PRIVATE PROTECTION OF PATENTABLE GOODS

*Jonathan M. Barnett**

INTRODUCTION

The contracts literature has recognized that informal and extralegal enforcement mechanisms, in the form of arbitral tribunals, commercial monitoring agents, financial bonding mechanisms, reputational pressures and industry custom, play a crucial role in inducing widespread compliance with contractual obligations.[1] The criminal law literature has similarly recognized that informal and extralegal enforcement mechanisms, in the form of security devices, privately employed security personnel, reputational sanctions and moral norms, play an essential role in inducing widespread compliance with criminal prohibitions.[2] The patent law literature, however, is largely bereft of any close examination of the informal, extralegal and non-patent protections employed by innovators[3] in order to prevent unauthorized third parties from imitating, and improving upon, their intellectual goods.[4] These alternative appropriation mechanisms are widespread

* Associate, Cleary, Gottlieb, Steen & Hamilton, New York; Adjunct Professor, Fordham University School of Law. This article was written while the author taught in the 2002-03 academic year as a visiting assistant professor at the Fordham University School of Law. Helpful comments were provided by participants at a faculty workshop at Fordham University School of Law and separately by Sonia Katyal, Mark Lemley, Yael Lustmann, Robert Merges, Ken Rosen, Steve Thel and Ben Zipursky. The views expressed herein are those of the author and do not necessarily reflect the views of Cleary, Gottlieb, Steen & Hamilton or any of its clients.

1 *See, e.g.*, David Charny, *Nonlegal Sanctions in Commercial Relationships*, 104 HARV. L. REV. 375 (1990); Benjamin Klein & Keith B. Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. POL. ECON. 615 (1981); Stuart Macaulay, *Non-contractual Relations in Business: A Preliminary Study*, 28 AM. SOC. REV. 55 (1963).

2 *See, e.g.*, Harold G. Grasmick & Robert J. Bursik, Jr., *Conscience, Significant Others, and Rational Choice: Extending the Deterrence Model*, 24 LAW & SOC'Y REV. 837 (1990); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997); Steven Shavell, *Individual Precautions to Prevent Theft: Private Versus Socially Optimal Behavior*, 11 INT'L REV. L. & ECON. 123 (1991).

3 As used herein, "innovators" refers to researchers, inventors and firms that develop, manufacture and distribute patentable products. "Innovation" is used in a broad sense to refer to the entire process of invention and product development, testing, manufacture and marketing.

4 For an exception, see Tom G. Palmer, *Intellectual Property: A Non-Posnerian Law and Economics Approach*, 12 HAMLINE L. REV. 261, 287-300 (1989). The copyright literature has recently devoted significant attention to the expanded use by software vendors and online distributors of literary, visual and musical materials of contractual mechanisms, technological

and include contractual agreements, joint ventures, technological opacity, copy-prevention technologies, secrecy practices, and, most importantly, various branding, product bundling, learning-by-doing and other commercial strategies that exploit a first-mover innovator's "lead time" advantage.[5] Industrial economists have compiled a considerable and consistent body of empirical evidence that suggests that these alternative protection mechanisms usually exceed in practical importance the formal protection supplied by patent entitlements.[6] This is especially true in what the innovation literature calls "complex" (also known as "cumulative" or "systems"[7]) technology sectors, in which innovations tend to arise and operate interdependently[8] (and which probably describes today's most technologically fertile industries[9]), and less true in what the innovation literature calls "discrete" technology sectors, in which innovations tend to exist independently. With the exception of a few such "discrete technology" sectors, informal protection mechanisms are widely perceived to be more effective, and consequently more heavily relied upon, than patents in shielding innovations from unauthorized imitation and improvement.[10]

Although the economics literature has devoted some attention to this evidence in its effort to formulate a reliable descriptive account of patenting behavior, the legal literature has generally not done so and few legal or economic commentators have addressed its normative

---

encryption and "digital rights management" devices as an alternative or supplement to copyright protection. *See, e.g.*, Tom W. Bell, *Escape from Copyright: Market Success vs. Statutory Failure in the Protection of Expressive Works*, 69 U. CIN. L. REV. 741 (2001); Glynn S. Lunney, Jr., *The Death of Copyright: Digital Technology, Private Copying, and the Digital Millennium Copyright Act*, 87 VA. L. REV. 813 (2001); Raymond Shih Ray Ku, *The Creative Destruction of Copyright: Napster and the New Economics of Digital Technology*, 69 U. CHI. L. REV. 263 (2002).

5 For a full discussion, see *infra* Part I.A.

6 For a review and discussion of this evidence, see *infra* Part II.B.1.

7 In a cumulative technology sector, a chain of related innovations often consists of a single first-mover that introduces a pioneering innovation and numerous second-movers that develop an advancing string of closely entangled subsequent improvements. In systems technology sectors, innovations are horizontally entangled in an interdependent web so that any individual innovation is made up of several essential components. *See* Robert P. Merges & Richard R. Nelson, *On Limiting or Encouraging Rivalry in Technical Progress: The Effect of Patent Scope Decisions*, 25 J. ECON. BEHAV. & ORG. 1, 7 (1994) [hereinafter Merges & Nelson, *Patent Scope*].

8 *See* WESLEY M. COHEN ET AL., PROTECTING THEIR INTELLECTUAL ASSETS: APPROPRIABILITY CONDITIONS AND WHY U.S. MANUFACTURING FIRMS PATENT (OR NOT) 6 (Nat'l Bureau of Econ. Research, Working Paper No. 7552, 2000); RICHARD NELSON & SIDNEY G. WINTER, AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE 255-62 (1982); Merges & Nelson, *Patent Scope*, *supra* note 7, at 7-10; Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990) [hereinafter Merges & Nelson, *Complex Economics*]. I am simplifying these commentators' line of reasoning since they generally draw a *four*-part distinction between cumulative, discrete, chemical and science-based technologies. *See* Merges & Nelson, *Patent Scope*, *supra* note 7, at 6-9; Merges & Nelson, *Complex Economics*, *supra*, at 884.

9 *See infra* note 99 and accompanying text.

10 *See supra* note 8.

implications for patent theory, policy, and law.[11] This Article seeks to begin remedying this gap. Evidence concerning the relatively limited effectiveness of patents as appropriation mechanisms, and the relatively greater effectiveness of informal protection mechanisms, poses an important challenge to the conventional theory that accounts for patenting behavior as a positive matter and justifies patent awards as a normative matter. Specifically, this evidence challenges the incentive theory on which scholars, judges, lawyers, legislators, lobbyists and international trade negotiators have relied to account for firms' significant investments in obtaining patent protection and to justify granting and strengthening patent entitlements. As state-created monopolies, patents impose significant social costs in the form of administrative and litigation expenses, monopoly pricing, rent-seeking expenditures by innovators seeking patent awards, and, most importantly, the reduced access of subsequent improvers to the existing knowledge base. Accordingly, some compelling justification is required. Some explanation is also required as a positive matter to explain why firms expend non-trivial amounts on obtaining patents and then sometimes expend far greater amounts in defending patents in infringement litigation. It is widely accepted that, without entitlements such as patents, innovators could not expect to recoup their development costs and enjoy a reasonable rate of return, and therefore would invest few resources in intellectual production relative to other activities having superior appropriative characteristics.[12] Patents are thought to cure this underinvestment problem by providing a secure appropriation device that enables innovators to recoup the costs of, earn a reasonable return on, and attract external financing for, commercially and technologically uncertain research projects. If, however, (1)

---

[11] For exceptions, see JAMES BESSEN & ERIC MASKIN, SEQUENTIAL INNOVATION, PATENTS AND IMITATION (MIT Dept. of Econ., Working Paper No. 00-01, 2000), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=206189 (last visited Oct. 17, 2002) [hereinafter BESSEN & MASKIN, SEQUENTIAL INNOVATION], arguing that, in industries characterized by sequential and complementary innovation activities, patents have little value as an incentive mechanism and may inhibit innovation by reducing competitors' imitative efforts, which would increase firms' future net profits; Richard C. Levin, *A New Look at the Patent System*, 76 AMER. ECON. REV. 199, 200-01 (1986) [hereinafter Levin, *New Look*], stating that patents may be useful for purposes other than establishing property rights, such as measuring the performance of research and development ("R&D") employees or gaining strategic advantage in interfirm negotiations or litigation; Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 626 (2002), arguing that the abundance of worthless patents suggests that patents must play a role other than that of an exclusionary device and proposing that patents play a signaling function in enabling patent holders to credibly indicate the quality of their technological assets and levels of R&D investment); Palmer, *supra* note 4, at 300-04, arguing that patents (and copyrights) represent an unnecessary intervention in the market because innovators have extralegal means by which to appropriate innovation returns.

[12] For a typical formulation of this argument, see Merges & Nelson, *Complex Economics*, *supra* note 8, at 843.

innovators often make use of private protection mechanisms to appropriate a significant portion of the proceeds of intellectual goods and (2) patents offer appropriability capabilities far weaker than those offered by these private mechanisms, then the incentives theory is at least presumptively weakened, both as a positive explanation of patenting behavior and as a normative justification for patent protection.

The logic of the conventional view is straightforward: the market fails to supply innovators with adequate appropriation mechanisms and patents fill this gap, thereby correcting the market failure by encouraging private parties to invest resources in high-risk innovation ventures. In this Article, I essentially reverse this logic. I argue that the market *already* supplies incumbent firms with abundant informal and non-patent appropriation mechanisms. Without the aid of patents, these established firms exert appropriation capabilities that sufficiently shield the proceeds earned on costly investments in product research and development. But there is a significant and unwelcome side effect: incumbents' potent appropriation capabilities significantly inflate development, production and distribution costs for—and thereby dissuade entry by—potential competitors. Simply put, the market failure lies not in an undersupply of appropriation mechanisms in the market generally, but rather, in an oversupply of appropriation mechanisms in a particular market segment. Patent protection is principally required in order to alleviate *this* market failure—specifically, by inducing entrants to invest in fundamental innovation projects in an attempt to undermine incumbents' dominant position. This "market-entry" thesis relies on the fact that the most effective private protection mechanisms are more widely available to incumbents than to small-firm entrants and, to a lesser extent, what I call large-firm "outsiders."[13] While entrants certainly have access to some informal appropriation mechanisms (such as secrecy practices and copy-prevention technologies), the most potent mechanisms take several years to develop and exploit lead-time advantage—e.g., reputational goodwill, learning-by-doing, manufacturing efficiencies and distribution relationships—that by definition are not immediately available to entrants. Consequently, established market leaders have a natural competitive advantage relative to later entrants. Without some non-trivial level of patent protection, potential entrants may have limited incentives to invest in the development of either improvements to the incumbent's technology (that is, complementary innovations) or radical innovations that would render obsolete the incumbent's technology (that is, substitute innovations). To the extent that large-firm incumbents

[13] As used herein, "outsiders" refers to large firms that have an established position in one market and are seeking to enter another unrelated market. Outsiders are distinguished from the remaining entrants, which are small firms without any established position in any market.

have extensive access to, and can establish secure market positions by relying on, effective private protection mechanisms that are largely foreclosed to potential entrants, patent protection may be required in order to facilitate entry into technology-dependent markets dominated by one or a handful of firms.

But why should the state facilitate entry into these markets? The provision of patent protection under the market-entry thesis rests on a compelling efficiency rationale. Obviously, a credible entry threat should generate an efficiency benefit by diminishing incumbents' supracompetitive pricing power. But more importantly from a social-welfare perspective, there is significant (albeit, not entirely undisputed) evidence suggesting that (1) small firms contribute disproportionately to technological advance by supplying a large percentage of fundamental innovations, and (2) large, incumbent firms are likely to have relatively weak incentives and/or organizational capabilities to supply such breakthrough innovations.[14] If this is the case, then inducing small and outsider firms to attempt entry or to expand market-share through risky innovation investments may likely increase the aggregate social value generated by the entire set of research and development ("R&D") investments. Whereas the incentives theory of patent protection posits an underproduction of innovative output and is therefore primarily concerned with the *rate* (or *quantity*) of private innovation investment, a market-entry theory of patent protection posits a misallocation of innovative inputs and is therefore primarily concerned with the *direction* (or *quality*) of private innovation investment.[15] Conventional wisdom argues that, without patents, the market would invest few resources in innovation and therefore, patents must increase the total amount of private innovation investment. By contrast, the market-entry thesis implies that, without patents, intellectual goods would consist largely of incremental improvements produced by entrenched incumbents, and therefore, patents are likely to increase the average quality of private innovation investment.

