephone Cases, at 77-86 (setting forth Grey's patent filings); *Id.* at 567-71 (rejecting charge that Bell copied some portion of Grey's filings). (A patent "caveat" was a filing allowed under nineteenth century patent practice to preserve priority while the inventor completed a more formal patent application).

[67] The Telephone Cases, at 546-67 (discussing and rejecting Drawbaugh's claim to priority); *cf. id.* at 573-77 (dissenting opinion that accepts Drawbough's claim to priority). While the four members of the Court majority ultimately rejected Drawbough's priority, they "did not doubt" that Drawbough had conceived of the idea and "was experimenting on the subject." *Id.* at 567.

[68] *Id.* at 551 (noting that "Grey, Edison, Dolbear, and others had either claimed for themselves or others had claimed for them, priority of invention and discovery, and Bell had thus far been sustained as

over three hundred witnesses were called just on Drawbough's priority claim.[69] Oral argument at the Supreme Court stretched over twelve days, and the report of the case fills an entire volume of the United States Reports.

The case of the telephone is not unique. Evidence of pre-patent rivalry abounds even for broad pioneering patents. Thus, in the case of the telegraph, the Supreme Court noted that "many eminent and scientific men in Europe, as well as in this country," were racing to invent the device, and that ultimately "four different magnetic telegraphs . . . [were] invented and made public so nearly at the same time that each has claimed a priority; and that . . . neither inventor can be justly accused of having derived any aid from the discoveries of the other."[70] For the incandescent electric light, at least a half-dozen men were seeking the invention in the late 1870's.[71] Similarly, in the case of the laser, at least three sets of competing inventors had plausible claims to the invention.[72]

One possible answer to this history of rivalry could be that the patenting should have been pushed back even earlier in these cases. But even among the very best scientists working at the very edge of human knowledge, rivalry poses real threats.[73] And while some inventions are created by

---

against them all.").

[69] *Id.* at 555.

[70] O'Reilly v. Morse, 56 U.S. 62, 107-08 (1854). The four inventors were Morse, the German Steinheil, and the British Wheatstone and Davy.

[71] Besides Edison, the other inventors seeking stable incandescent lamps were William Sawyer, Albon Man, J.W. Starr, Hiram Maxim, Joseph Swam, St. George Lane-Fox.. *See* Josephson, *supra* note 54, at 179, 183, 191 (1959).

[72] See Gould v. Schawlow, 363 F.2d 908 (CCPA 1966); Hecht & Teresi, *Laser: Supertool of the 1980s* (1982).

[73] One famous example is the nearly simultaneous formulation of evolution theories by Charles Darwin and Alfred Wallace.

a single person without others in close pursuit, those cases are more plausibly explained as situations where competition drove the inventor to gamble on highly risky and unpromising technologies. This is, for example, certainly true in the case of xerography.

In seeking a better way to copy images, Chester Carlson, the eventual inventor of the xerographic process, decided to investigate electrostatic methods of copying because he knew that "a lot of big companies were deeply involved in research using chemical or photographic processes, and [in the inventor's words] — 'Who was I to compete against Eastman Kodak.'"[74] Though today xerography is seen as a great invention of the twentieth century, that was not true even in 1959, nearly two decades after Carlson had received his first patent. Then, as Xerox was introducing its first plain-paper copier, the conventional wisdom was that the new machine would "find plenty of competition" in the "crowded field" of office copying, and that Xerox's business strategy was a "calculated risk" and a "gamble."[75] That some such gambles pay off handsomely does nothing to demonstrate that rents are preserved, for the many less famous failures must be considered.[76] Any claim that competitive rivalry poses a diminished threat at some stage of technological development is speculation, supported by neither intuition nor empirical proof.

In an economy with well functioning capital markets and free entry into the search for technological advance, the McFetridge and Smith critique is persuasive. Rents associated with

---

[74] *The Invention Nobody Wanted*, 66 FORTUNE 155 (1962).

[75] *Out to Crack Copying Market*, 1568 BUSINESS WEEK 86, 86-89 (Sept. 19, 1959). *See also* Richard Hammer, *There Isn't Any Profit Squeeze at Xerox*, 66 FORTUNE 151 (1962) (describing the "hair-raising risk originally taken by the [Xerox] company" in pursuing Carlson's work).

[76] Technical failures include, for example, IBM's infamous and expensive attempt to develop a practical "Josephson Junction" computer in the 1980s, AT&T's research into an optical computer in the late 1980s, and unsuccessful deployment of the "Iridium" satellite system in the 1990's.

patent grants will be dissipated by rivalry. If the prospect theory has value, it is not in solving Barzel's common pool problem.

B. Channeling Rivalry: Exchanging Duplication for Acceleration in Races-to-Invent.

Even if all private rents associated with a patent grant will, on average, be dissipated by competition to obtain the patent, rents can be dissipated in at least two very different ways: (1) accelerating innovative efforts to an earlier time, and (2) duplicating the efforts of others. While both duplication and acceleration dissipate rents, society should not be indifferent in choosing between the two. Duplication is likely to be a pure waste of resources,[77] but accelerating innovative efforts can have positive externalities due to limits on the patent term and scope. Under this view, the prospect features of the patent system are useful not so much because they eliminate competitive rent dissipation, but because the can *channel* rent-seeking behavior into activities having positive externalities.[78] All private rents will be dissipated, but some social surplus preserved.

---

[77] *See* Partha Dasgupta & Joseph Stiglitz, *Industrial Structure and the Nature of Innovative Activity*, 90 ECON. J. 266, 267 (1980) (assuming that duplicative efforts are socially wasteful). Tandon demonstrates that, assuming that firms cannot communicate with each other, ex post duplicative research may not be inefficient because running simultaneously several lines of research, even those that turn out to be duplicative, will increase the probability of research success in any given time period. *See* Tandon, *supra* note 13, at 153, 155, 157. This is true, but it is a second-best solution: Duplicative efforts are wasteful but not necessarily inefficient if the cost of the waste is less then the cost to eliminate the waste through, for example, better communication between the firms.

[78] This view of the prospect theory differs from current proponents of the theory, who argue that prospect patents help to curb rent dissipation and preserve rents. *See, e.g.*, Kenneth W. Dam, *The Economic Underpinnings of Patent Law*, 23 J. LEG. STUD. 247, 264 (1994) (contending that the prospect features of patent law "tend to reduce rent seeking by inducing early elimination or redirection of R&D by rival firms on issuance"); Grady & Alexander, *supra* note 12, at 317-19 (arguing that the patent system is designed help to preserve rents by granting rights at a time when a few prospectors "are inherently swifter than the rest)); Lueck, *supra* note 21, at 417-18 (1995) (arguing that rules granting patent rights are designed to "address the potential for wasteful races by granting ownership 'early,' when claimant heterogeneity is still large," and that therefore the patent system "is not a race but rather a process by which a low- cost developer of ideas gains ownership under a priority rule"). In contrast, the view expressed here assumes that competing firms

McFetridge & Smith did not consider the effects of duplicated efforts because they relied on Barzel's model,[79] which assumes that "[o]nly one innovator is associated with a given innovation."[80] Even in the section of his paper devoted to relaxing his initial assumptions, Barzel did not consider the wasteful effects of duplication but instead reiterated that, "[a]s considered here, the basic wasteful effect of competition lies not in duplicating the use of resources but in using these resources prematurely, when they would have earned a higher return elsewhere in the economy."[81]

The efficiency gains associated with early patenting policies can be observed, however, only with models of patent races that include acceleration, duplication and finite patent lengths. Two such models are presented below. Both models will assume that firms are competing to complete an invention that will yield a present discounted value of expected royalties $R(t_I)$, where t_I is the time of invention. Royalties are assumed to increase with time due to exogenous economic expansion. Completing the invention requires a pre-patent investment, I, and a post-patent expenditure, C, to commercialize the innovation. For expositional purposes, the model will make the further simplifying assumptions that I and C are constants in time; that expenditure I occurs at time t_I; and that expenditure C occurs at a time infinitesimally later that t_I (so that I and C do not need to be discounted to present value at t_I). A more mathematically rigorous treatment of the problem is presented in the Appendix, but the fundamental results are not changed.

With no duplication of efforts, a firm making the investments in innovation at t_I could expect

are homogenous, and that all rents will be dissipated. Still, the prospect patent system is desirable, because it channels rent seeking behavior into early, rather than duplicative, technological development.

[79] McFetridge & Smith, *supra* note 13, at 199 (starting analysis with the "Barzel model").

[80] Barzel *supra* note 4, at 350.

[81] *Id.* at 352.

profits (Π) of:

(1) $$\Pi(t_I) = R(t_I) - C - I$$

Rivalry will insure that all profits are dissipated (i.e., $\Pi(t_I) = 0$), and so investment in innovation will occur at the time when expected royalties just equal the costs of innovation and commercialization ($R(t_I) = I + C$). Where duplicative investments are possible, however, each firm must take into account that possibility. The firms' behavior will be analyzed with a simple model, which does not permit firms to take probabilistic research strategies, and a probabilistic model.