The market-entry thesis, and the related concern with the quality (rather than the quantity) of innovation investment, runs into an empirical obstacle. As a positive matter, this thesis predicts that large-

---

[14] For a discussion of these points and the related empirical literature, see *infra* Part III.B. Note that I state that these innovation-related losses are more important than supracompetitive pricing from a social-welfare perspective on the basis of empirical research showing that the social loss incurred as a result of underinvestment in innovation projects is most likely far greater than the social loss incurred as a result of supracompetitive pricing. *See* Joseph F. Brodley, *The Economic Goals of Antitrust: Efficiency, Consumer Welfare, and Technological Progress*, 62 N.Y.U. L. REV. 1020, 1025-26 (1987) [hereinafter Brodley, *Economic Goals of Antitrust*]; F. M. Scherer, *Antitrust, Efficiency, and Progress*, 62 N.Y.U. L. REV. 998, 1001-02 (1987).

[15] On the distinction between the "rate" and the "direction" of R&D investment, see Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY 609, 611 (Nat'l Bureau of Econ. Research Report, 1962).

firm incumbents would not incur the significant costs of acquiring patents (and certainly not the even more significant costs of defending patents) because they already have access to more effective, non-patent appropriation mechanisms. As a normative matter, this thesis recommends that patents should principally be awarded to small or outsider firms that seek to contest incumbents' market shares through the risky development of radical technological innovations. But the leading applicants for patent protection consist of large, established incumbents.[16] Since the widespread availability of effective, private protection mechanisms apparently eliminates any plausible incentives explanation, I offer three other explanations for why this behavior is occurring and then, as a normative matter, explore the efficiency implications of each explanation. First, incumbents may wish to erect a barrier against potential entrants by assembling a wide patent portfolio that discourages existing and potential competitors from investing R&D resources in certain technological fields. This explanation suggests that much incumbent patenting is socially inefficient, since an expanded patent portfolio enhances incumbents' first-mover advantage against potential entrants without any offsetting increase in ex ante innovation incentives (given the availability of non-patent appropriation mechanisms). Second, incumbents may be locked into an "arms race" in which a patent portfolio is necessary to compete with, and maintain research access to the industrial knowledge base held by, competitors who are similarly accumulating patents. This explanation is normatively ambiguous: On the one hand, it suggests that established firms may suffer from a collective action problem in which each firm would be better off if all firms could credibly agree not to patent; on the other hand, it suggests that patents may enable competitors to negotiate and enter into knowledge-sharing joint ventures on the basis of legal entitlements covering technological knowledge contributed to the joint venture. Finally, large firms may sometimes act as "small firms" when entering into a new market where they will not have a first-mover advantage and therefore require patent protection in order to contest the incumbent's dominant position. This explanation suggests that some large-firm patenting may be socially efficient to the extent that it sustains credible entry threats to the dominant position of incumbent firms in unrelated markets.

This Article proceeds as follows. In Part I, I describe existing evidence regarding the use of private protection mechanisms in industries producing goods that are generally eligible for patent protection. In Part II, I explore the implications of this body of evidence for the standard incentives theory of patent protection. In Part

[16] *See infra* note 172 and accompanying text.

III, I propose the market-entry thesis as an alternative theory of patent protection. I also address the principal objection to this thesis posed by the prevalence of patenting among large, incumbent firms. In Part IV, I derive some preliminary policy implications from the market-entry thesis and also consider a possible extension of this Article's thesis to copyrightable goods and its more general implications for the analysis of intellectual property law.

I. PRIVATE PROTECTION

Generally speaking, private protection takes four principal forms: (1) first-mover commercial strategies, (2) copy-prevention technologies, (3) contractual devices, and (4) social norms. Relying on evidence from the marketing, business economics and industrial economics literatures, the discussion below describes each of these methods in detail.

A. *First-Mover Strategies*

Private parties may employ a variety of commercial strategies that exploit a pioneer's lead-time or first-mover advantage[17] and raise the costs incurred by third-party competitors in bringing to market an identical or close substitute product. "Lead time" refers to the period from the start of product development until an imitative substitute or close substitute product enters the market. A substantial body of empirical work suggests that a pioneer's first-mover advantages are effective in a variety of industries at raising barriers to entry and preserving a pioneer's long-term market share despite entrants' product improvement and pricing strategies.[18] Just as patent protection is

[17] There is well-developed business economics and marketing literature on the first-mover advantage. For a general introduction to the concept and related empirical findings, see Marvin B. Lieberman & David B. Montgomery, *First-Mover Advantages*, 9 STRATEGIC MGMT. J. 41 (1988). For reviews of this literature, see Roger A. Kerin et al., *First-Mover Advantage: A Synthesis, Conceptual Framework, and Research Propositions*, 56 J. MARKETING 33 (1992); William T. Robinson et al., *First-Mover Advantages from Pioneering New Markets: A Survey of Empirical Evidence*, 9 REV. OF INDUS. ORG. 1 (1994). On the historical importance of the first-mover advantage in the development of modern industrial capitalism, see ALFRED D. CHANDLER, JR., SCALE AND SCOPE: THE DYNAMICS OF INDUSTRIAL CAPITALISM 34-35 (1990).

[18] *See* COHEN ET AL., *supra* note 8 (referring to studies of the aircraft and semiconductor industries showing that "gaining lead time and exploiting learning curve advantages are the primary methods of appropriating returns"); Frank R. Kardes et al., *Brand Retrieval, Consideration Set, Composition, Consumer Choice, and the Pioneering Advantage*, 20 J. CONSUMER RES. 62 (1993) (stating that research has shown that the first brand to enter a new market often gains a long-term market share advantage over later entrants, both in consumer and industrial markets and in growing and saturated markets); Kerin et al., *supra* note 17, at 36 (reviewing empirical studies showing that first-movers tend to have higher market shares than

thought to provide a successful innovator with a period during which it is sheltered from competitors' entry, the first-mover advantage plays a similar function in ensuring such a monopoly (or quasi-monopoly) period. Of course, a first mover does not always have an insuperable advantage. It is not unusual for a capable later entrant to overcome a pioneer's first-mover advantages through product differentiation, improved technology or superior management skills[19] or by "free riding" on the pioneer's heavy initial marketing, manufacturing and

later entrants); Mary Lambkin, *Order of Entry and Performance in New Markets*, 9 STRATEGIC MGMT. J. 127 (1988) (stating that it is widely believed by academics and corporate management that early entrants into newly developing markets enjoy an "enduring competitive advantage" over later entrants and that research backs up this belief); Gary L. Lilien & Eunsang Yoon, *The Timing of Competitive Market Entry: An Exploratory Study of New Industrial Products*, 36 MGMT. SCI. 568, 569 (1990) (describing research showing that pioneering entrants generally maintain their market share advantage and that pioneer entry is one of the major determinants of the long-term success of a new product). For examples of studies empirically assessing the first-mover advantage, see Lambkin, *supra*, at 132, using sample of 129 start-up firms and 187 "adolescent firms" and finding that, on average, twenty-four percent of market leaders among start-up firms were pioneers and thirty-three percent of market leaders among adolescent firms were pioneers versus ten percent and nineteen percent for "early followers" respectively and ten percent and twenty-five percent for "late entrants" respectively; Lieberman & Montgomery, *supra* note 17, at 47, noting research showing that twenty of twenty-five leading consumer brands in 1923 were still in first place approximately sixty years later, and two-thirds of the pioneers in eighteen U.K. grocery product categories since 1945 retained market leadership positions through the mid-1970s; William T. Robinson & Claes Fornell, *Sources of Market Pioneer Advantages in Consumer Goods Industries*, 22 J. MKTG. RES. 305, 310 (1985), using a sample of 371 mature consumer goods businesses and finding that "market pioneers" had an average market share of twenty-nine percent, "early followers" seventeen percent and "late entrants" twelve percent and that more than eighty percent of the pioneers had been in the market for twenty years or more; Richard W. Shaw & Susan A. Shaw, *Late Entry, Market Shares and Competitive Survival: The Case of Synthetic Fibers*, 6 MANAGERIAL & DECISION ECON. 2 (1984), finding that early entrants who established major market shares early in the Western European synthetic fibers industry maintained their leadership positions nearly twenty years later and that almost all late entrants failed to achieve significant market shares and often withdrew from the market altogether; Glen L. Urban et al., *Market Share Rewards to Pioneering Brands: An Empirical Analysis and Strategic Implications*, 32 MGMT. SCI. 645, 647, 654 (1986), interviewing 300 respondents with respect to large group of frequently purchased consumer brands and finding that pioneer's share falls to 58.5% after the first entrant enters, 43.6% after the second entrant, and 35.7% after the third entrant. For minority objections to the above body of evidence, see Peter N. Golder & Gerald J. Tellis, *Pioneer Advantage: Marketing Logic or Marketing Legend?*, 30 J. MARKETING RES. 158 (1993), arguing that evidence showing a first-mover advantage may suffer from (1) sampling bias, since two of the main databases used for such research only include survivors (thus possibly overstating the first-mover advantage); and (2) self-perception bias, because these databases rely on self-classifications of firms as pioneers or late entrants rather than using an objective measure; Robinson & Fornell, *supra*, at 309, raising the possibility that empirical studies may overstate the first-mover advantage by using a database consisting only of successful firms.

[19] *See* Golder & Tellis, *supra* note 18 (stating that pioneers may have first-mover advantage if they have advantages in terms of costs and product-line breadth, but late entrants can sometimes "leapfrog" pioneers through superior technology or branding); Kardes et al., *supra* note 18 (noting that pioneering brands are not always successful); Kerin et al., *supra* note 17, at 47-48 (emphasizing that a first-mover does not have an unbeatable position and that the relationship between market share and market-entry is complex, sometimes working to the advantage of the first-mover and sometimes to the advantage of later entrants).

R&D costs.[20] Even taking into account these caveats, however, the literature tends to agree that first-movers often enjoy a strong competitive advantage relative to all later entrants, including entrants that offer competing products with superior price or quality attributes. Various manifestations of the first-mover advantage are discussed below.

1. Branding and Bundling Strategies

First-movers may bundle together various related products and services, which, combined with a strong brand image and firm goodwill, impose switching costs on customers who jump (or even contemplate jumping) to a competing product. Even if a later entrant can promptly match a first-mover's product quality and distribution costs, it may have difficulty replicating in the short-term the first-mover's favorable reputation, broad product line and extensive sales and service support.[21] For example, even if a particular innovation is subject to easy imitation, it may be offered for sale together with a host of complementary services, such as customer support, updates and repair services or a line of complementary accessories and related products, all of which are difficult to replicate quickly. Two forms of formal (but non-patent) protection make this difficult task even more burdensome. Although there is debate as to its effective strength in various jurisdictions, trade secret protection probably discourages at least the most blatant forms of industrial espionage, thereby delaying, for example, replication of a first-mover's production methods.[22] Of even greater importance, even limited enforcement of trademark protection for company logos and

---

[20] *See* Lieberman & Montgomery, *supra* note 17, at 47; Kerin et al., *supra* note 17, at 35-36, 47, Lambkin, *supra* note 18; *see also* Lilien & Yoon, *supra* note 18 (describing market-entry decision as attempt to balance the risks of premature entry and the risks of missed opportunities as the result of entering too late, and noting that the pioneer must bear most of the costs and risks of developing the initial product). Additionally, the increased mobility of trained personnel and enhanced communication technologies may have speeded up the diffusion of industrial and scientific knowledge, thereby reducing the quasi-monopoly period in which first-movers can safeguard their intellectual and human capital. *See* Rajshree Agarwal & Michael Gort, *First-Mover Advantage and the Speed of Competitive Entry, 1887-1986*, 44 J.L. & ECON. 161, 164-68 (2001). Agarwal and Gort back up this claim with empirical data showing a significant decline in the length of the "monopoly interval"—that is, the period during which a first-mover or group of first-movers enjoys control of the market until the entry of imitative competitors. In a sample of approximately forty-five products, the authors find that the monopoly interval declined from thirty-three years at the turn of the twentieth-century to 3.4 years in 1967-86. *See id.* For similar results, see Golder & Tellis, *supra* note 18, at 167 (finding an average "monopoly interval" of nineteen years for products introduced prior to World War II and five years for products introduced after World War II).

[21] *See* Robinson & Fornell, *supra* note 18, at 314; Kerin et al., *supra* note 17, at 38.

[22] *See* Ronald J. Gilson, *The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not To Compete*, 74 N.Y.U. L. REV. 575 (1999).

other branding devices should make it nearly impossible for a third party to replicate perfectly a competitor's brand image and therefore to acquire quickly the competitor's solid reputation and loyal customer base.[23]

The marketing literature argues that this first-mover advantage can prevail even in the face of objectively superior competing products. Depending on the size and frequency of the purchase, consumers may rationally entertain a strong presumption in favor of the incumbent product in order to minimize product search and evaluation costs[24] (and, following purchase, product learning and adoption costs[25]). Unless a competitor can offer either significantly superior and readily ascertainable product quality or materially lower prices, rational consumers are unlikely to exchange (or even consider exchanging) the first-mover's known and tested commodity for the later entrant's unknown and untested (albeit, even objectively identical) commodity.[26] Empirical evidence provides support for the "switching cost" and related "purchaser inertia" thesis. First, switching costs can explain well why first-movers' market shares tend to be larger both in consumer markets for goods retailing at low sales prices, and in industrial goods markets for big-ticket items.[27] In the former case, low purchase prices

---

[23] On a much more global level, this observation raises the possibility that trademark protection may surprisingly turn out to be the most valuable type of formal intellectual-property protection in many industries.

[24] For the principal exposition of this argument, see Richard Schmalensee, *Product Differentiation Advantages of Pioneering Brands*, 72 AMER. ECON. REV. 349 (1982); *see also* Robinson & Fornell, *supra* note 18, at 309 (noting that consumers will rationally favor the incumbent brand with respect to products that have a low purchase price and a low purchase frequency, since it will not be worthwhile to incur search costs); Kerin et al., *supra* note 17, at 45-46 (noting that consumers will tend to stick with the incumbent product where product search and evaluation costs are high and the potential cost of a purchase mistake is large). For a further variant on this line of argument, see Kardes et al., *supra* note 18, at 63 (arguing and presenting experimental evidence that a pioneer's market share advantage may be partially attributable to its prominent brand and initial monopoly position to the extent that (1) consumers' "retrieval set" (that is, the set of products that are accessible from memory) is limited to more familiar product; and (2) consumers will seek to reduce the costs of a complex purchase decision by limiting the "consideration set" (that is, the set of products to be evaluated) to the most well-known products).

[25] *See* Lieberman & Montgomery, *supra* note 17, at 46 (arguing that, in the case of products that take some time and effort to learn how to use, a consumer may be disinclined to switch to a competing product that will require an additional incurrence of learning costs).

[26] *See* Robinson & Fornell, *supra* note 18, at 7; *see also* Kerin et al., *supra* note 17, at 35. Such buyer inertia is not limited to unsophisticated parties. Doctors continue to prescribe more expensive branded drugs over their generic equivalents because they believe (with little or no empirical support) that generic drug producers maintain inferior quality-control standards relative to those of the branded producers. *See* M.A. Hurwitz & R.E. Caves, *Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals*, 31 J.L. & ECON. 299, 305 (1988).

[27] *See* Robinson & Fornell, *supra* note 18, at 10-11; *see also* Schmalansee, *supra* note 24, at 360 (stating that the first-mover advantage attributable to consumers' uncertainty about the quality of later entrants' imitative products suggests that this advantage will be increased in the case of (1) products that are purchased infrequently and have a high risk relative to unit cost; and

do not justify search expenditures; in the latter case, risk-averse consumers do not want to "take a gamble," learning costs may be high, and the related support package may have added importance.[28] Second, the ability of pharmaceutical manufacturers to maintain significant market share at supracompetitive prices even after the expiration of patent protection and the entry of less expensive generic substitutes,[29] as well as the broader finding that many consumers are willing to pay more for national than identical generic brands,[30] strongly suggests that consumers value highly the incumbent's brand name and associated reputation, attributes which cannot be imitated easily by competitors. Finally, evidence shows that consumers are especially likely to favor pioneer products in entirely novel markets, where consumers may be particularly uninformed about the attributes of the new product and will treat the pioneer as highly or uniquely representative of the entire product class.[31] Latecomer competitors, even if offering an objectively superior product, are perceived as mere "copycats" and must struggle to overcome the distinctive position occupied by the pioneering brand as the ideal prototype product.[32]

---

(2) products with low unit cost that are "convenience goods").

[28] *See supra* note 27.

[29] *See* Andrea Coscelli, *The Importance of Doctors' and Patients' Preferences in the Prescription Decision*, 48 J. INDUS. ECON. 349, 363 (2000); Paul K. Gorecki, *The Importance of Being First: The Case of Prescription Drugs in Canada*, 4 INT'L J. OF INDUS. ORG. 371, 371-72 (1986); John Hudson, *Generic Take-Up in the Pharmaceutical Market Following Patent Expiry: a Multi-Country Study*, 20 INT'L REV. L. ECON. 205, 216 (2000). It appears that these results may be attributable, at least in part, to uncertainty about the quality of the generic product and doctors' and pharmacists' resulting fear of liability for prescribing a generic product. A study of market penetration of generic pharmaceuticals into the Canadian provincial markets finds that generics' market share is drastically higher when the provincial government certifies that the generic and the brand drug are interchangeable (which, under the relevant law, then absolves doctors and pharmacists of liability). *See* Gorecki, *supra*, at 376-83. Even where there is interchangeability, however, a significant percentage of doctors continue to prescribe the brand name, indicating customer loyalty and potentially lingering uncertainty regarding product quality and professional liability. *See id.* at 388-89. It should be noted that there is some disagreement as to whether branded drugs and their generic equivalents are identical, principally because generics manufacturers are claimed to invest fewer resources in quality control. Doctors apparently subscribe to this view, perhaps partly due to fears of malpractice liability. *See* Hudson, *supra*, at 207.

[30] *See* ROBERT BARSKY ET AL., WHAT CAN THE PRICE GAP BETWEEN BRANDED AND PRIVATE LABEL PRODUCTS TELL US ABOUT MARKUPS? (Nat'l Bureau of Econ. Research, Working Paper No. 8426, 2001) (showing large price premia for "name brands" over essentially identical generic brands, across a wide range of consumer products and excluding product types where the house brand might be of lower quality).

[31] *See* Gregory S. Carpenter & Kent Nakomoto, *Consumer Preference Formation and Pioneering Advantage*, 26 J. MKTG. RES. 285, 287-88 (1989); Kardes et al., *supra* note 18, at 62; *see also* Lilien & Yoon, *supra* note 18, at 569 (noting research showing that pioneer products have an advantage because they are the first to establish a reputation in the marketplace and a stable customer base).

[32] *See* Carpenter & Nakomoto, *supra* note 31, at 287-88.

2. Tacit Knowledge

A survey study that sought to assess imitation costs found that third parties must incur roughly two-thirds of the costs incurred, and roughly seventy percent of the time taken, by the original innovator in inventing the relevant product.[33] As this finding suggests, replicating a particular innovation is often far from automatic, demands a preexisting skill set among firm personnel, and can require a significant investment of time and money on the part of the imitator firm in order to arrive at the tacit knowledge that lies behind the pioneering innovation.[34] Especially problematic is the fact that much valuable technological know-how is often bound up with the "human capital" of a particular star researcher or team of researchers,[35] which is obviously impossible to imitate

---

[33] *See* Edwin J. Mansfield et al., *Imitation Costs and Patents: An Empirical Study*, 91 ECON. J. 907, 909-10 (1981). Based on interviews conducted with managers of firms in the chemical, drug, electronics and machinery industries concerning the cost and time of legally imitating forty-eight product innovations, Mansfield reports that imitation costs average about sixty-five percent of innovation costs, and imitation time averages about seventy percent of innovation time (although there is considerable interproduct variation). *Id.* at 909. "Imitation costs" are defined to include all costs of developing and introducing the imitative product, up to and including manufacturing and marketing "start-up." *Id.* at 907. About seventy percent of the subject products were patented and thus, with respect to these products, imitation costs include "inventing around" costs. *Id.* at 909. "Imitation time" is defined as the length of time from the beginning of the imitator's applied research (if any) until the date of the imitative product's commercial introduction. *Id.* at 907.

[34] *See* James Bessen, *Adoption Costs and the Rate of Return to R&D* (arguing that adoption costs are significant and may imply that technology is often not a non-rival good that can be accessed at little or no cost and further arguing that high adoption costs may enhance appropriability), *available at* http://www.researchoninnovation.org (last visited Oct. 24, 2002); *see also* Richard C. Levin et al., *Appropriating the Returns from Industrial Research and Development*, *in* 3 BROOKING PAPERS ON ECONOMIC ACTIVITY 783 (Martin Neil Baily & Clifford Winston eds., 1987) (noting that industry managers report that typically only a few other firms are capable of duplicating their innovations and that learning curves provide an obstacle to imitation); Raphael Kaplinsky, *Firm Size and Technical Change in a Dynamic Context*, 32 J. INDUS. ECON. 39, 46-48 (1983) (stating that, without patent protection, software producers are able to appropriate proceeds by virtue of the fact that (1) it takes considerable time and investment to imitate a software product; and (2) there is a shortage of qualified software writers); Levin, *New Look*, *supra* note 11, at 199-200 (citing the aircraft and semiconductor industries as examples of industries in which, as a result of a steep learning curve and high cost of reverse engineering, powerful incentives to innovate exist despite the absence of strong patent protection); Richard Reed & Robert J. DeFillippi, *Causal Ambiguity, Barriers to Imitation and Sustainable Competitive Advantage*, 15 ACAD. MGMT. REV. 88, 90-94 (1990) (arguing that the existence of significant tacit knowledge, complexity of organizational routines and multiple product components, and various types of physical and human asset specificity may frustrate competitors' attempts to imitate the products and services offered by an incumbent); Richard Zeckhauser, *The Challenge of Contracting for Technological Information*, 93 PROC. NAT'L ACAD. SCI. 12743 (1996) (stating that information goods often do not fully satisfy the public good criterion of nonexcludability because, even where there are no access restrictions, the information or technology may be costly to master and adopt).

[35] *See* Bharat N. Anand & Alexander Galetovic, *Weak Property Rights and Holdup in R&D*, 9 J. ECON. & MGMT. STRATEGY 615, 616 (2000). Some empirical work has been done to assess

perfectly or even closely. Even if a later entrant can perfectly replicate a first-mover's innovation, a first-mover's lead time still gives it the opportunity to develop production and distribution efficiencies which may take third-party imitators considerable time to replicate.[36] On the production side, a first-mover may, as a result of either "learning-by-doing" and/or economies of scale, enjoy cost efficiencies that enable it to offer prices below imitators and preserve its market share.[37] On the distribution side, a first-mover may cultivate arrangements with resellers and other retail agents that may include specially tailored and unusually favorable contractual provisions, not to mention the more amorphous "human capital" that accrues within the context of a long-term business relationship.[38] Exemplifying this point, imitation barriers in the form of learning-by-doing may explain why investment banks have historically been willing to invest in the development of novel

---

the impact of a "star" researcher or research team to the economic growth of the relevant firm. *See* ASHISH ARORA ET AL., MARKETS FOR TECHNOLOGY: THE ECONOMICS OF INNOVATION AND CORPORATE STRATEGY 28-29 (2001) [hereinafter ARORA ET AL., MARKETS FOR TECHNOLOGY] (describing empirical evidence showing that firms with affiliated scientists showed significantly higher average growth in employment relative to firms without such ties but noting that causality is ambiguous since higher-performing firms may attract higher-performing scientists).