Simple Model With Pure Strategies: In this model, firms cannot communicate with each other to coordinate their activities, and firms are not permitted to adopt mixed (probabilistic) strategies. Each firm assumes that its own decision to invest in innovation will be duplicated by a total of all N identical firms in the industry (including that firm).[82] The patent on the innovation is randomly assigned or randomly divided among the N firms making the investment.[83] Under these assumptions, each firm would expect that the costs of innovation would be I and the benefits would be the value of the patent ($R(t_I) - C$) times the probability of getting the patent ($1/N$). In other words, the model is like a raffle where the cost of a ticket is I, the prize is ($R(t_I) - C$), and the number of players, each with one ticket, is N. In such a game, each participant can expect profits to be:

(2) $$\Pi_I = \frac{R(t_I) - C}{N} - I$$

---

[82] *See* Tandon, *supra* note 13, at 153 (similar model also assuming no communication between firms).

[83] Patent rights are, of course, assigned or divided not on a random basis, but on the basis of rules governing priority and claim scope in interference proceedings. If, however, firms do not have good information ex ante (i.e., at the time of the decision to invest) concerning whether they will prevail in a patent interference, they would be rational to model the allocation mechanism as a random process.

If free entry and rivalry drive profits to zero, then each firm will invest in innovation when its expected reward from innovation equals the cost of innovation—i.e., when

(3) $$I = \frac{R(t_I) - C}{N}$$

(4) $$R(t_I) = C + NI.$$

Equation (4) can be understood as stating simply that investment will not occur until the expected royalties from the invention ($R(t_I)$) equal the total amount that would be expended by the industry to produce the invention ($C + NI$).

We may assume that government policy cannot affect the discounted value of the royalties ($R(t_I)$) nor the total costs of creating it ($I + C$). Patent policy can, however, affect the allocation of costs between I and C: Granting a patent earlier in the course of research and development reallocates costs from I to C, because the patent is granted when significant commercial development is still needed. If such an early patenting policy is pursued, lower royalties are needed to justify investments in innovation. As can be seen from equation (4), every dollar removed from I and added to C reduces the right hand-side of the equation by N - 1 dollars. Thus, shifting dollars from I to C decreases the royalties needed to justify the investment in innovation. Because the present value of royalties increases with time, lower royalties correspond to innovation at an earlier time.

Consider a prospective technology that costs a total of \$90 to invent and commercialize and that is expected to generate royalties of \$100 at time t_I. If there were no problem of duplication, some firm would make the \$90 investment and reap the \$10 surplus. But with duplication, the result depends crucially on the allocation of the \$90 between pre-patent and post-patent expenditures. If a mere two firms can be expected to play (N=2), and it takes $40 dollars to meet the patentability standard (I=40), then plainly the firms would lose if they invest:

(5)
$$\Pi(t_I) = (R(t_I) - C) * (1/N) - I = (100-50)(1/2) - 40$$
$$\Pi(t_I) = 50/2 - 40 = -15$$

Firms would not invest until the present value of the royalties from the technology equaled $130, which will occur later in time because, as noted above, the present value of the technology increases with time. Thus, a threat of duplication delays investment.

If, however, the patent were granted after a smaller investment were made—after, for example, only $10 is invested in the technology—then firms would invest at time t_I, even though the cost of commercializing the invention will be higher ($C = 80$):

(6)
$$\Pi(t_I) = (R(t_I) - C) * (1/N) - I = (100-80)(1/2) - 10$$
$$\Pi(t_I) = 20/2 - 10 = 0$$

Note also that the value of the patent ($R(t_I) - C$) is higher in the first case ($50) than in the second ($20), because the winner in the second case still has much work to do in developing the invention. The result can be intuitively understood with an analogy to a raffle: No rational person would pay $40 to enter a raffle where the prize is $50 and at least one other person is certain to purchase a ticket. But if the cost of entry is reduced to $10, it is not irrational to play even if the prize is reduced by the same amount.

Reducing the cost of entry further makes the game more attractive. Thus, if the investment needed to meet the patent standard is reduced to $5, innovation is accelerated even more. Then only $95 in royalties would be sufficient to attract investment, so investment will occur before time t_I.

This simple model — or rather, the more complete version set forth in the mathematical appendix—can be used to estimate the delay caused by a threat of duplicative investment and the effect of early patenting on the delay. If early patenting is not allowed—i.e., if pre-patent expenditures are much larger than post-patent expenditures (*I* is much larger than *C*)—then the delay between the time when a firm would make the investment in innovation with no possibility of duplication ($t_{I\ 1}$) and the time when firms would invest with duplication by *N* firms ($t_{I\ N}$) will approach a limit given by the following:

(7) $$\text{For I} >> \text{C, delay in investment} = t_{IN} - t_{I1} \cong \frac{\ln N}{g}$$

where *g* is the growth rate of royalties, which is assumed to equal the general growth rate in the economy.[84] This result suggests that the threat of duplication can produce long delays in innovation. For example, in an economy with a growth rate of 5%, the delay in investment is *13.9 years* where N=2, and *32 years* where N=5. (Those delays are doubled in an economy with a growth rate of 2.5%.) In contrast, where early patenting policies are followed—i.e., *I* is negligible compared to *C* — and patent terms are finite, the delay caused by duplicative investment approaches zero.[85] Thus, the delay shown in equation (7) provides a fair approximation of the acceleration of

[84] See the appendix for a proof.

[85] See the appendix for a proof.

innovation produced by early patenting.[86]

This simple model shows that permitting early patenting — i.e., patenting of innovations that still require large investments to commercialize — reduces the social costs caused by duplicative efforts. If the patent scope were sufficient to capture the full value of the invention and patent terms were infinite, acceleration would not be preferable to duplication. Both would be equally wasteful. This is, in essence, the point made by the McFetridge and Smith critique, which followed Barzel's model in assuming infinite patent length. But patents are not infinite in term. They are partial property rights, limited in both scope and time. Thus, as competing innovators accelerate patenting earlier in time, they also lose a greater and greater portion of the rents associated with the innovation because the innovation will enter the public domain sooner. Limitations on patent scope may also decrease the portion of rents that may be captured by an early patentee. These positive externalities make acceleration of innovation socially preferable to duplication.

Under this view, the patent system can be usefully compared to Demsetz's proposal that, in industries with continuously declining average costs, the government should auction exclusive franchises to the bidder offering the best service in terms of price and quality. The insight of Demsetz's proposal is that competition to gain the exclusive right can be harnessed to reduce private

---

[86] Even where patent terms are infinite, early patenting policies trade duplication for acceleration. The effect is reduced, however. With late patenting policies, the delay caused by threat of duplication reaches the same limit shown in equation (7), but some delay remains even with early patenting policies. For infinite patent life, the minimum delay is:

$$\text{For } C >> I,\ \text{delay in investment} = t_{IN} - t_{I1} = \frac{\ln N}{r}$$

(See appendix for a proof.) Thus, for infinite patent lengths, switching from late to early patenting policies reduces delay by the quantity (ln N/g - lnN/r).

rents while increasing social surplus.[87] Applying a Demsetzian analysis to the patent system is appropriate since the marginal cost of using intellectual property, once it has been created, is usually assumed to be zero. Thus intellectual property presents a situation of continuously declining average costs.

Where patent rights are limited in time, a competition to patent earlier will resemble a Demsetzian auction in that, by trying to be first to patent, the competing inventors are also vying to diminish their rents by placing the patent in the public domain sooner.[88] Social surplus is also increased because the output restriction imposed by the patent ends sooner. As with Demsetzian auctions, competition for the exclusive right should drive the sum of the rents conferred by the exclusive franchise toward the total average cost of obtaining the good. Thus, the winning bidders or inventors will not, on average, obtain supracompetitive returns on investment.

There are three principal differences between Demsetzian auctions and patent races. First, the "auctions" created by the patent system occur only over the course of time. At any given moment, the intellectual property right is being offered at a specific set of terms: rents associated with the innovation will go to the patentee for a set number of years and to the public thereafter. But this temporal structure makes little economic difference. Since bidders are free to advance the time of innovation arbitrarily early or to delay innovation arbitrarily late (with early and late innovation

---

[87] If price discrimination is not possible and elasticity of demand is positive, deadweight loss will not eliminated because the bidders will offer service at their average, not marginal costs.

[88] I assume here that the patent is awarded to the inventor seeking the earliest expiration date. That assumption is accurate for most patent systems in the world because priority in those systems is awarded to the first to file and the patent term is defined as a number of years following the filing date. In the United States, the assumption is good approximation because the patent term is now defined as a number of years after filing and, even though priority is in theory awarded to the first to invent, in practice it is difficult to obtain a priority date much before the time of filing. The analysis set forth in this paper suggests that, if a first-to-invent priority is to be kept, the expiration date of a patent should be closely correlated to the priority date claimed.

correlated to, respectively, greater and lesser rents dedicated to the public), a full range of possible bids are available over the course of time.[89] Nor is running an auction over such a long period of time administratively difficult or costly, for no "bids" will be filed until the time when the discounted rents associated with the patent right begin to exceed the development costs.