[36] On this point generally, see David B. Audretsch & Zoltan J. Acs, *Technological Regimes, Learning and Industry Turbulence*, *in* ENTREPRENEURSHIP, TECHNOLOGICAL INNOVATION AND ECONOMIC GROWTH 314-17 (F. M. Scherer ed., 1992), discussing the view that entry will be difficult for new and fringe firms in industries where knowledge is largely accumulated through experience in the industry and is therefore not easily accessible to outsiders. *See also* Edwin J. Mansfield, *How Rapidly Does New Industrial Technology Leak Out?*, 34 J. INDUS. ECON. 217, 221 (1985) (stating that "[i]t often takes considerable time to invent around patents (if they exist), to develop prototypes, to alter or build plant and equipment, and to engage in the manufacturing and marketing start-up activities required to introduce an imitative product or process").

[37] *See* Kerin et al., *supra* note 17, at 41; Robinson & Fornell, *supra* note 18, at 307; *see also* Agarwal & Gort, *supra* note 20, at 162 (stating that some scholars have argued that scale economies can help first-movers to maintain market share against later entrants that cannot replicate such scale economies until their competing products have gained sufficient acceptance among consumers); Lilien & Yoon, *supra* note 18, at 569 (describing research showing that production costs for the pioneer tend to be lower than those for later entrants as a result of the "experience curve effect"); Lieberman & Montgomery, *supra* note 17, at 42-43 (stating that tacit knowledge of pioneer entrant, in addition to cost efficiencies as the result of high output, can erect substantial barrier to entry by imitators).

[38] *See* Robinson & Fornell, *supra* note 18, at 307; *see also* Robinson et al., *supra* note 17, at 8 (noting study showing that, in a survey of 145 resellers of supermarket goods, resellers were more favorably inclined toward the first-mover, market-leader brands as having greater potential to contribute to sales); Kerin et al., *supra* note 17, at 38 (stating that empirical studies in the semiconductor industry and related research show that links between vendors and buyers and certain buying practices increase switching costs on vendors and thereby raise barriers to later entrants). Jay Barney offers an excellent illustration of this point. During World War II, the U.S. military selected Caterpillar to provide it with heavy equipment worldwide and assisted the company in developing a global distribution and support network. Because no other U.S. heavy-equipment manufacturer put into place such an extensive distribution network, Caterpillar enjoyed after the war a competitive advantage that was difficult for any competitor to overcome. *See* Jay B. Barney, *Competence Explanations of Economic Profits in Strategic Management: Some Policy Implications*, *in* DYNAMIC COMPETITION AND PUBLIC POLICY 48-49 (Jerry Ellig ed., 2001).

securities products even without patent protection, which only became available in the late 1990s.[39] To the extent that being the first to market with a novel security instrument enables an investment bank to develop cost efficiencies, tacit knowledge or marketing capabilities that in turn allow it to offer lower prices or superior performance relative to its imitative competitors, the innovating bank can capture significant returns and minimize knowledge spillovers to its competitors.[40] Of course, many such learning-by-doing advantages necessarily deteriorate over time as internal firm knowledge leaks out through employee movement, research publications, reverse engineering, trade shows and other avenues.[41]

3. Inter-firm Alliances

Competing innovators may enter into strategic alliances, joint ventures and other cooperative relationships with potential or actual competitors in order to share knowledge and pool R&D, production and distribution functions.[42] These ventures can increase the participants' abilities to appropriate innovation returns by more speedily bringing a completed product to market and minimize knowledge "giveaways" to competitors by more quickly developing a broad range of related

---

[39] Prior to the Federal Circuit's *State Street* decision in 1998, the patent office occasionally granted patents on financial products but many market participants questioned their validity given judicial case law holding that business methods were unpatentable. The *State Street* decision and subsequent legislation established the patentability of financial and other business methods and the patent office has issued almost 500 patents (primarily to financial institutions) for financial product innovations. *See* Josh Lerner, *Where Does* State Street *Lead? A First Look at Finance Patents, 1971-2000*, 57 J. FIN. 901-04 (2002). However, in response to public alarm over the growing number of business-method patents, the PTO instituted a second level of review for business-method applications, and the number of business-method patents issued fell from approximately 900 in 2000 to approximately 430 in 2001. *See* Evan P. Schultz, *Too Many Patents?*, LEGAL TIMES, Mar. 21, 2002.

[40] *See* Peter Tufano, *Financial Innovation and First-Mover Advantages*, 25 J. FIN. ECON. 213, 231-35 (1989). Tufano explains that an innovating bank most likely enjoys cost efficiencies because innovators can exploit their knowledge of the buyers of the new security product to become the dominant market-maker of the security. Tufano also shows that investment banks do earn superior returns as a result of financial innovation—specifically, pioneers in the study sample capture market shares for the relevant security product almost two and one-half times as large as their imitators. *See id.* Note that the same logic should explain law firms' willingness to develop novel contractual forms even when competing firms will be able to imitate the forms at a much lower cost relative to the cost incurred by "first-mover" law firms in developing such forms.

[41] *See* Lieberman & Montgomery, *supra* note 17, at 43; Reed & DeFillippi, *supra* note 34, at 96-97.

[42] For general explorations of strategic alliances from the management literature, see JOSEPH L. BADARACCO, JR., THE KNOWLEDGE LINK: HOW FIRMS COMPETE THROUGH STRATEGIC ALLIANCES (1991); PETER LORANGE & JOHAN ROOS, STRATEGIC ALLIANCES: FORMATION, IMPLEMENTATION, AND EVOLUTION (1992).

products.[43] Moreover, by pooling market uncertainty,[44] technical uncertainty[45] and innovation costs, these cooperative ventures bring down total product development, marketing and manufacture costs for the alliance members and thereby raise entry costs for all other competitors or potential competitors not party to the alliance (or a similar cooperative venture).[46] Knowledge-sharing and competency-sharing alliances are widely employed with great success in the biotechnology industry where small biotechnology firms, university research departments and large pharmaceutical firms often collaborate.[47]

4. Network Effects

In certain industries, first-movers can exploit the lock-in attributes of "network technologies"—that is, technologies where each consumer who has already purchased the relevant product experiences increasing returns as more consumers adopt it. For example, a consumer enjoys greater benefits from the purchase of a telephone as more and more consumers purchase telephones. If a particular brand of telephones can only "understand" telephones of the same brand, then the first brand on

---

[43] *See* Jonathan M. Barnett, *Cultivating the Genetic Commons: Imperfect Patent Protection and the Network Model of Innovation*, 37 SAN DIEGO L. REV. 987, 1017-21 (2000); Janusz Ordover, *A Patent System for Both Diffusion and Exclusion*, 5 J. ECON. PERSP. 43, 54-55 (1991) (noting that theoretical economic research has shown that when knowledge spillovers are moderately high and property rights weak, research joint ventures can be an effective means for internalizing such spillovers).

[44] Some innovation theorists distinguish between two forms of uncertainty. "Market uncertainty" refers to the fact that innovators may be uncertain as to whether there will be sufficient demand for the planned product and whether the input costs of production will lie at cost-effective levels. "Technical uncertainty" refers to the fact that innovators may be uncertain as to whether the project will result in a technically feasible application. *See* CHRIS FREEMAN & LUC SOETE, THE ECONOMICS OF INDUSTRIAL INNOVATION 242-45 (3d ed. 1997).

[45] *See id.*

[46] *See* Barnett, *supra* note 43; Gary Pisano, *The R&D Boundaries of the Firm: An Empirical Analysis*, 35 ADMIN. SCI. Q. 153, 155 (1990); *see also* Wesley M. Cohen & Steven Klepper, *A Reprise of Size and R&D*, 106 ECON. J. 925, 948-49 (1996) [hereinafter Cohen & Klepper, *A Reprise of Size and R&D*] (noting that smaller firms can overcome the superior ability of large firms to spread R&D costs by entering into cooperative ventures to conduct R&D, thereby spreading the costs among several participants).

[47] On this topic, see JOSH LERNER & ROBERT P. MERGES, THE CONTROL OF STRATEGIC ALLIANCES: AN EMPIRICAL ANALYSIS OF BIOTECHNOLOGY COLLABORATION (Nat'l Bureau of Econ. Research, Working Paper No. 6014, 1997); JULIA PORTER LIEBESKIND ET AL., SOCIAL NETWORKS, LEARNING AND FLEXIBILITY: SOURCING SCIENTIFIC KNOWLEDGE IN NEW BIOTECHNOLOGY FIRMS (Nat'l Bureau of Econ. Research, Working Paper No. 5320, 1995); Barnett, *supra* note 43, at 1015-18; Rebecca S. Eisenberg, *Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research*, 82 VA. L. REV. 1663 (1996); Walter W. Powell, *Inter-Organizational Collaboration in the Biotechnology Industry*, 152 J. INSTITUTIONAL & THEORETICAL ECON. 197 (1996); Kenneth Sutherlin Dueker, *Biobusiness on Campus: Commercialization of University-Developed Biomedical Technologies*, 52 FOOD & DRUG L.J. 453 (1997).

the market may enjoy an unbeatable "lock-in" advantage.[48] In a network technology market, the first-mover advantage is especially strong since even a superior but latecomer technology may not induce customers to incur the switching costs of abandoning the old but inferior technology that is far more widely used than its proposed substitute.[49] Moreover, even where a first-mover does not ultimately offer the sole product embodying the "locked-in" technology, during the lead-time interval it can at least set the industry standard in a manner tilted in favor of its technological and organizational competencies.[50]

B. *Copy-Prevention Technologies*

Innovators invest resources in obtaining and developing technologies to prevent unauthorized re-engineering, copying, production, distribution and other third-party appropriations of their intellectual goods. These technologies obviously take different forms in each product market but the foremost copy-prevention technologies are the secrecy and nondisclosure practices used in almost every industry. Any manufacturer can institute procedures that limit physical access to its research laboratories and data network (or, in the case of a software producer, the source code),[51] non-disclosure practices that limit its researchers' rights to publish scientific findings, "non-compete" agreements that limit its employees' abilities to work for a competitor[52] or incentive mechanisms that discourage employee attrition.[53]

---

[48] *See* W. Brian Arthur, *Competing Technologies, Increasing Returns, and Lock-In by Historical Events*, 99 ECON. J. 1116, 1116-17 (1989). The commercial victory of the VHS system over the Betamax system in the home videocassette recorder market is often thought to be an example of a lock-in mechanism. For an important critique of the theory that even a latecomer technology, although superior, may not be able to overcome a first-mover's "network effect" advantage (and a related rejection of the empirical argument that Betamax offered technology superior to that of VHS), see S. J. Liebowitz & Stephen E. Margolis, *Path Dependence, Lock-In, and History*, 11 J.L. ECON. & ORG. 205, 208 (1995).

[49] *See* Arthur, *supra* note 48; Agarwal & Gort, *supra* note 20, at 165.

[50] *See* Lieberman & Montgomery, *supra* note 17, at 47. For similar thoughts, see Richard N. Langlois, *Technological Standards, Innovation and Essential Facilities: Toward a Schumpeterian Post-Chicago Approach*, *in* DYNAMIC COMPETITION AND PUBLIC POLICY, *supra* note 38; *see also Case History: Trapeze Artists*, *in* THE ECONOMIST, TECHNOLOGY Q., Dec. 14, 2002, at 23 (noting that the makers of Adobe Acrobat software have given away free programming tools in order to roster a group of outside developers who then have a vested interest in making sure that the Adobe PDF standard remains the dominant format).