Second, the patent system does not attempt to limit the patentee's monopoly power during the term of the patent. In contrast, Demsetz's proposal requires the government to insure that the winning bidder adheres to the price and (what might be most difficult) the terms of service proposed in its bid. While this feature reduces the administrative burden of the patent system, the simplicity comes at a cost: The deadweight loss during the term of the patent might be greater than under a Demsetzian system. On the other hand, after expiration of a patent, the cost of using the intellectual property falls to its marginal cost (zero) and deadweight loss is eliminated. It is an empirical question whether, in any particular case, the allocation of fixed costs produced by Demsetzian auctions will be superior to that produced by the patent system.

The final difference is that, in order to obtain the exclusive franchise, the patent system requires investment sufficient to produce an innovation meeting the patentability standard. Demsetzian auctions merely require investment sufficient to produce an auction bid. But it is precisely here that the prospect features of the patent system are important. If the patent system required an innovation to be completely developed prior to patenting, it would still resemble a

[89] While in theory the temporal structure of the patent auction could lead to inefficiently early or late development of a technology, these effects are likely to be insubstantial provided that (1) the costs of innovation are negligible compared to the costs of commercialization (i.e., pre-patent expenditures are much less the post-patent expenditures or $I \ll C$); and (2) the patent term, d, exceeds $[\ln (r/g)]/[r-g]$, where r is the relevant discount rate and g is the growth rate in the economy. The first condition ensures that technical development does not occur too early; the second, that it does not happen to late. Note that a twenty year patent term satisfies the second condition for many reasonable values of g and r. See the Appendix for further elaboration and proof.

Demsetzian auction process, but a very inefficient one. It would resemble an auction in which bidders were required to build their utility plants *before* bidding on the exclusive franchise. The inefficiency is obvious: Competitive bidding could not occur without the very wasteful duplication of production facilities that the Demsetzian system tries to avoid.

Early patenting policies increase the resemblance between the patent system and Demsetzian auctions. In a hypothetical auction for a utility franchise, bidders must invest resources to formulate their bids. They must determine their costs of constructing and maintaining plants, estimate future variable costs, project demand for service and develop or evaluate the contractual language specifying their obligations. These are largely expenditures for knowledge — specifically, for knowledge needed to define and evaluate the value of the franchise. The research necessary to obtain a patent is not radically different in kind, for determining the value of the patent franchise requires, as a prerequisite, the creation and definition of the technology that the franchise will cover. Early patenting policies merely insure that, once applicants have met that prerequisite, they will not have to make additional expenditures prior to award of the exclusive right.

Thus, the "prospect" features of the patent system have value not because patents resemble mineral claims, but because they do not. Mineral claimants can continue to exploit their claim until all mineral value is depleted. Patent claimants have their rights terminate at the end of the patent term even if, as is often the case, the intellectual property would continue to produce positive rents. Thus, patent competitions increase the share of the rents captured by the public — an effect wholly alien to the "gold rushes" created by the mineral claim system, but familiar to the competition fostered by a Demsetzian auction.[90] The prospect features of the patent system are important not for

[90] Indeed, a good analogy to the patenting system is not the mineral claiming system but the exclusive franchises awarded in the 18th and 19th centuries to encourage the production of bridges. Like patents, these

the suppression of rivalry, but for the accentuation and direction of it.

Probabilistic (Mixed Strategy) Model: The first model does not describe a Nash equilibrium. Each firm, if it believed that all other firms would wait until time $t_{I\,N}$ to invest, would have an incentive to "jump the gun" and invest at any time after $t_{I\,1}$, because the expected profits after $t_{I\,1}$ are positive if no other firm shares the prize. But if each firm has imperfect information concerning the strategies of its competitors (as is likely), each firm will rationally fear that others may also invest early, and that possibility will deter the firm from investing at time $t_{I\,1}$, though the firm may not wait until $t_{I\,N}$. The model could therefore overestimate the delay caused by a threat of duplication.

A model that permits firms to adopt mixed (probabilistic) strategies produces a symmetric equilibrium solution and provides additional insight into the value of early patent policies. In this model, each firm may purchase a probability *p* of achieving the innovation for an investment *I(p)*. As in the first model, if more than one firm achieves the invention, the patent is randomly assigned (or evenly divided) among the successful firms. A simple version of the model assumes that investment and probability of success are linearly correlated—just as they are, for example, in a lottery where players may choose numbers between one and one thousand, with one number selected as the "winning number." (Each additional ticket purchased increases the probability of winning by .001, provided the player does not make the mistake of duplicating numbers.) Thus, investment and probability of success are correlated:

(8) $$I(p) = p * I_T$$

---

franchises were limited in time, and like patents, a condition on the exclusive right was that, at the end of the franchise period, the right holder had to dedicate the relevant property to the public domain. *See*, *e.g.*, *Charles River Bridge v. Warren Bridge*, 36 U.S. 420, 423 (1837) (quoting the 1785 charter granted to the Charles River Bridge company, which provided for a forty year term and required that "at the end of the said term, the said bridge shall revert to, and be the property of the commonwealth").

where I_T is the investment needed to the obtain certainty (i.e., buy all of the tickets). If all firms pursue identical strategies, the expected profits for each firm will be:[91]

(9) $$\Pi(t,p) = \frac{1-(1-p)^N}{N}[R(t)-C] - I(p) = \frac{1-(1-p)^N}{N}[R(t)-C] - (p * I_T)$$

This model produces a symmetric equilibrium solution — i.e., there is a level of investment where, if all firms invest to that same level, no firm would have an incentive to change its investment if it knew the strategies of the other firms.[92] The equilibrium level of investment occurs where the profits of each firm equal zero—i.e. where each firm invests to achieve a probability of success p_o, as defined by the following equation:

(10) $$\Pi = 0 = \frac{1-(1-p_o)^N}{N}[R(t_p)-C] - I(p_o)$$
$$0 = \frac{1-(1-p_o)^N}{N}[R(t_p)-C] - (p_o * I_T)$$

That equation provides the relationship between the equilibrium level of probability p_o, the time when that probability occurs, t_p, and $R(t_p)$, C, and I_T (the expected royalties at time t_p, the costs of commercialization, and the investment needed for certain success). The equation can be reformulated in a form similar to equation (4) from the first model:

[91] Equation (9) can also be used to calculate the expected value of a standard lottery, where each of N players purchases sufficient tickets to obtain a probability p of winning; each ticket selects a number from a finite set of number combinations; I_T is cost of purchasing tickets corresponding to the entire set of possible number combinations; no player purchases duplicate tickets; and winnings are divided among all players picking the winning number combination.

[92] See the Appendix for a proof.

(11) $$R(t_p) = C + \frac{p_o N}{1-(1-p_o)^N} I_T$$

The factor $p_oN/(1\text{-}(1\text{-}p_o)^N)$ is basically a measure of the duplicative efforts; it ranges from values close to 1 (no duplication) for small probabilities, to values approaching N as the probability of success approaches 100%. The factor is the ratio of expected number of research successes (p_oN) to the probability of at least one research success $(1\text{-}(1\text{-}p_o)^N)$.

This model produces two important results. First, while early patenting policies always accelerate investment in innovation (as in the simple model), the amount of acceleration becomes small where the equilibrium chance of research success is small.

To see this point, we begin as before with the assumption that patent policy can affect only the allocation of costs between pre- and post-patent expenditures, not the total cost ($I_T + C$) needed to complete the technology. From equation (11), we can see that shifting dollars from I to C will decrease the right hand side of equation, because the quantity $p_oN/(1-(1-p_o)^N)$ is always greater than 1.[93] A decrease in the right hand side of the equation means that lower royalties are needed to achieve any given level of equilibrium probability. Since royalties are assumed always to increase with time, a lower level of royalties corresponds to an earlier time period, t_p. Thus, as in the simple model, shifting costs to C—i.e., allowing early patenting—leads to earlier investment in innovation.

The effect of shifting costs from I to C will, however, be insignificant where the probability of success is small and the number of firms is small to moderate. For example, where 10 firms are

[93] The denominator $(1-(1-p_o)^N)$ is equal to the probability of at least one success in N tries at a game with probability p_o of success, while the numerator (p_oN) is equal to the expected number of successes for N tries. Since the expected number of successes is always greater than the probability of one success, then the quantity $(p_oN/(1-(1-p_o)^N))$ will always be greater than 1.

engaged in research and the equilibrium probability of research success is 1/100, the factor $p_o N/(1-(1-p_o)^N)$ is less than 1.05, so shifting the entire cost of research from *I* to *C* decreases the level of royalties needed to achieve that equilibrium probability by less than 5%. Where royalty growth occurs at the same rate as economic expansion, royalties will increase by 5% in less than two years for an economy with a growth rate of 2.5%.

This result is easily interpreted: If the probability of research success is small and the number of firms is not too large, duplication is not a serious problem, and so policies that shift costs from the pre-patent stage to the post-patent stage of technological development make less of a difference. Such policies still have some value because the acceleration of investment could be inefficient. By limiting the amount of investment necessary to obtain a patent, early patenting policies insure that the dominant effect of a patent race is to diminish the private share of appropriable surplus (because the invention is dedicated to the public sooner), rather than to reduce the total social surplus by encouraging inefficiently early investment. In other words, early patenting policies still help to increase the efficiency of the temporal auctions created by patent races.

The second implication of this model is that, where probabilistic strategies are possible, the time of innovation becomes dependant on how long each "round" of a game takes and how the players' knowledge of prior game rounds evolves with time. While such inquiries are generally outside the scope of this paper, it is worth noting that, in the limit where each game round takes an infinitesimally small period of time to play, the players can make infinitesimally small investments in innovation, and the players instantaneously know the results of the prior round, the problem of duplication disappears and the result is the same as in Barzel's model: Innovation occurs as soon as the expected profits from the patent (R - C) exceed the investment needed to obtain the patent (I_T), and no resources are wasted on duplicative projects. This limiting case makes sense, since the limit approaches perfect knowledge. Because duplication is not a problem, the benefits of a prospect patenting policy will be limited to more modest effect noted in the last paragraph: The policy will help to insure that the dominant effect of a patent race is to reduce private rents, not to diminish the overall social surplus through inefficiently early investment.

The last point has significance. Fast communication of research success can minimize duplication of efforts and thereby further the same goals sought to be advanced by early patenting. This is another important limit on the prospect theory: Where government policies already permit early patenting, changes in policy to permit faster dissemination of information (for example, by early publication of patent applications) may be more valuable than policies that permit patenting at an even earlier stages of research.

C. Partial Property Rights and a Pioneer Patent Holder's Ability to Coordinate Further Investment in Innovation.

The limited scope of patent rights is the other major difficulty for the prospect theory. Unlike the fully exclusive rights given to mineral claimants, patentees obtain only exclusionary rights; future inventors may claim "blocking" exclusionary rights on improvements to an earlier patented technology. Thus, after the grant of a patent, the right to innovate remains a common right even within the claims of the patent.

The prospect theory provides no explanation for this fundamental limit on patent rights. Kitch's theory is that, "by awarding exclusive and publicly recorded ownership of a [technological] prospect shortly after its discovery,"[94] the patent system, like the mineral claim system, puts the patent holder "in a position to coordinate the search for technological and market enhancement of the patent," and that this "increases the efficiency with which investment in innovation can be managed."[95] But why then does the patent system allow other inventors to continue rivalrous and uncoordinated searches for patentable improvements within the bounds of existing patent claims?

The view advanced here provides an answer: The patent system does not seek to cut off rivalry, but only to direct it away from duplication and toward acceleration of innovative efforts. The partial property rights granted to a pioneer patent holder allow for some coordination of further innovation but, as discussed in an earlier paper, this power to coordinate is contingent on the pioneer continuing to accelerate innovation to a time earlier than when uncoordinated rivalry would produce the innovation.[96] The firm holding the pioneer patent will not have complete control over the timing of investments in patentable improvements because it will not be insulated from competition. But that is as it should be, because the competition for improvement patents forces the pioneer to innovate early.

The competitive limitation on the pioneer's power to control further innovation is consistent with the overall view presented in this paper, for the patent law does not attempt to preserve royalties by shutting off rivalry in innovation—it seeks only to insure that rivalry is directed toward activities, such as accelerated innovation, that are likely to have positive externalities and thereby to dissipate private rents faster than social rents.

---

[94] Kitch*, supra* note 1, at 266.

[95] *Id.* at 276.

[96] John Duffy, *A Prospect Patent's Power to Coordinate Further Innovation*, pp. 2-3 (Feb. 2001 draft).

III. Policy Implications.

The view set forth in this paper suggests that the patent system attempts not to solve the common pool problem nor to limit rivalrous dissipation of rents associated with a patent grant but to channel rivalry into activities with large externalities. Like a Demsetzian auction, the patent system thereby insures that the competition diminishes private rents faster than social rents. This view of the patent system has important policy implications, many of which diverge from the implications of the prospect theory.

A. The Correlation Between Patent Priority and Patent Expiration.

Patent races will serve the same purposes as Demsetzian auctions — channeling competition toward decreasing private rents while preserving social surplus — if the law recognizes as the winner in the patent race the inventor seeking the patent with the earliest expiration date. In other words, a central goal of the patent law should maintain a strict correlation between patent priority and patent expiration. In the United States, the historical trend has been toward a tighter correlation between patent priority and patent expiration; in other countries, a strict correlation between priority and expiration already exists.

For most of the twentieth century, the term of U.S. patents was seventeen years from the date on which the patent *issued*.[97] Patent priority, on the other hand, was determined by the filing date or, under certain circumstances, by an even earlier date of invention. Thus, patent priority and patent expiration were not necessarily correlated. Patent applicants had an incentive to file early to gain priority in the patent race, but then to delay issuance of their patents as long as possible to

[97] Id. at 41-42.

capture the larger rents associated with later periods of time. Some of these so-called "submarine" patents would surface *decades* after they were filed.

In such circumstances, the patent system failed to function as an efficient Demsetzian auction because the winner of the patent race was not necessarily the party willing to provide the patent to the public for the least rents. Patent applicants had an incentive to delay the patent office bureaucracy. While such a strategy realized private returns, it did so only by increasing the rents that would be paid by consumers. It was rent-seeking activity that generated negative, not positive externalities.[98]

In adopting the TRIPs treaty in 1994, the United States changed this feature of domestic patent law. Patents filed on or after June 8, 1995, will expire twenty years after the date of the patent application.[99] Because the United States follows a "first-to-invent" rule for establishing patent priority (granting priority for the earliest date of invention, not the earliest filing date), this change does not perfectly correlate patent priority and patent expiration. But in practice the correlation is usually fairly good, because American patent law treats the filing date as the presumptive date of invention and numerous doctrines restrict the ability of an inventor to establish any earlier priority date.[100]

---

[98] The Federal Circuit has recently embraced a legal doctrine of "laches" which can invalidate a patent if the inventor delayed the issuance of the patent for too long a period of time. See *Symbol Technologies. Inc. v. Lemelson Medical, Education & Research Fnd.*, 277 F.3d 1361 (Fed. Cir. 2002). This doctrine helps to maintain a correlation between patent priority date and patent expiration date, although it is more administratively complex than a simple rule that calculates the patent expiration date based on the priority date (i.e., the rule now used by nearly every country in world).

[99] See 35 U.S.C. § 154(a)(2). The change was necessary to conform U.S. law to Article 33 of the TRIPS agreement (*see* Agreement on Trade-Related Aspects of Intellectual Property Rights, Dec. 15, 1993, 33 I.L.M. 81 (1994)), which requires the United States and other signatories to provide a term of patent protection extending at least twenty years from the application filing date.

[100] *See, e.g.*, MARTIN J. ADELMAN, PATENT LAW PERSPECTIVES §2.3 at n.36 (2d ed. 1992) (noting that

In most major foreign patent systems, including the European and Japanese systems, priority is awarded to the first inventor to file a patent application.[101] Because patents expire twenty years after the filing date in those systems, the patent priority and patent expiration are strictly correlated, and those patent systems are better able to function as an efficient Demsetzian auction. This analysis does not necessarily suggest that a first-to-file priority system is better than a first-to-invent system. But it does suggest that, if a first-to-invent priority is to be kept, the expiration date of a patent should be closely correlated to the priority date claimed.

The goal of correlating priority and expiration dates has one other policy implication: Patent terms should not be extended retroactively — i.e., any legislated extension of a patent term should not apply to already completed inventions. Here the Demsetzian auction analogy is particularly instructive. Retroactive patent term extensions are comparable to post hoc renegotiations of the auction terms under a Demsetzian system. If such renegotiations are permitted, then rents associated with the exclusive right will, in part, be dissipated by expenditures on political lobbying and, as Oliver Williamson has noted, a Demsetzian auction system begins to exhibit some of the undesirable features of administrative regulation.[102] Although nineteenth century patent law did have a general administrative mechanism for extending patent terms (and Congress was also more willing to grant

---

"in reality the United States has a modified first-to-file system" because the statutory bars in 35 U.S.C. § 102(b) establish "a basic and large class of prior art . . . that cannot be challenged byt he inventor based on his date of invention"); Edward G. Fiorito, *Highlights of Selected Recommendations of the Advisory Commission on Patent Law Reform*, 1 Tex. Intell. Prop. J. 11, 16 (1992) (noting the sense in industry that "there is already a de facto first-to-file system in place" in the United States since "less than one percent of applications get into interference; and most of the time the senior party [the first to file] wins").

[101] ROBERT PATRICK MERGES, PATENT LAW AND POLICY 37 (2nd ed. 1996).

[102] *See* OLIVER E. WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM 352-364 (1985) (setting forth a case study of CATV franchise bidding in Oakland, California, where award of the initial franchise merely marked the beginning of efforts to lobby political bodies for a renegotiation of the franchise terms; concluding that "in practice, franchise bidding for CATV . . . has many of the qualities of regulation").

special extensions),[103] the patent terms today are relatively well defined, with extensions of term being permitted only for specific regulatory delays that are largely outside the control of the patentee.[104]

B. How Long?: Patent Term Length

The classic analysis of patent term length is found in William Nordhaus's work, which views the problem of optimal patent term in terms of achieving a balance between the incentives necessary to encourage innovation and the inefficiencies associated with the exclusive right:

> As the [patent] life is increased, two opposite forces affect the level of economic welfare. First, a longer life increases invention and thus gives on balance a larger amount of output for a given level of inputs. This is a positive effect. Second, a longer life means that the monopoly on information lasts longer and thus there are more losses from inefficiencies associated with monopoly. The optimal life of a patent is that point at which the two forces balance at the margin.[105]

This approach has at least two major flaws. First, it is not sensitive to the actual structure

---

[103] *Bloomer v. McQuewan*, 55 U.S. 539 (1852) (noting multiple extensions of a patent on planing machines through both administrative renewal of the patent and special legislation by Congress).