[51] *See* Kaplinsky, *supra* note 34, at 49-50 (stating that software producers seek to appropriate innovation proceeds by limiting physical access to the source code).

[52] For a discussion on non-compete agreements and the varying enforceability of these agreements depending on the relevant jurisdiction, see Gilson, *supra* note 22.

[53] *See* Kaplinsky, *supra* note 34, at 50-51 (stating that software companies sometimes attempt to discourage employee attrition, either negatively, by forcing an employee to resell his stock to the company if he leaves employment, or positively, by offering stock option and bonus plans).

Producers of software[54] may install "lock out" technologies that prevent all but the most sophisticated users from easily making unauthorized copies[55] or a disabling feature that deactivates the program automatically after a specific time period unless renewed for a fee pursuant to the governing license.[56]

C. *Contract*

Innovators may employ a variety of contractual devices in order to obtain third parties' commitment to refrain from taking actions that would impede control over the proceeds flowing from the relevant product. These contracts sometimes confer rights over unpatented or unpatentable subject matter, or alternatively, cover a product that is already patented and extend or tailor the rights already held by a patent owner with respect to a particular licensee. Software vendors often attach "shrink-wrap licenses" to their products, which restrict the consumer's scope of use of the product and often effectively extend the vendor's rights beyond those it already enjoys under the patent or copyright laws.[57] Technology licensing agreements typically include provisions placing geographic restrictions on where the licensor may sell the licensed product or "grant-back" clauses that require the licensee to share with the licensor any improvements it makes on the licensed technology.[58] Similarly, inter-firm alliances in the biotechnology industry are often founded on licensing agreements that contain restrictions on the licensee's use of certain patented technologies contributed to the alliance, such as "grant-back" clauses as described above or "field of use" restrictions limiting the licensee's

---

[54] Private protection mechanisms with respect to software are included in this discussion because software applications are sometimes eligible for both patent and copyright protection. For a detailed discussion, see ROBERT P. MERGES ET AL., INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE 911-1051 (2d ed. 2000).

[55] *See* Lunney, *supra* note 4, at 826-28. Note that the ability to install encryption technologies may be limited by the fact that as encryption becomes more sophisticated, it tends to interfere with the functioning of the encrypted application. *See id.*

[56] *See* Note, Timothy P. Heaton, *Electronic Self-Help Software Repossession: A Proposal to Protect Small Software Development Companies*, 6 B.U. J. SCI. & TECH. L. 8 (2000). In yet another variation, software companies are developing a distribution model in which software is "hosted" on the service provider's server as opposed to the user's desktop, which gives the software company far greater control over the user's access to its product. This is known as the "application service provider model" or "ASP". *See* SOFTWARE & INFORMATION INDUSTRY ASSOCIATION, SIIA'S REPORT ON GLOBAL SOFTWARE PIRACY 2000 6.

[57] *See* Mark A. Lemley, *Intellectual Property and Shrinkwrap Licenses*, 68 S. CAL. L. REV. 1239, 1246-48 (1995).

[58] *See* Richard E. Caves et al., *The Imperfect Market for Technology Licenses*, 45 OXFORD BULL. OF ECON. & STAT. 249, 260 (1983).

right to apply the licensed technology to certain biomedical "targets."[59] While such contract-based alliances often rely on technological assets that are (at least nominally) covered by patents[60] (which perhaps makes many joint ventures a "dirty" form of non-patent protection),[61] participants in these collaborative product development arrangements are clearly employing non-patent legal devices (namely, contractual instruments) to shield innovation proceeds that would otherwise be more easily appropriable by potential co-venturers.

D. *Industry Norms*

To the extent that many innovating firms are repeat players intent on participating for many years in the relevant industry, they may be subject to sharing or courtesy norms. These norms encourage market participants to accumulate long-term reputational capital by foregoing the short-term gains from unauthorized appropriation of a competitor's intellectual capital. Thus, researchers and firms in certain industries may follow norms of knowledge exchange that impose reputational penalties on researchers and firms that appropriate ideas without authorization.[62] Even where these reputational penalties do not prevent infringement and the filing of a defensive suit, repeated interaction among market participants may promote relatively speedy settlement of patent disputes without resort to the expensive discovery and trial process.[63] Exemplifying the possible influence of such industry norms,

---

[59] *See* Gregory B. Abbott, *Pharmaceutical and Biotechnology Licensing and Joint Ventures*, 514 PLI/PAT. 37, 54-57.

[60] *See* LERNER & MERGES, *supra* note 47.

[61] As I explain later in this Article, even where contractual devices are employed as an appropriation strategy in the joint venture context, some participants may be induced to exchange intimate technological knowledge with the other joint venture partners partially due to the presence of reasonably secure levels of patent protection, see *infra* note 177 and accompanying text. On this point generally, see Robert P. Merges, *Intellectual Property and the Costs of Commercial Exchange*, 93 MICH. L. REV. 1570, 1573-74 (1995) (arguing that intellectual property rights may render contract-based joint ventures more attractive than completely integrated organizational structures insofar as such rights lower the risk of sharing knowledge in such transactions and allow parties to tailor terms governing the use of contributed intellectual assets).

[62] *See* LIEBESKIND ET AL., *supra* note 47, at 11 (stating that social networks, and especially academic norms regarding knowledge exchange, within the biotechnology community may provide protection against the unauthorized appropriation of unprotected knowledge). For a study of social norms in the scientific research community and how such norms have changed since the onset of industrial sponsorship of biomedical research and the increased patenting of research results, see Rebecca S. Eisenberg, *Proprietary Rights and the Norms of Science in Biotechnology Research*, 97 YALE L.J. 177 (1987).

[63] *See* JEAN O. LANJOUW & MARK SCHANKERMAN, ENFORCING INTELLECTUAL PROPERTY RIGHTS 4-5, 21-22 (Nat'l Bureau of Econ. Research, Working Paper No. 8656, 2001). The authors note that most patent suits settle soon after the suit is filed and argue that this may be

software patent litigation has been surprisingly sparse since the introduction of patent protection for software in the early 1980s,[64] which may reflect many market participants' desire to maintain industry norms of cooperative behavior rather than to aggressively exert patent rights.[65]

II. THE INCENTIVES THESIS

In this Part, I review the standard incentives argument for patent protection and explore how this argument is challenged by empirical evidence concerning the widespread availability of private protection devices and the limited effectiveness of patent protection.

A. *The Incentives Thesis*

As a legally created monopoly over an intellectual good, a patent entitlement imposes significant social costs. These costs include: (1) administrative costs (incurred by the patent office[66] and innovators) of prosecuting patents,[67] issuing patents and adjudicating disputes relating to patent infringement, (2) rent-seeking costs incurred by innovators seeking to win a patent, (3) supracompetitive pricing power exerted by the patent holder (or, more specifically, the deadweight loss resulting from the patent holder's output restrictions), and (4) restricted access to the patented good by subsequent improvers. As will be discussed further, some of these costs—in particular, supracompetitive pricing power—may decline to the extent that patent rights are insecure or narrowly defined and therefore, do not preclude competitors from developing and offering close substitute products.[68] Nonetheless, there

attributable to repeated interaction among competitors and the resulting incentives to act cooperatively. The authors find some quantitative evidence to back up their hypothesis, finding that where a technology sector is more concentrated, there is less litigation, suggesting that companies in concentrated industries are less likely to allow patent disputes with their competitors to reach litigation. *See id.*

64 *See infra* note 90 and accompanying text.

65 *See* BESSEN & MASKIN, SEQUENTIAL INNOVATION, *supra* note 11, at 17.

66 I am, of course, aware that the "patent office" in the U.S. is known as the Patent & Trademark Office, or PTO. Throughout the Article, I use the generic term, "patent office", for purposes of convenience.

67 "Patent prosecution" is the term of art that refers to the process of applying for a patent from the patent office.

68 This possibility has been demonstrated empirically in the pharmaceutical sector. *See* FRANK R. LICHTENBERG & TOMAS J. PHILIPSON, THE DUAL EFFECTS OF INTELLECTUAL PROPERTY REGULATIONS: WITHIN- AND BETWEEN-PATENT COMPETITION IN THE US PHARMACEUTICAL INDUSTRY (Nat'l Bureau of Econ. Research, Working Paper No. 9303, 2002). The authors find that pharmaceutical firms selling a patented therapeutic product often face

is strong reason to believe that all or most of these costs reach significant levels. Patent prosecution and litigation costs alone are estimated to equal roughly $4.33 billion and $1 to $2.1 billion annually, respectively.[69]

Some compelling justification is therefore required to explain why these enormous costs are socially worthwhile. This is where the incentives argument comes in. Assume that the costs of imitating intellectual goods are generally far lower than the costs required initially to develop such goods and to bring them to market. Further assume that an innovator cannot prevent its competitors from studying and copying the innovation once it enters the market (or once it is revealed in other settings, such as in the course of business negotiations). Given these assumptions, no innovator can expect to recover its development costs plus a reasonable return. As a result, a socially undesirable "waiting game" develops where no innovator is willing to act as the first-mover who sinks funds into an innovation project. Here is where the state must step in to correct market underinvestment. Once the state supplies a legally enforceable right to control exclusively the production and distribution of certain intellectual goods, potential innovators can expect to be able to prevent a good deal of unauthorized imitation and improvement and therefore recoup at least development costs plus a reasonable return. As a result, patents restore adequate incentives for innovation investment and solve the waiting game problem.

---

competition from other patented therapeutic products that offer alternative methods of treating the same disease or disorder and are therefore fall outside the patent scope. According to the authors, such "between-patent" competition reduces a pharmaceutical firm's returns to a greater extent than the entry of generic substitutes ("within-patent" competition) following patent expiration. *See id.*

69 About 1,600 patent lawsuits are filed each year and about 100 patent cases per year make it to trial. The median cost of patent litigation to each side is $799,000 through the end of discovery and $1,503,000 through trial and appeal. Taking a rough conservative estimate, Mark Lemley arrives at an estimate of $2.1 billion spent on patent litigation each year, not including amounts paid in judgment. *See* Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495 (2001) [hereinafter Lemley, *Rational Ignorance*]. Josh Lerner estimates that patent litigation begun in 1991 will lead to total legal expenditures of about $1 billion. *See* Josh Lerner, *Patenting in the Shadow of Competitors*, 38 J.L. & ECON. 463, 470 (1995) [hereinafter Lerner, *Patenting in the Shadow of Competitors*]. Neither estimate includes the social costs of spent judicial resources. Estimates for patent prosecution are even higher. In a conservative estimate, Lemley gives $4.33 billion as the total annual cost of U.S. patent prosecution. *See* Lemley, *Rational Ignorance*, *supra*. Elsewhere, Lemley and co-author John Allison note an estimate of over $5 billion for annual patent prosecution costs. *See* John R. Allison & Mark A. Lemley, *Who's Patenting What? An Empirical Exploration of Patent Prosecution*, 53 VAND. L. REV. 2099, 2100 (2000).

B. *The Limited Relevance of Patent Protection*

While cogent theoretically, the incentives theory runs into some problems when applied in the real world. Note that the incentives theory relies on three crucial assumptions. The first two assumptions are: (1) high invention costs incurred by the first-mover, relative to the low imitation costs incurred by third-party latecomers; and (2) high exclusion costs incurred by the first-mover in preventing unauthorized third-party appropriation. Where these cost assumptions hold, no potential innovator will want to suffer the first-mover penalty and a waiting game problem should logically develop. The third assumption is: patents will cure the waiting game by drastically inflating third-parties' imitation costs and drastically lowering first-movers' exclusion costs. Where this assumption holds, potential innovators will rationally sink resources into uncertain R&D ventures and patents will indeed have cured the waiting game problem. In this Part, I explore a growing body of empirical evidence, which, together with the evidence discussed above concerning private protection methods, contests each of these assumptions. This evidence includes: (1) survey studies showing industry participants' low estimation of the relative effectiveness of patent protection, (2) studies showing no or limited correlation between R&D investment and modifications to the strength of available patent protection, (3) the formidable definitional barriers facing patent applicants in various technology settings, and (4) the low frequency with which patents are renewed.