[104] *See* The Drug, Price Competition and Patent Term Restoration Act, codified in part at 35 U.S.C. § 156 (also known as the Hatch-Waxman Act).

[105] *See*, *e.g.*, WILLIAM D. NORDHAUS, INVENTION, GROWTH, AND WELFARE 76 (1969) (describing the optimal patent life as the point where the positive effect of increased invention balances at the margin with the greater losses from inefficiencies associated with the longer monopoly). Nordhaus's work remains a standard citation for a theory of optimal patent life. *See*, *e.g.*, Jean Tirole, THE THEORY OF INDUSTRIAL ORGANIZATION 400 (1988). Nordhaus is not alone in believing the issue of patent term length to be a trade off between the greater incentives for invention and increased social cost of a longer monopoly. *See* C. Michael White, *Why a Seventeen Year Patent?*, 38 J. PAT. OFF. SOC'Y 839, 840 (1956) ("Ideally the legal life of a patent should represent a balance between the additional incentive of another year and the social cost of a longer monopoly."); Richard A. Epstein, *Property Rights in cDNA Sequences: A New Resident for the Public Domain*, 3 U. CHI. L. SCH. ROUNDTABLE 575 (1996) ("[I]t hardly follows that his patent protection should be infinite either in scope or duration . . . . The basic social choice is strictly functional and instrumental. We must decide whether the incentives for the creation of these inventions and discoveries are large enough to justify the restrictions on output that follow when rights over invention or discovery are vested exclusively in the inventor and discoverer.").

of legal rights. Where the legal system allows patenting to occur before commercialization, a longer life for a patent need not generate either increased output or greater monopoly inefficiencies. Rather, longer life may mean merely earlier innovation followed by a longer delay in commercialization. Second, the approach takes a static view of innovation — i.e., it does not take into account the inefficiencies associated with either early or late innovation.

The actual model employed by Nordhaus compounds these problems. The model eliminates any consideration of competition between inventors by explicitly assuming that innovators can set their level of research to maximize profits.[106] Moreover, the forces leading to innovation are completely static under the model, so investments in innovation occur either at time zero or never. Thus, to derive an optimal patent life, the model assumes that initial investments in innovation generate supracompetitve returns, and that increased investment suffers diminishing returns at a rapid rate (under Nordhaus's assumptions, the first unit of research expenditure yields a return greater than the next 1000 units combined).[107] Increasing the patent term creates more marginal inventions but at the cost of extending the output efficiencies for the infra-marginal inventions that will be induced by even very short patent lives. As a result, the model predicts that patent terms of *six months to one year* will produce 80-90% of all possible welfare gains from a patent system.[108] That counterintuitive prediction is directly traceable to the assumptions that some infra-marginal

---

[106] *See* id. at 74, equation 5.4 (assuming that an inventor has the ability to choose the level of research that maximizes profits).

[107] Nordhaus assumed that the returns to investment in innovation increase with the 10th root of the investment. Id. at 80 (assuming the cost reduction attributable to innovation is a function of research expenditures to the α power, where $\alpha = .1$). Thus, the first unit of research expenditure yields the same return as the next 1023 units combined.

[108] *See* id. at 83-84.

investments in innovation yield supracompetitive returns and that competition cannot push very fruitful investments in innovation to earlier time periods.

If innovation is modeled as a process that is both dynamic and competitive, then patent racing will tend to drive innovators' profits to zero. Moreover, *all investments in innovation will be marginal at the time they are made*. Determining the optimal patent length is then transformed into an issue of timing: Longer patent terms may encourage inefficiently early innovation, and shorter patent terms may cause innovation to be inefficiently late. With the assumptions of the Barzelian model for innovation, a minimum value for an optimal patent term can be calculated to be:

$$d_{\min} = \frac{1}{r-g}\ln\frac{r}{g}$$

where r is the discount rate and g is the rate of economic expansion.[109] This minimum time will be the optimal patent term in the best case scenario where the patentee captures all social surplus from the invention during the patent term and no resources are wasted on duplicative investment.

Three points are worth noting about this minimum for the optimal term. First, if the patent term is shorter than this minimum, innovation will occur inefficiently late even where the patentee captures all social surplus during the patent term and no duplicative investment occurs. In other words, the effect identified by Barzel of inefficiently accelerated innovation will not occur in a patent system with term lengths at or below this minimum. Because the patent terms traditionally allowed under Anglo-American law (i.e., between 14 and 20 years) are quite close to this minimum

[109] See Appendix, Equation 28 and accompanying text. This threshold patent term approaches a lower limit of $1/r$ where g approaches r. For a complete discussion, see John F. Duffy, *A Minimum Optimal Patent Term* (2003) (available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=354282).

threshold,[110] the patent system can be structured to encourage accelerated innovation as relentlessly as possible without the effect identified by Barzel becoming a major problem.

Second, if the innovator is not able to capture the full social surplus during the patent term (i.e., where deadweight loss exists because the patentee is unable to engage in perfect price discrimination), the optimal patent length is *longer* than the minimum given above. The analysis in Nordaus suggests that, where deadweight loss exists, the optimal patent term should be *shorter* than the optimum term for situations where there is no deadweight loss. That analysis, however, does not hold under a dynamic model where innovators compete for patents.

To see this, consider a patent system that has a patent term of X years, which is the minimum optimal patent life under the analysis set forth above. Where the patentee cannot engage in perfect price discrimination and the elasticity of demand is positive and finite, the patentee will not be able to capture the full social surplus, and the patent term of X years will induce innovation at an inefficiently late time. A consumer surplus will exist, and the exclusive right will create some deadweight loss. Would decreasing the patent term increase social welfare? The answer is clearly no.

Decreasing the patent term will push both the time of innovation and the time of patent expiration to a later time. The change does not affect the welfare of the innovator because, by assumption, competition drives profits to zero. But social welfare is decreased. Because the time of patent expiration is later, deadweight loss is created during a time period when it did not previously exist. In addition, some consumer surplus is lost because the time of innovation is later. Thus, shortening the patent term does not, as Nordhaus assumed, reduce the inefficiencies associated

[110] For realistic discount rates and economic growth rates (i.e., $5\% < r < 10\%$ and $2\% < g < 5\%$), the threshold minimum ranges from roughly 13 years to 30 years.

with monopoly; rather, it introduces those inefficiencies into a period when they would otherwise have been eliminated.

A third and final point is that, where the investment needed to achieve the patentable innovation is much less than the investment needed to commercialize, increasing patent terms above the optimal term may have little effect. Competition for the exclusive franchise will push back the time of patenting, but will leave largely unaffected the time of commercialization and the time of patent expiration. The expected rents associated with the franchise will still equal the total average costs of producing the innovation, and the output inefficiency associated with the exclusive right will be largely unaffected. Keeping patent terms close to the optimal is desirable mainly because shorter patent terms economize on administrative costs — i.e., longer patent terms require the patent rights to be defined sooner, which is likely to be more expensive.

C. How Early?: No Need to Avoid Rent Dissipation

Because the prospect theory was formulated as a response to Barzel's common pool problem, it stressed the early grant of broad patents. No limits on how early the grant should occur were acknowledged. In his response to McFetridge and Smith, Kitch made clear that his theory seeks to push patenting back to "the early stages of innovation"—a point early enough to avoid rent dissipation. So too Grady and Alexander have sought to push patenting back to a time when a "few [prospectors] are inherently swifter than most" so that "[r]ents are preserved."[111] But to push patenting ever earlier in the hope of preserving rents is to chase a phantom. There has been no satisfactory demonstration that rivalrous rent dissipation can be avoided where the right to invent

[111] Grady & Alexander, *supra* note 12, at 317.

remains a common right, and the evidence shows quite the contrary.

If patent races are seen as Demsetzian auctions, the press for earlier patenting eases. The goal of the patent law is then to push patenting back to a time not before competitive pressures are felt, but before much duplication of effort occurs. While competition among even a few rivals may be sufficient to dissipate all rents associated with a patent, it may not produce much duplication if the probability of success for each research approach is low. Indeed, some evidence suggests that, in researching difficult technical problems, the value of avoiding possible duplication can be outweighed by the costs of coordinating.[112]

Moreover, while allowing inventors to patent in the early stages of R & D efforts tends to create greater efficiency by trading duplication for acceleration, other factors begin to counterbalance this effect as patenting is pushed earlier and earlier. The first, and most obvious, of these offsetting factors is the steep rise in administrative costs because of the difficulties in defining and enforcing rights to an embryonic invention.

Another effect limiting the benefit of very early patenting is that, in the early stages of research, firms may communicate which research avenues they are pursuing. Kitch assumed the opposite. He argued that secrecy would be a particularly acute problem in technological research and that it would exacerbate the common pool problem identified by Barzel: "[U]nlike fisheries, public roads, and other types of goods usually considered, technological information can be used

---

[112] For example, Western Union underwrote research on telephony by at least two separate researchers, Thomas Edison and Elisha Grey, without requiring coordination of efforts. *See* Josephson, *supra* note 54, at 139. (Another Western Union researcher, G.W. Phelps, was also assigned to investigate the technology, but he may have directed Grey's work. *Id.* at 138-39 & 80. Edison does not appear to have collaborated with the Grey/Phelps team.)

without signaling that fact to another."[113] While "[f]ishing boats can be detected, and one who is considering entry can take into account the magnitude of his competitor's activities," this is not so in the area of technological innovation, Kitch claimed, because "it is possible for a firm working in secrecy to enter upon a 'prospect,' investigate it extensively, and depart without a trace."[114] Yet a firm may have difficulty investigating a technological frontier in complete secrecy. Most obviously, the firm will have to enter a rarified labor market—the market for PhDs and other highly educated professionals in that particular field. Competition for qualified workers will itself convey information to other firms about the technological prospects being investigated by the firm.

Moreover, even if a firm *could* investigate technological prospects in secrecy, disclosure may be in the firm's best interest in the early stages of research, because the disclosure may deter entry by competitors more than it attracts them. The models set forth in Part II assumed that the players would not communicate with each other. This assumption is justified if the private costs of communication outweigh the private benefits, but that assumption may become less appropriate in the early stages of research.

Two effects must be considered. By communicating the research projects that it is pursuing, a firm may signal others that it has information (possibly from prior research projects) that the particular project is promising, and this would attract competitors. But identifying research projects also tends to discourage other firms from engaging in similar projects because, if two firms of similar competence choose the same project, then each firm's expected return decreases by one-

---

[113] Kitch, *supra* note 1, at 276.

[114] *Id.*

half.[115]

In the very early stages of research, the later effect may dominate: If a firm has not yet done much work investigating the prospect, other firms may not interpret the selection of the project as indicating some special information about the project's potential. Also, in the early stages of research, there may be many possible avenues of research with similar (low) probabilities of success. A firm could then easily avoid the avenues chosen by competitors in favor of research projects not yet investigated.

Even in later stages of research, prospective inventors may have incentive to communicate at least the *area* in which they are working, if such a disclosure would not disclose information useful to competing inventors. For example, in the race to invent a practical incandescent light, while prospective inventors "were not aware of the detailed progress of the their rivals, . . . [they] each knew that others were entered in the competition."[116] And, when Thomas Edison decided that he too would enter the race to invent, he disclosed his intention in a newspaper interview, declaring "I have let the other inventors get the start of me in this matter, somewhat, . . . but I believe I can catch up to them now."[117] Such disclosures allow other firms considering entry into a field to take into account the efforts of others.

In sum, awarding patents too early may reduce duplicative efforts little while increasing

---

[115]For example, consider a game where players can choose a number between 1 and 100, with the winning number chosen randomly and the prize divided among winners if more than one player chooses the winning number. If the first player chooses 66 and discloses that to the other players, then the expected value of choosing 66 for the next player is one-half the expected value of the others numbers.

[116] PAUL W. KEATING, LAMPS FOR A BRIGHTER AMERICA 9 (1954).

[117] ARTHUR A. BRIGHT, JR., THE ELECTRIC-LAMP INDUSTRY: TECHNOLOGICAL CHANGE AND ECONOMIC DEVELOPMENT FROM 1800 TO 1947 58 (1949) (quoting NEW YORK TRIBUNE, Sept. 28, 1878, p. 4.)

administrative costs for defining and enforcing the right. The timing should be based on the need to avoid duplication, not on a quixotic hope of limiting rivalry to preserve rents.

D. Compulsory Licensing, Blocking Patents and Experimental Use.

Compulsory licensing proposals would, under certain circumstances, compel pioneer patentees to grant licenses to improvers who obtain patents within an earlier pioneer patent. As Kitch noted, the effect of compulsory licensing is that "[t]hird parties can search for ways to increase the value of the patent and when they find it force the owner to license the patent at the regulated rate."[118] Kitch criticized such proposals as "destroy[ing] the prospect function."[119]

While it is true that compulsory licensing may further facilitate uncoordinated searches by third parties, such searches are already encouraged by the patent doctrines that permit third parties to experiment with patented technologies and to obtain blocking patents within the claims of existing patents. Kitch conspicuously failed to consider these features of patent law, even though they cause the same type of damage to the prospect function as would be caused by compulsory licensing.[120] For a patent system with such features, compulsory licensing cannot be dismissed merely because it would encourage third parties to search for improvements to issued patents.

Compulsory licensing proposals can be evaluated only with an understanding of why the

---

[118] Kitch, *supra* note 1, at 287. For purposes of this article, compulsory licensing proposals will refer to proposals that force licenses only in a case of blocking patents.

[119] *Id.*

[120] *See* Lemley, *supra* note 2, at 1047 (noting the inconsistency between the allowance of blocking patent and the prospect theory); Rebecca S. Eisenberg, *Patents and the Progress of Science: Exclusive Rights and Experimental Use*, 56 U. CHI. L. REV. 1017, 1044 (1989) (noting that "an experimental use exemption [to infringement] arguably undermines the prospect function" because it permits "unauthorized (and uncoordinated) research" by inventors seeking to improve the prospect patent).

patent system permits experimental uses and blocking patents. Those features of the patent law are best explained by the view presented here—that the patent system does not try to prevent uncoordinated competition in innovation, but only to force competition toward acceleration of innovation. Permitting uncoordinated searches for innovations to continue within the claims of an earlier patent then makes sense. Competition for patentable improvements to existing patents is little different from the competition for the initial patent. With both, rents are dissipated. With both, acceleration is more likely to have externalities than duplication. Increasing a patentee's rights over patentable improvements merely trades rent dissipation at the improvement stage for rent dissipation in seeking the pioneer patent. If both forms of rent dissipation have similar externalities, no social gain is realized.

Under current law, a pioneering firm will have some power to coordinate further innovation to the extent that the costs of prior coordination are lower than the transactions of licensing blocking patents. Compulsory licensing may decrease the transaction costs of blocking patents and thereby decrease the pioneer's coordination power somewhat. Compulsory licensing would not, however, eliminate transaction costs. Most obviously, both sides would bear legal costs associated with the compulsory licensing system. Because transaction costs continue to exist, the pioneering firm still has a greater incentive to make improvements than all other firms, and this is true without regard to the level royalties set in the compulsory licensing proceeding. Thus, even though the pioneer's power to coordinate further innovation—which is already checked substantially—may merely be further diminished by compulsory licensing, it would not be eliminated. Compulsory licensing does not effect a categorical change, and it cannot be rejected simply because it is inconsistent with the prospect theory.[121]

E. The Standard for Invention.

Under blackletter patent law, a researcher must, to receive a patent, discover something not only that is new and useful, but also that meets some threshold standard of inventiveness. Since the novelty requirement is uncontroversial, and utility is found upon a quite trivial showing, the requirement of inventiveness is considered the "final gatekeeper of the patent system" and "the ultimate condition of patentability."[122] Traditionally, this requirement ("nonobviousness" or "inventive step" as it is now called in, respectively, the American and European patent systems) has been justified as a corollary to the "reward theory" of patent law: New and useful, but also relatively obvious, creations do not deserve the reward of a patent because the social benefits of the invention are outweighed by the social costs of the patent monopoly.[123]

Though Kitch claimed that his prospect theory was proposed only as a supplement to the reward theory,[124] his discussion of inventive standard demonstrates how far his theory departs from the traditional reward theory. Kitch claimed that "the arguments for a property right in technological

---

[121] Other reasons may compel rejection compulsory licensing proposals. If, for example, compulsory licensing royalties are set too low, the value of improvement patents may be high enough to trigger excessive duplication.

[122] ROBERT PATRICK MERGES, PATENT LAW AND POLICY 479 (2nd ed. 1996); JOHN F. WITHERSPOON, NONOBVIOUSNESS—THE ULTIMATE CONDITION OF PATENTABILITY: PAPERS COMPILED IN COMMEMORATION OF THE SILVER ANNIVERSARY OF 35 USC 103 (1980).

[123] The best historical account of the development of the nonobviousness doctrine remains Kitch's 1966 article, *Graham v. John Deere Co.: New Standard for Patents*, 1966 SUP. CT. REV. 293. In his prospect theory article, Kitch recanted his conclusions from this article. *See* Kitch, *supra* note 1, at 282 ("my 1966 effort to derive a manageable standard from the economic literature was a failure").

[124] Kitch, *supra* note 1, at 267.

information *all* depend on the assumption that investment in the search for ways to enhance the value of the information is needed."[125] Ultimately, he endorsed a "substantial novelty" test for invention because, he believed, any substantially new information may need some inquiry into ways to enhance its value. Such a view justifies patent protection based solely on post-patent effects; it does not consider the pre-patent incentives to invent that may be created by the patent as a reward.