As discussed in the preceding Part and discussed further below, available evidence shows that in many industries innovators have effective means other than patents of raising third parties' imitation costs and shielding a good portion of the innovation proceeds from third-party imitators. This suggests that assumption (1) (low imitation costs) sometimes may not be true, and assumption (2) (high exclusion costs) often may not be true. If first-mover innovators can erect significant non-patent barriers to unauthorized third-party appropriation, then, even in a world without patents, (1) competitors sometimes *do* incur significant costs in seeking to imitate and/or improve upon a first-mover innovation, and (2) innovators often *do not* face exorbitant exclusion costs in seeking to prevent unauthorized imitation. As a result of the powerful first-mover advantage and other informal appropriation devices, an innovator may enjoy a strong market share even in the most extreme case where a later entrant can either replicate the first-mover product exactly or offer a competing product with superior price or quality attributes. Thus, there often may not be any or

much of a waiting game problem for patents to cure. But even where such a waiting game problem does exist, empirical evidence confirms several commentators' intuition that patents may not offer much of a cure[70]—meaning, patents may have little effect in either increasing imitation costs incurred by third parties or reducing exclusion costs incurred by first-movers. That is, the third assumption may also sometimes not be true. As discussed further below, with the exception of a few industries, it appears that patents have remarkably limited effectiveness in significantly enhancing a patent holder's ability to capture innovation proceeds.

1. Survey Evidence

There are several well-known studies that examine the perceptions of industry participants regarding the effectiveness of patents relative to other appropriation mechanisms. These studies, published in 2000 (the "Cohen et al. study"),[71] 1987 (the "Yale study"),[72] 1986 (the "Mansfield et al. study")[73] and 1973 (the "U.K. Study"),[74] report impressively

---

[70] *See* Long, *supra* note 11 (stating that it is "undeniable" that "worthless patents abound" when the value of intellectual property rights is measured solely in terms of exclusion and rents); John Baldwin et al., *Determinants of Innovative Activity in Canadian Manufacturing Firms: The Role of Intellectual Property Rights*, at 8 (Mar. 7, 2000) (stating that "[d]espite the widespread belief that the existence of intellectual property protection is critical to the innovative process, empirical evidence as to the beneficial effects on innovative activity is sparse" and that "there is empirical evidence to suggest the opposite"), *available at* http://papers.ssrn.com/sol3/Delivery.cfm/000524351.pdf?abstractid=229792 (last visited Mar. 19, 2004) ; Edmund W. Kitch, *Property Rights in Inventions, Writings, and Marks*, 13 HARV. J L. & PUB. POL'Y 119, 122-23 (1990) (stating that "[t]he claims of most issued patents are so narrow that competitors can devise many ways of achieving the same thing as the subject matter of the claim", and that "most issued patents are worthless, or very nearly worthless . . . [with] no market value, much less market power"); Lieberman & Montgomery, *supra* note 17, at 43 (stating that in most industries patents are not meaningful, because either they offer weak protection, are easy to "invent around" or quickly become obsolete due to technological change); Urban et al., *supra* note 18, at 645 (stating that patents are often ineffective because of difficulties of protecting patent rights and the ability of other firms to "invent around" the patent as the relevant technology advances).

[71] *See* COHEN ET AL., *supra* note 8. In 1994, sample data was collected by surveying R&D lab managers randomly drawn from a sample of all R&D labs located in the U.S., which conducted R&D as part of a manufacturing firm; responses were received from 1,478 labs, representing a response rate of fifty-four percent.

[72] *See* Levin et al., *supra* note 34. Sample includes all publicly traded firms in the U.S. with R&D expenses in excess of either one percent of sales or thirty-five million dollars as of 1981. This sample includes firms in more than one hundred manufacturing industries and 130 lines of business.

[73] *See* Edwin J. Mansfield, *Patents and Innovation: An Empirical Study*, 32 MGMT. SCI. 173 (1986) [hereinafter Mansfield, *Patents and Innovation*]. The sample includes one hundred randomly selected firms from twelve industries in the U.S., excluding firms with less than twenty-five million dollars in sales with respect to inventions or products commercially introduced during 1981-83.

[74] *See* C. T. TAYLOR & Z. A. SILBERSTON, THE ECONOMIC IMPACT OF THE PATENT SYSTEM:

consistent results. All these studies use large sample sets consisting of medium to large manufacturing firms in the U.S., with the exception of the U.K. study (which uses a smaller sample of medium to large manufacturing firms in the U.K.).[75] These studies report that firm managers in most industry sectors believe they can raise imitation costs significantly through non-patent devices such as lead time, learning advantages, secrecy and complementary manufacturing capabilities, and that patents are generally one of the least effective tools for appropriating returns from R&D investments.[76] Except for pharmaceuticals and chemicals, industry participants believed that patents were generally not essential for the development or introduction of new products.[77] It is also noteworthy that the Cohen et al. study reports similar results for 1994,[78] which means that industry managers

---

A STUDY OF THE BRITISH EXPERIENCE (1973).

[75] The U.K. study's dataset consists of about forty-five medium to large manufacturing firms in the U.K. chemicals, pharmaceuticals, oil refining, electrical and electronic equipment, mechanical equipment and textile industries. *See id.* at 82. However, responses in some cases were less than complete, so that the study's results are sometimes based on an even smaller sample set.

[76] *See* COHEN ET AL., *supra* note 8 (finding that most industries viewed patents as the least effective mechanism for directly appropriating the economic returns from an innovation); Levin et al., *supra* note 34 (surveying manager expectations and finding that managers often view patents as ineffective mechanisms to protect intellectual property from imitators, and that managers use alternative devices to appropriate returns from R&D investments). The findings of the above survey studies received further indirect confirmation by a study conducted by Mansfield, who surveyed U.S. firms' willingness to enter into joint ventures for the production or sale of goods in Spain, Japan and various emerging market jurisdictions. Mansfield found that a firm's degree of unwillingness to produce or market its products in jurisdictions with no or weak patent protection varies with industry type, such that companies in many industries are still willing to invest even where protection is weak, tracking the results of Mansfield's early survey evidence, pharmaceutical and chemical companies are far more unwilling than all other industries. In all cases, survey respondents indicated that patent protection was only one of many factors relevant to such an investment decision. *See* Edwin Mansfield, *Unauthorized Use of Intellectual Property: Effects on Investment, Technology Transfer and Innovation*, *in* NATIONAL RESEARCH COUNCIL, GLOBAL DIMENSIONS OF INTELLECTUAL PROPERTY RIGHTS IN SCIENCE AND TECHNOLOGY 107 (Mitchel B. Wallerstein et al. eds., 1993).

[77] *See* Mansfield, *Patents and Innovation*, *supra* note 73 (finding that, based on interviews with firm managers in a random sample of 100 U.S. manufacturing firms, in all industries other than chemicals and pharmaceuticals, corporate managers believed that less than thirty percent of the inventions that firms developed during 1981 to 1983 would *not* have been developed without patent protection). In a more recent study that uses data collected in the Cohen et al. study and then seeks to estimate the value of patent protection to the patenting firm, Arora et al. found that, other than in a few select industries, patenting the typical innovation is not profitable once patenting costs are taken into account. *See* ASHISH ARORA ET AL., R&D AND THE PATENT PREMIUM 29-31, 35 (Nat'l Bureau of Econ. Research, Working Paper No. 9431, 2003) [hereinafter ARORA ET AL., R&D]. The industry-variable importance of patent protection finds further support in the fact that the ratio of patents issued per innovation produced is relatively higher in the chemical and petroleum industries and relatively lower in certain other industries. *See* Zoltan J. Acs & David B. Audretsch, *Innovation in Large and Small Firms: An Empirical Analysis*, 78 AMER. ECON. REV. 678, 683 (1988) [hereinafter Acs & Audretsch, *Innovation in Large and Small Firms*].

[78] *See* COHEN ET AL., *supra* note 8. This finding partly responds to some commentators'

have not revised their low opinion of the relative effectiveness of patent protection despite the creation in 1982 of the Court of Appeals for the Federal Circuit which is widely believed to have increased the strength of available patent protection.[79] Using its smaller sample set, the U.K. study similarly finds that, outside pharmaceuticals and chemicals, industrial managers seldom believe that patent protection is a necessary precondition for going ahead with a research project.[80]

Given market participants' skepticism concerning the value of patent protection, it is not surprising to learn that they also tend to believe that patents have a limited effect on raising competitors' imitation costs and extending competitors' imitation time. The Yale study reported that almost all patented innovations are duplicated within five years.[81] Similarly, in an additional study, Mansfield et al. found that sixty percent of patented innovations are imitated within four years of their initial introduction and that industry participants felt that patents generally delay imitators' entry by less than a few months.[82] More specifically, Mansfield et al. found that industry participants believe that patents only increase the costs of imitation by a median eleven percent, ranging from seven percent in electronics and machinery, to ten percent in chemicals and thirty percent in pharmaceuticals.[83] Interestingly,

---

observations that the survey studies' findings may change given the strengthening of patent protection since the creation of the Federal Circuit. *See* David C. Mowery, *Global Intellectual Property Rights Issues in Perspective*, *in* NATIONAL RESEARCH COUNCIL, GLOBAL DIMENSIONS OF INTELLECTUAL PROPERTY RIGHTS IN SCIENCE AND TECHNOLOGY, *supra* note 76, at 369-70 (noting that existing survey studies regarding managers' views of the effectiveness of patents need to be updated to reflect the creation of the Federal Circuit and the subsequent strengthening of patent protection and to reflect the extension of patent protection to previously ineligible subject matter).

[79] Between 1953 and 1978, circuit courts affirmed sixty-two percent of district court decisions holding patents to be valid and infringed, and reversed twelve percent of the decisions holding patents to be invalid or not infringed. Between 1982 and 1990, the Federal Circuit affirmed ninety percent of district court decisions holding patents to be valid and infringed, and reversed twenty-eight percent of the decisions holding patents to be invalid or not infringed. *See* JEAN O. LANJOUW & JOSH LERNER, THE ENFORCEMENT OF INTELLECTUAL PROPERTY RIGHTS: A SURVEY OF THE EMPIRICAL LITERATURE (Nat'l Bureau of Econ. Research, Working Paper No. 6296, 1997).

[80] *See* TAYLOR & SILBERSTON, *supra* note 74, at 195-98. The authors found that in twenty-four out of the thirty-two returns, it was estimated that R&D was either not, or was minimally, affected by prospective patent protection. The six returns in which R&D was believed to be substantially affected by patent protection were in the pharmaceuticals, chemicals and heavy industrial plant sectors. When these results are converted into figures representing percentages of firm R&D, it appears that only pharmaceuticals and chemicals are especially dependent on patent protection.

[81] *See* Levin et al., *supra* note 34.

[82] *See* Mansfield et al., *supra* note 33. The study is based on data obtained from firms in the chemical, drug, electronics and machinery industries concerning the cost and timing of imitating forty-eight product innovations. The firms were chosen "more or less randomly" from among the major firms in these four industries in the northeastern U.S. and the new products were chosen "more or less randomly" from among those introduced recently by these firms. *See id.*

[83] *See id.*

econometric estimates of the private value of patents (derived from renewal fees and renewal rates in particular jurisdictions) reach roughly similar figures.[84] Even given these gloomy data regarding the effectiveness of patent protection, it would be possible to surmise that patents may be of help in industries that lack market-generated, informal appropriation mechanisms (for example, in industries where new technologies are transparent and inexpensive to imitate). But the Cohen et al. study considers and rejects this possibility, finding that patents offer almost no aid in industries where non-patent appropriation mechanisms are ineffective and patents are generally only helpful in industries that are already able to exploit a variety of non-patent appropriation mechanisms.[85] Thus, the marginal value of patent protection as a cure to any waiting game appears to be severely limited.

2. R&D Investment

The incentives theory would predict that, *ceteris paribus*, increasing the strength of available patent protection should induce more R&D investment, while decreasing the strength of available patent

---

[84] Several researchers have attempted to estimate the private value of applying for and maintaining a patent based on renewal frequency and renewal fees. This econometric estimating procedure relies on the assumption that a patent holder will only renew a patent if it believes that the renewal fee is exceeded by the expected discounted present value of future benefits attributable to the patent in the period until the next renewal and to the option to renew the patent again. In a survey of patents issued in France, Mark Schankerman found that the private value of patents represents, on average, an implicit subsidy of 15.6% expressed as a percentage of the total company-funded and government-funded R&D conducted by the relevant patent holder. *See* Mark Schankerman, *How valuable is patent protection? Estimates by technology field*, 29 RAND J. ECON. 77, 96-97 (1998). That is, without patent protection, a subsidy equal to 15.6% of existing R&D expenditures would need to be given in order to sustain existing levels of R&D investment. In another study, Ariel Pakes finds an implicit subsidy of 15.5% with respect to patents issued in France, 13.8% with respect to patents issued in Germany and eleven percent for patents issued in the U.K. *See* Ariel Pakes, *Estimates of the Value of Holding European Patent Stocks*, 54 ECONOMETRICA 755, 778-79 (1986) [hereinafter Pakes, *Patents as Options*]. Jean Lanjouw estimates equivalent subsidy rates for company-funded and government-funded R&D, with respect to patents issued in Germany, of 6.8% for pharmaceuticals, 10.4% for computers, 11.5% for engines and 38.3% for textiles. *See* Jean Lanjouw, *Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value*, 65 REV. ECON. STUD. 671 (1998) [hereinafter Lanjouw, *Patent Protection in the Shadow of Infringement*]. Using an alternative methodology applied to data collected by the Cohen study, Arora et al. found an average equivalent subsidy of seventeen percent (with biotechnology, drugs and medical instruments having the highest equivalent subsidy rates), a result roughly consistent with that reached by the above-described studies. *See* ARORA ET AL., R&D, *supra* note 77, at 32-33.

[85] *See* Levin et al., *supra* note 34. The Yale study searched for "clusters" of industries that display similar appropriability characteristics. In doing so, the Yale study identified three clusters: (1) in cluster 1, no appropriation mechanism (including patents) was effective except for sales and service efforts; (2) in cluster 2, only lead time, learning curve, and sales and service efforts were effective appropriation mechanisms; and (3) in cluster 3, all the above appropriation mechanisms, in addition to patents, were effective.

protection should have the opposite effect. But the evidence gleaned from some natural experiments strongly suggests otherwise. Researchers have consistently found little correlation between increases or decreases in the strength of patent protection (as measured by judicial and legislative action) and, respectively, increases or decreases in R&D investment, innovative output, or industry growth rates. Industries that are widely believed to lack strong or legally certain patent protection, such as the software, computer, and semiconductor industries, have grown rapidly despite the absence of such protection.[86] It may appear possible to argue that such growth may have been even greater if patent protection were made available or strengthened. But the software industry experienced no growth or a reduction in growth and R&D intensity (the ratio of R&D spending relative to sales or output) following court decisions in the early 1980s that extended patent protection to software applications[87] and the subsequent rise in the number of software patent applications.[88] Although patenting volume does seem to move in response to changes in the strength of available patent protection,[89] several studies have shown that historical declines and increases in patenting in the U.S. do not appear to correspond to declines or increases in R&D investment. Despite the enormous rise in patenting since the creation of the Federal Circuit (and the simultaneous strengthening of patent protection and expansion of patentable subject

---

[86] With reference to software and computers, see BESSEN & MASKIN, SEQUENTIAL INNOVATION, *supra* note 11, at 1-2, 15-16. With reference to semiconductors, see Bronwyn H. Hall & Rosemarie H. Ziedonis, *The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979-1995*, 32 RAND J. ECON. 101 (2001) (reviewing patenting by approximately 100 publicly traded U.S. semiconductor firms from 1975 to 1998 and finding that "incumbent" firms (i.e., firms who entered the industry before 1983) show a slight decline in R&D intensity and entrants show some increasing R&D intensity at first, followed by a decline).

[87] *See* BESSEN & MASKIN, SEQUENTIAL INNOVATION, *supra* note 11, at 14-16. The authors find that R&D spending relative to sales and output in the software industry has remained roughly steady or even declined since the courts ruled that software was eligible (to a certain extent) for protection by patents and not just the weaker protections afforded by copyright. The authors obtained these results from three samples of firms: the top ten U.S. software patentees in 1995 (accounting for thirty-five percent of software patents issued to U.S. companies in that year), certain industry groupings for computer hardware and programming services, and a group of computer, telecommunications and electronic components companies. Relying on other studies, the authors eliminated alternative explanations for reduced R&D spending such as a decline in technological opportunities or a rise in R&D costs. Of course, it is possible to argue that the decline in R&D spending and growth rates in the software industry subsequent to the extension of patent protection could be the result of other factors. However, Bessen and Maskin argue that this is probably not the case by confirming that there was no contemporaneous decline in technological opportunity or increase in the cost of performing R&D during the relevant period. *See id.* at 18-20.

[88] Applications for software-related patents went from 250 in 1980 to 21,000 in 1999. *See* Evan P. Schultz, *Too Many Patents?*, LEGAL TIMES, Mar. 21, 2002.

[89] *See* SAMUEL KORTUM & JOSH LERNER, STRONGER PROTECTION OR TECHNOLOGICAL REVOLUTION: WHAT IS BEHIND THE RECENT SURGE IN PATENTING? 1 (Nat'l Bureau of Econ. Research, Working Paper No. 6204, 1997).

matter),[90] the rise in U.S. industrial R&D spending since that time predates 1982, and the intensity of research efforts (usually measured by the ratio of R&D expenditures to sales) remained static in the late 1980s and early 1990s.[91] Contrary to what would be predicted by the incentives thesis, firms do not apparently alter R&D investment in response to changes in available patent protection, but simply elect to patent a smaller or larger percentage of their innovations. Thus, researchers have found that the surge in patenting since the mid-1980s did not correspond to an increase in the number of inventions but rather an increase in the percentage of inventions patented.[92] Conversely, the drop in patenting in the U.S. from the late 1960s to the early 1980s, which approximately corresponds to a decline in the strength of patent protection as a result of a pattern of hostile judicial rulings, apparently corresponds to a decline in the percentage of inventions patented rather than a decline in the number of inventions.[93]

3. Definitional Barriers

Because a patented innovation, unlike a parcel of land, generally cannot be precisely or even reasonably delineated without expenditure of very considerable time and expense by the patent applicant and the patent examiner, most issued patents are vague to a significant degree and therefore vulnerable to infringement attempts by, and boundary disputes with, third-party innovators.[94] But evidence shows that patent

---

[90] *See id.* The authors note that applications for U.S. patents by U.S. inventors have risen more, in absolute and percentage terms, from 1985 to 1996 than in any other previous decade. As of the date of the study (1996), the number of patents issued had reached an all-time high in 1996. From 1983 to 1991, total patenting in the U.S. rose by almost seventy percent. *See id.* at 1, 22.

[91] *See id.*

[92] *See id.* The authors reach this conclusion largely by relying on the fact that (1) jurisdictions outside the U.S. have not similarly increased available patent protection; (2) U.S. inventors have been patenting at similarly increased rates both in the U.S. and in foreign jurisdictions; and (3) the U.S. has not increased as a destination for patents by foreign inventors. If the surge in patenting were thought to be a function of increased available patent protection, it would be expected that the U.S. would attract an increased number of patent applications from foreign inventors (given the U.S.'s more drastic increases in available patent protection) and that U.S. inventors would not patent at similar rates outside the U.S. (given that foreign jurisdictions have not matched the U.S. increases in available patent protection). To reach this conclusion, the authors examined data on patent applications for 1955-93 in Germany, France, the United Kingdom, Japan and the United States by inventors from each of these five countries or elsewhere. For a critical discussion of Kortum & Lerner's results, see Hall & Ziedonis, *supra* note 86.

[93] *See* Mansfield, *Patents and Innovation*, *supra* note 73. The study reports an increase in the percentage of patentable inventions that are patented from 1965-69 and 1980-82. The study therefore concludes that the decline in patenting from the late 1960 to the early 1980s must represent a drop in the number of inventions.

[94] *See* Long, *supra* note 11, at 633-35 (noting that much of the patent literature agrees that patent protection is often incomplete, partly because perfect delineation and enforcement of

examiners devote remarkably little time to each patent that issues.[95] Even if patent examiners' resources were greatly expanded, however, it is not clear that many patents' definitional specificity would be significantly improved. The degree to which an invention can be defined with clarity depends closely on whether the relevant innovation arises in a "discrete" or a "complex" (also known as "cumulative" or "systems") technology sector.[96] In the former case, which would generally include products consisting of a single or a few components, there may be a very close correspondence between the patent and the underlying product, which may provide the patent holder with a strong monopoly (just as the incentives argument proposes).[97] In the latter case, which would generally include products consisting of numerous components and, consequently, hundreds or even thousands of patents, any innovator intent on securing a dominant market position based primarily on patent protection probably will have great difficulty in doing so.[98] Empirical research suggests that many, if not most, and certainly today's most technologically fertile industries, fall into the "complex technology" category.[99] Moreover, even in the case of a discrete invention (and especially in the case of a complex innovation), patents tend to be granted fairly early in the inventive process.[100] This means that patent claims may often overstate, understate or otherwise misconstrue the exact contours of the final product and its immediately related applications, thereby frustrating later attempts to enforce the patent against third-party infringers.

These arguments are supported by empirical evidence. The survey studies of industrial managers described above find that patents are generally considered to be more effective in the case of innovations such as chemicals and simple mechanical products in which it is

---

patent rights is impossible); Anand & Galetovic, *supra* note 35, at 618 n.5 (noting that "intellectual-property rights may be weak when it is difficult to clearly specify the boundaries of the knowledge being contracted upon").

95 *See* Lemley, *Rational Ignorance*, *supra* note 69 (stating that the total average time a PTO examiner spends on a patent that issues over a typical two to three-year prosecution period is eighteen hours, and that "[i]t is not surprising . . . [that] the PTO issues many patents that would have been rejected had the examiner possessed perfect knowledge").

96 For an explanation of these terms, see *supra* notes 7-9 and accompanying text.

97 For a discussion of this point, see James Bessen, *Patent Thickets: Strategic Patenting of Complex Technologies*, at 1 (Aug. 2002), *available at* http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID327760_code020924670.pdf?abstractid=327760 (last visited Mar. 19, 2004) [hereinafter Bessen, *Patent Thickets*]; Robert P. Merges, *Intellectual Property Rights and the New Institutional Economics*, 53 VAND. L. REV. 1857, 1859 (2000) [hereinafter Merges, *New Institutional Economics*].

98 *See* Bessen, *Patent Thickets*, *supra* note 97; Merges, *New Institutional Economics*, *supra* note 97.

99 *See* Bessen, *Patent Thickets*, *supra* note 97.

100 *See* Roberto Mazzoleni & Richard R. Nelson, *The Benefits and Costs of Strong Patent Protection: A Contribution to the Current Debate*, 27 RES. POL'Y 273, 276-77 (1998).

relatively easy to distinguish between the patented product and an allegedly infringing product.[101] In the case of these discrete technologies, it is both difficult for competitors to develop non-infringing "invent around" technologies (since, for example, even a minute change to a chemical molecule results in an entirely non-substitutable product) and easy for patent holders to detect and show infringement by third parties. By contrast, because it is difficult to assemble a patent portfolio or to draft a single patent covering all elements of a cumulative technology, competitors can more easily "invent around" patents covering such technologies or can block further development[102] (usually contingent on negotiating a mutually beneficial cross-licensing or other technology-exchange arrangement[103]). Thus, as the complexity of a technology increases, the strength of its patent protection tends to fall. The more complex the underlying technology, the less clear the boundaries of the applicable patent and the greater the insecurity of the patent as an appropriation mechanism, even after the considerable effort and financial sums expended in the patent prosecution process by patent applicants and examiners.