Grady and Alexander have taken Kitch's views on patentability to their logical conclusion. Because the prospect theory justifies patents solely on the need to coordinate and control searches for further enhancements, the theory leads to a sort of reverse Occam's Razor—"a simple and elegant solution" incapable of improvement should receive no patent protection.[126] That conclusion is bizarre, and it highlights the prospect theory's inability to replace the reward theory as the central justification for the patent system. For if holding out a patent reward limited in both time and scope encourages a simple and elegant solution, society is better off providing the reward in exchange for the solution.

Under the view presented in this paper, the reward theory still provides the central justification for the patent system. The so-called prospect features of the patent system—especially the preference for early patenting of technical discoveries—have a subservient function: They channel the inevitable competition for patent rewards into the accelerated creation of nonobvious developments, which are likely to have the largest externalities. Races to invent can be viewed as bidding wars between the competitors, with the patent rents going to the firm willing to offer the

[125] *Id.* at 283.

[126] Grady & Alexander, *supra* note 12, at 325; *see also id.* ("Just as philosophers have applied "Occam's Razor" to their analytical problems, so have patent courts applied its reverse.").

innovation to society in exchange for the least valuable (i.e. earliest) franchise period.

The innovation for which society is conferring such a reward must, however, be of sufficient importance so that externalities do exist. If rewards are provided for obvious developments—developments that would have been made without the patent system—then competitors will compete to create those. The benefit to society for such obvious creations may be slight, because the creation would have occurred without the patent incentive and because the exclusive rights of the patent will impose some social costs. Thus, the nonobviousness test for invention recognized by current law has advantages over Kitch's substantial novelty test.

IV. Conclusion.

The prospect theory focuses attention on the timing of patent grants and provides the insight that early patenting serves an important economic function. However, the theory does not provide a satisfying solution to the common pool problem identified by Barzel.

This paper has offered an alternative explanation for the patent law's preference for granting patent rights in the early stages of technical development. Early grants of patent rights channel rent seeking behavior into accelerated efforts to innovate. Because of the limits on patent scope and term, innovation has positive externalities for society. Although competition dissipates all private rents associated with a patent grant, some social surplus is preserved, and patent races are harnessed to serve a function similar to the Demsetzian auction process. Such a system is rational given the difficulty of thwarting rivalrous dissipation of rents associated with innovation. The view of the patent system presented here is a realistic one, which recognizes both the special constraints of creating property rights in innovations and the basic differences between patent law and prospecting systems with contrasting structures.

Mathematical Appendix

Appendix for Part II: Models for Duplication and Acceleration in Patent Races.

The two models presented in this section examine the effect that early patenting policies have on the acceleration and duplication of efforts in patent races. The models assume that innovators must make pre-patent investments (I) to meet the patent standard, and post-patent investment (C) to commercialize the product. The total cost of producing the innovation ($I + C$) cannot be changed by government policy. The government can, however, affect the allocation of costs between pre- and post-patent stages (between I and C) by choosing to permit patenting in the early stages of research and development, when significant additional expenditures are still needed to develop a commercial product incorporating the relevant innovation, or in the late stages, when little additional development is needed. The models also assume that (1) patents may be finite in length, (2) competing innovators may duplicate each other's efforts, and (3) the competing innovators cannot communicate with each other to coordinate their pre-patent activities.[127]

1) *Simple Model for Effect of Duplicative Efforts and Finite Patents.*

A first, very simple model is based on the assumptions (i) that all firms competing in the patent race pursue pure (non-probabilistic) strategies and (ii) that each firm assumes that its own decision to invest in innovation will be duplicated by the N other firms in the industry.

The model assumes that a research investment I undertaken at time t_I will produce an invention that meets the legal standard of patentability. The value of the patent is $R – C$, where R and C are, respectively, the discounted royalties from the invention and the discounted costs of

[127] See Tandon, at 153 (similar model also assuming no communication between firms).

commercializing the invention. The stream of royalties does not begin until the patentee makes the investment necessary to commercialize the invention, which occurs at time t_C. Royalties are assumed to grow exponential according to the growth rate of the economy, *g*, which is assumed to be less than the discount rate, *r*.

With these assumptions and definitions, a firm making an investment in research expect that its investment will be duplicated by *N* other firms each of which has an equal chance of winning the patent right, then the firm could expect profits of:

(1) $$\Pi = \frac{R - C}{N} - I$$

where *R*, *C*, and *I* are defined as:

(2) $$\begin{aligned} R &= \int_{t_C}^{t_I + d} R_o e^{-(r-g)t} dt \\ &= \frac{R_o}{(r-g)} \left[e^{-(r-g)t_C} - e^{-(r-g)(t_I + d)} \right] \\ C &= C_o e^{-rt_C} \\ I &= I_o e^{-rt_I}. \end{aligned}$$

The time period *d* is the length of the patent term which, for convenience, is assumed to run from the time of innovation. R_o is the base rate of royalties per year, with royalties increasing at rate *g*. The model assumes that innovation and commercialization can take occur at any time for lump sum expenditures I_o and C_o. Both royalties and expenditures are discounted to present value at the same rate *r*.

The patentee will try to set time t_C, the time of commercialization, to maximize the value of the patent, *R* - *C*, subject to the constraint that t_C \$$t_I$. The patent value will be maximized at t_C*

as defined by the following:

(3) $$\left.\frac{d(R-C)}{dt_C}\right|_{t_C^*} = 0$$

Substituting from the equations in (2) and differentiating produces:

(4) $$R_o e^{-(r-g)t_C^*} = rC_o e^{-rt_C^*}$$, which yields

(5) $$t_C^* = [\ln(C_o r / R_o)] / g\ .$$

Time t_C will equal t_C* where t_C* $ t_I, and will equal t_I otherwise. Intuitively, this means that, where large post-patent investments are required to begin the royalty stream (C_o is large), the patentee may find it more profitable to delay commercialization until the rate of expected royalties equals the interest stream from C_o. In all other circumstances, the patentee will maximize profits by commercializing as soon after innovation as possible.

If competition drives profits to zero (Π= 0), the time of invention t_I can be determined by the following equation:

(6) $$\Pi = \frac{R-C}{N} - I = 0$$

or:

(7) $$R - C - NI = 0$$

Solving equation (7) for t_I produces different results for the case where commercialization occurs at time of invention (i.e., t_C* # t_I, so $t_C = t_I$), and the case where commercialization occurs at a later time (i.e., $t_C^* > t_I$, so $t_C = t_C^*$).

1.a) *Early Commercialization.*

Consider first the range where t_C* # t_I, so that $t_C = t_I$. Time of invention then can be determined as follows:

(8) $$0 = \frac{R_o}{(r-g)}\left[e^{-(r-g)t_I} - e^{-(r-g)(t_I+d)}\right] - C_o e^{-rt_I} - NI_o e^{-rt_I}$$

(9) $$0 = \frac{R_o}{(r-g)} e^{gt_I}\left[1 - e^{-(r-g)d}\right] - C_o - NI_o$$

(10) $$e^{gt_I} = \frac{(r-g)(C_o + NI_o)}{R_o(1 - e^{-(r-g)d})}$$

(11) $$t_I = \frac{1}{g}\left[\ln(r-g) + \ln(C_o + NI_o) - \ln R_o - \ln(1 - e^{-(r-g)d})\right]$$

From equation (11), the delay caused by threat of duplication by *N* firms can be calculated to be:

(12) $$delay = t_{IN} - t_{I1} = \frac{1}{g}\left[\ln\left(\frac{C_o + NI_o}{C_o + I_o}\right)\right]$$

For $C_o << I_o$ (i.e., where early patenting is not permitted), delay will approach the following limit:[128]

(13) $$delay = t_{IN} - t_{I1} \rightarrow \frac{1}{g}\ln N, \text{ as } C_o \rightarrow 0$$

For $I_o << C_o$ (i.e., where early patenting is permitted), threat of duplication will cause negligible delay:[129]

$$delay = t_{IN} - t_{I1} \rightarrow \frac{1}{g}\ln\frac{C_o}{C_o} = 0, \text{ as } I_o \rightarrow 0$$

Thus, prospect features of the patent law can accelerate innovation by a maximum of lnN/g.

1.b) *Late Commercialization.*

Where t_C* is greater t_I, equation (7) will produce a different solution, though the limits of delay are not changed if patents are finite. Substituting equations (2) into equation (7) produces the following:

---

[128] Note that, where $C_o << I_o$, this limit for delay caused by threat of duplication applies in the case of infinite patent length. Where $C_o << I_o$, equation (11) accurately defines the time of innovation for infinite patent length, although the last term in the expression goes to zero as d goes to infinity.

[129] In the case of infinite patent length, the delay caused by threat of duplication does not go to zero because, where $I_o << C_o$ and patent term is infinite, the patentee will always be able to commercialize at time t_C* rather than at t_I.

(14) $$0 = \frac{R_o}{(r-g)}\left[e^{-(r-g)t_C^*} - e^{-(r-g)(t_I+d)}\right] - C_o e^{-rt_C^*} - NI_o e^{-rt_I}$$

(15) $$R_o e^{-(r-g)(t_I+d)} + NI_o(r-g)e^{-rt_I} = (R_o e^{-(r-g)t_C^*} - C_o r e^{-rt_C^*}) + C_o g e^{-rt_C^*}$$

Equation (4) shows that the first quantity on the right hand side on equation (16) is zero. Substituting the value for t_C given in equation (5) produces:

(16) $$R_o e^{-(r-g)(t_I+d)} + NI_o(r-g)e^{-rt_I} = +C_o g e^{-rt_C^*}$$

(17) $$\frac{R_o}{(r-g)}\left[e^{-(r-g)(t_I+d)}\right] + NI_o e^{-rt_I} = \left(\frac{rC_o}{R_o}\right)^{-\frac{r}{g}}\left(\frac{gC_o}{(r-g)}\right)$$

Equation (18) defines t_I where commercialization occurs at t_C*. For a finite patent term, t_I will approach the following limit as I_o approaches zero:[130]

(18) $$t_I = \frac{r/g}{(r-g)}\ln\left(\frac{rC_o}{R_o}\right) + \frac{1}{(r-g)}\ln\left(\frac{R_o}{gC_o}\right) - d$$

Since t_I does not depend on N, it is obvious that threat of duplication will cause no significant delay

[130] The first term in equation (17) is the discounted value of the royalties foregone because of the limited patent term; the second is the discounted pre-patent expenditures by all *N* firms. The expression for t_I given in equation (18) is accurate if the discounted pre-patent expenditures are much smaller than the discounted royalties foregone.

where I is sufficiently small and patent term is finite:

(19) $$delay = t_{IN} - t_{I1} \to 0, \text{ as } I_o \to 0$$

This result can be easily understood: Where patent terms are finite and pre-patent expenditures minimal, the dominant cost of accelerating innovation will be not the increase in the discounted cost of innovating earlier, but the loss of royalties due to the earlier expiration of the patent.

Note that where patent terms are infinite, then the first term in equation (17) is zero. Time t_I and delay $t_{IN} - t_{I1}$ are then given by the following equations:

(20) $$t_I = \frac{1}{g}\ln(C_o r / R_o) - \frac{1}{r}\ln[C_o g / NI_o(r-g)], \text{ for } d = \infty \text{ and } t_C^* \geq t_I$$

(21) $$delay = t_{IN} - t_{I1} = \frac{\ln N}{r}$$

Again, the last result is easily understood: If patent terms are infinite, the only cost of accelerating innovation will be to increase the discounted cost of pre-patent expenditures. Pre-patent expenditures are subject to wasteful duplication, and thus the threat of duplication by N firms will always cause some delay. Because time t_C* will always be later than time t_I if I_o is sufficiently small (see note 5, *infra*), equation (22) gives the minimum delay attributable to duplication that can be achieved for infinite patents.

For $C_o << I_o$ (i.e., where early patenting is not permitted), t_C* will always be less than t_I; commercialization will not occur at time t_C*, but at time t_I; and delay caused by the threat of duplication will be as shown in equation (13) above. To see this, we can use equations (5) and (11) to find the conditions under which the optimal time of commercialization, t_C*, will be earlier than the time of innovation, t_I:

(22) $$t_C^* < t_I$$

(23) $$[\ln(C_o r / R_o)] / g < \frac{1}{g}\left[\ln(r-g) + \ln(C_o + NI_o) - \ln R_o - \ln(1 - e^{-(r-g)d})\right]$$

(24) $$C_o r(1 - e^{-(r-g)d}) < (r-g)(C_o + NI_o)$$

(25) $$\frac{NI_o}{C_o} > \frac{g - re^{-(r-g)d}}{r-g}$$

Inequality (26) shows that the optimal time of commercialization will be earlier than the time of innovation where either of two conditions is satisfied.[131] First, NI_o is of large magnitude in comparison to C_o, where the relevant magnitude is defined by the right hand side of the inequality.[132] Thus, where I_o is very large compared to C_o ($I_o >> C_o$), t_C* will always occur before the time of innovation, and the time of innovation, t_I, and the delay caused by threat of duplication will be given by equations (11) - (13) above.

Because the quantities NI_o, C_o, and $(r - g)$ are positive, inequality (26) will always be satisfied where the numerator on the right hand side is zero or negative. Thus, t_C* will be earlier

---

[131] For infinite patent length, only one condition need be met for time t_c* to be later than t_I: NI_o / C_o must be less than $g / (r - g)$. Accordingly, for sufficient small I_o, the patentee will always be able to commercialize at the optimal time.

[132] The right hand side of the inequality will always be less than one provided that the growth rate is less than one-half the discount rate ($g < r/2$). Thus, for a broad range of growth rates NI_o will have to be less than C_o.

than t_I for all patent terms below a minimum determined as follows:

(26) $$g - re^{-(r-g)d} < 0$$

(27) $$d < \frac{1}{r-g} \ln \frac{r}{g} \equiv d_{\min}$$

Inequality (28) establishes a lower bound on the patent term necessary so that the patentee will commercialize the invention at the efficient time. Indeed, the significance of inequality (28) can be stated much more broadly. If $d < d_{min}$, then the time of innovation, t_I, will occur after the efficient time of commercialization, t_C*, *for any allocation of research and development expenditures between I and C.* Where the entire cost of research and development is allocated to C (i.e., $I_o = 0$), t_C* is, under the assumptions in the Barzelian model, the general socially optimal time for innovation and commercialization to occur. Thus, if the patent term is less than d_{min}, the effect of inefficient acceleration of innovation identified by Barzel does not occur, and the invention will be made inefficiently late even where the investment needed to obtain the exclusive patent right is infinitesimal and the patentee reaps all social benefits during the patent term.

For realistic values of r and g (i.e., $5\% < r < 10\%$ and $2\% < g < 5\%$), d_{min} ranges from roughly 13 years to 30 years.[133] The effect of inefficient acceleration identified by Barzel may not arise frequently with patent terms of lengths traditionally allowed under Anglo-American law (i.e., between 14 and 20 years).

2. *Probabilistic Model:* The probabilistic model varies the first model only in that firms may

[133] The term d_{min} has a lower bound of $1/r$, which it approaches as g approaches r.

purchase any given probability of research success. For simplicity, investment and the probability of research success are assumed to be linearly correlated according to the following:

(28) $$I(p) = I_T * p$$

The constant I_T corresponds to the investment needed to purchase certain research success. Again for simplicity, expenditure *I* will be assumed to occur at time t_I, and expenditure *C* to occur at a time infinitesimally later that t_I (the assumption eliminates the need for discounting of the expenditures). At any given time, each firm knows whether any other firm has yet succeeded in its research. A symmetric equilibrium will then exist.

Where all firms in the industry follow identical strategies of investing to obtain probability *p*, then the profits across the entire industry of N firms will equal the value of the patent (R - C) times the probability that at least one firm succeeded $(1 - (1 - p)^N)$, minus the investment made by *N* firms:

(29) $$\sum_{j=1}^{N} \Pi_j(t, p) = [1 - (1 - p)^N][R(t) - C] - NpI_T$$

Since all firms are identical, profits per firm are:

(30) $$\Pi(t, p) = \frac{1 - (1 - p)^N}{N}[R(t) - C] - pI_T$$

Define p_o as the value of *p* for which firm profits equal zero:

(31) $$\Pi_j(t, p) = \frac{1 - (1 - p_o)^N}{N}[R(t) - C] - p_o I_T = 0$$

A Nash equilibrium exists if, given the strategies of all other firms, no firm has an incentive to change its strategy. This condition is met where all firms invest with probability p_o. To see this, consider a firm deciding whether to change its probability of investing in an industry where all other firms invest with probability p_o. The profits of this firm would be:[134]

(32) $$\Pi_j(t, p_o, p_j) = p_j \left[\frac{1-(1-p_o)^N}{p_o N} [R(t) - C] - I_T \right]$$

Taking the derivative of this expression demonstrates that, if the strategy of all other firms is to invest with probability p_o as defined in equation (10) above, the firm would have no incentive to deviate from a strategy of investing with probability p_o:

(33) $$\frac{d\Pi}{dp_j} = \frac{1-(1-p_o)^N}{p_o N} [R(t) - C] - I = 0$$

Thus, there is a symmetric equilibrium in which all firms invest with a probability of p_o and profits for each firm are driven to zero.[135]

---

[134] A proof that equation (33) accurately represents the profits of the firm is available from the author.

[135] Under this model, firms also have no incentive *not to deviate* from that strategy. The equilibrium found here is consistent with Tandon's observation for a similar model that, "if it is profitable for a firm to sample even once, . . . it will sample the entire set of research lines." Tandon, 14 Bell J. Econ. at 159. If it is profitable (i.e., if equation (34) is *greater* than zero), then the firm will set the probability of undertaking the research project to 100%. Note also that, if there are diminishing returns to research investment (i.e. *dI/dp > I/p* for some *p*), then it is possible to have a symmetric equilibrium in which each firm has a disincentive to deviate from the equilibrium solution. As demonstrated by Tandon, that equilibrium will occur where the marginal and average costs of obtaining addition probability of research success are equal for each firm—or, equivalently, where average cost is at a minimum. See also Herman C. Quirmbach, *R&D: Competition, Risk and Performance*, 24 RAND J. OF ECON. 157, 159-160 (1993) (noting that, in a competitive market, research projects of all firms will be set to where average cost is minimized).