4. Renewal Data

The low rates at which patents are renewed gives some limited evidence as to the low value of patents as an appropriation mechanism. Most jurisdictions require that patent holders pay renewal fees in order to maintain their patent.[104] Data regarding the percentage of patents that are renewed in various jurisdictions show that patents apparently lose their value quickly since most patents are not renewed to the full statutory term and a large percentage are not renewed even early in the statutory term. In the U.S., more than half of all patents are not

---

101 *See* Levin et al., *supra* note 34, at 798; COHEN ET AL., *supra* note 8; Mansfield, *supra* note 36.

102 *See* Bessen, *Patent Thickets*, *supra* note 97; Merges, *New Institutional Economics*, *supra* note 97; *see also* Peter C. Grindley & David J. Teece, *Managing Intellectual Capital: Licensing and Cross-Licensing in Semiconductors and Electronics*, 39 CALIF. MGMT. REV. 8, 9 (1997) (noting that, in cumulative technology fields such as electronics and semiconductors, "[t]here are inevitably overlapping developments and mutually blocking patents").

103 For an extensive discussion of cross-licensing practices and strategies, see Grindley & Teece, *supra* note 102.

104 All U.S. patents issued as a result of applications filed after December 1980 must be maintained by the payment of renewal fees in increasing amounts at periods of three and a half years, seven and a half years, and eleven and a half years after the patent issues. *See* 35 U.S.C. § 41(b) (West Supp. 2000). Failure to pay these fees effectively represents abandonment of the patent since a patent holder cannot bring a patent lawsuit if it has not paid the maintenance fees. *See* Lemley, *Rational Ignorance*, *supra* note 69, at 9-10. Renewal fees have been in force in European countries for over forty years. *See* KORTUM & LERNER, *supra* note 89, at 26.

renewed within ten years of the date of patent application[105] and more than two-thirds of all patents are not renewed to the end of the twenty-year statutory term.[106] Similarly low rates of renewal have been found with respect to German, French and U.K. patents.[107] Such evidence of low renewal rates may be attributed either to an industry perception that patents offer limited ability to appropriate further returns or to high rates of technological obsolescence. Either of these possibilities leads patent holders to conclude that the cost of the renewal fee is greater than the expected discounted value of future returns appropriable through patent protection.[108] As a further alternative, low renewal rates may have little to say about the relative effectiveness of patent protection to the extent that they simply reflect the fact that innovations are patented early in the innovative process and most turn out to have no or limited commercial application.[109]

III. THE MARKET-ENTRY THESIS

In this Part, I propose a market-entry theory of patent protection, both as a positive account of patenting practices and as a normative justification for patent awards given the widespread availability of non-patent appropriation mechanisms and the limited effectiveness of patent entitlements. I then consider, and respond to, an important objection to

---

105 *See* Francesca Cornelli & Mark Schankerman, *Patent renewals and R&D incentives*, 30 RAND J. ECON. 197 (1999).

106 *See* Lemley, *Rational Ignorance*, *supra* note 69, at 10. Despite the general impression that industry accords patents greater significance than in the recent past, renewal rates for U.S. patents actually fell in the 1990s. *See* KORTUM & LERNER, *supra* note 89, at 26 (noting study indicating that percentage of U.S. patents for which fourth year renewal fees were paid fell from eighty-four percent in 1991 to seventy-nine percent in 1994, and percentage of U.S. patents for which eight year renewal fees were paid fell from seventy-four percent in 1991 to sixty-six percent in 1994).

107 Reaching similar results, Jean Lanjouw has found that less than thirty-five percent of all German patents are maintained until the maximum term and fewer than fifty percent of German patents are maintained more than ten years. *See* Lanjouw, *Patent Protection in the Shadow of Infringement*, *supra* note 84, at 693. The sample set consists of a group of over 2000 West German patents issued based on applications filed during 1953-88 and covering the computer, textiles, engines and pharmaceutical industries. Similarly, Ariel Pakes has found that only seven percent of French patents and eleven percent of German patents are maintained until the expiration date. *See* Pakes, *Patents as Options*, *supra* note 84, at 774 (Fig. 4). The author's sample includes data relating to renewal of roughly one million patents applied for in France between 1951 and 1979 and roughly half a million patents issued in West Germany between 1952 and 1972.

108 *See* KORTUM & LERNER, *supra* note 89, at 26-27; *see also* Grindley & Teece, *supra* note 102, at 27 (noting that in the early phase of the semiconductor industry, many firms often did not patent products or process because product life cycles were short and therefore there was little economic justification for expending resources on obtaining patents that would soon be obsolete).

109 Note that the distribution of the economic value of patented innovations is known to be highly skewed, with a few patented innovations having enormous economic value and many having little or no economic value. *See infra* note 154.

the market-entry thesis in connection with the widespread use of the patent system by large, incumbent firms that have easy access to alternative protection devices.

A. *The Positive Thesis*

The large body of empirical evidence concerning the relative effectiveness of patent and non-patent appropriation mechanisms obviously poses a serious challenge to the standard incentives argument. The stronger the evidence exists that (1) innovators already have access to effective, non-patent means for appropriating a significant portion of the proceeds from intellectual goods and (2) even where innovators do not have access to such alternative appropriation devices, patents offer little help in capturing additional innovation proceeds, the weaker the account offered by the incentives theory of the patent system. Put simply, existing evidence suggests that the waiting game is not a serious danger in most industries, and furthermore, where it does arise, patents are of little help in alleviating any incentives shortfall. This evidence certainly calls for a revised positive account of patenting behavior (and, as discussed further below, a revised normative explanation of patent awards). But it does not call for concluding that patents lack any significant exclusionary function[110] (and, as a normative matter, certainly does not call for jumping to the extreme of patent abolitionism).[111]

Such extreme conclusions are ruled out by an important wrinkle in the relevant empirical evidence. Although it is clear that alternative protection devices are widely available generally speaking, it appears unlikely that these devices are *equally* available to all firms. In particular, there is compelling reason to believe that entrants in technology markets necessarily have no or significantly less access to the potent appropriation mechanisms related to the first-mover advantage. Although a small-firm or large-firm entrant can adopt secrecy practices and other copy-prevention technologies just as easily as a large-firm incumbent, it cannot enjoy the manufacturing efficiencies, reputational capital, customer loyalty, brand image, distribution relationships, and learning advantages that a firm can only acquire over a considerable period of successful participation in the

---

[110] For arguments that patents generally or sometimes lack any significant exclusionary function, see Long, *supra* note 11; BESSEN & MASKIN, SEQUENTIAL INNOVATION, *supra* note 11.

[111] For a particularly strident proposal to entirely abolish the patent system on a mixture of utilitarian and libertarian grounds, see Palmer, *supra* note 4, at 300-04. For an interesting (and cautionary) history of "patent abolitionism" in England and the U.S., see Mark D. Janis, *Patent Abolitionism*, 17 BERKELEY TECH. L.J. 899 (2002).

relevant market.[112] Additionally, large-firm incumbents that have already interacted repeatedly with their competitors, and can credibly commit to doing so in the future, are far more able to exploit industry and other reputational norms in order to quickly resolve patent and other intellectual property boundary disputes without recourse to expensive formal litigation.[113] This appropriation method is almost entirely foreclosed to small firms and largely foreclosed to large-firm outsiders to the extent that their reputational capital is not especially applicable to the relevant industry (e.g., Coca-Cola enters the guided missile market). Because potential entrants have access to a much more limited and less effective set of alternative appropriation devices, patents may offer significantly greater marginal value as a protection device.[114]

There is some important empirical evidence to bolster these theoretical arguments. First and foremost, the survey evidence regarding the limited effectiveness of patents all use sample sets that exclude small firms (generally defined as firms with less than approximately $25 million in annual sales).[115] As the survey authors as well as commentators note, these surveys do not test, and therefore may underestimate, the perceived effectiveness of patent protection relative to non-patent appropriation devices among small firms that lack an established market position.[116] Second, although the marginal costs of

---

[112] On this point, see ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 261-62, stating that smaller firms lack the means available to larger firms to protect their innovations, such as "manufacturing and commercialization assets"; Zoltan J. Acs & David B. Audretsch, *Innovation as a Means of Entry: an Overview*, *in* P.A. GEROSKI & J. SCHWALBACH, ENTRY AND MARKET CONTESTABILITY 224-25 (1991), stating that entry is less viable when innovation depends more heavily on tacit knowledge accumulated through market experience; Jonathan B. Baker, *Fringe Firms and Incentives to Innovate*, 63 ANTITRUST L.J. 621, 622 (1995), noting that "fringe firms"—that is, firms with smaller market shares—are likely to lack some of the advantages that the leading firms possess, such as reputation for quality, access to inexpensive or high quality inputs, or effective distribution; COHEN ET AL., *supra* note 8, noting that, among non-patent appropriation mechanisms, lead time and complementary manufacturing capabilities take more time and money to acquire than secrecy.

[113] *See* LANJOUW & SCHANKERMAN, *supra* note 63, at 4-5.

[114] For a similar but more extreme version of this argument, see ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 261-62, arguing that patents are of greater value to smaller firms because they have no means other than patents to appropriate rents from their innovation.

[115] For a description of these sample sets, see *supra* notes 71-75.

[116] *See* Richard Gilbert & Zvi Griliches, *Appropriating the Returns from Industrial Research and Development: Comments and Discussion*, *in* 3 BROOKING PAPERS ON ECONOMIC ACTIVITY, *supra* note 34, at 831 (stating that patents may be much more important for start-ups because they give the firm an asset that can be sold if the firm later wants to sell out to an acquirer); Levin et al., *supra* note 34, at 797 (stating that patents may be the principal or sole asset of small firms and their sole means of appropriating returns since investment in alternative means, such as complementary sales and service efforts, may not be feasible); Mazzoleni & Nelson, *supra* note 100, at 276 (noting that survey studies showing limited effectiveness of patents suffer because the survey sample is generally restricted to large firms already having established market share and therefore may underestimate the importance of patents for smaller firms attempting to penetrate an established market).

patent prosecution must be greater for small firms,[117] several researchers have found that small firms exhibit a greater propensity to patent, relative to larger, more established firms[118] and take greater legal efforts to defend their patents against infringers.[119] Other researchers have found that smaller unlisted companies and domestic individuals in the U.S. are more likely to file a patent suit than larger domestic listed firms, and small listed companies are far more likely to file suits than larger listed companies.[120] Aggressive litigation defense by small firms suggests that patents are of greater marginal value to these firms, especially considering the fact that litigation costs are more burdensome for a smaller firm with lower cash reserves and a weaker ability to raise external financing. Unlike large-firm incumbents, small firms cannot easily rely on non-patent methods, such as leverage exerted as a result of long-standing business relationships, to dissuade potential infringers or to settle patent boundary disputes, and are therefore forced to resort more frequently to the expensive litigation process.[121] These data

---

117 *See* F.M. Scherer, *Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions*, 55 AMER. ECON. REV. 1097, 1104-05 (1965) [hereinafter Scherer, *Firm Size*] (stating that the marginal cost of patenting may be more for smaller firms dependent upon more expensive outside counsel, unlike large firms that often employ less expensive in-house patent counsel).

118 *See* Long, *supra* note 11, at 643-44; Zoltan J. Acs & David B. Audretsch, *Innovation and Technological Change: An Overview*, *in* INNOVATION AND TECHNOLOGICAL CHANGE: AN INTERNATIONAL COMPARISON 11-12 (Zoltan J. Acs & David B. Audretsch eds., 1991) [hereinafter Acs & Audretsch, *Innovation and Technological Change*]; Scherer, *Firm Size*, *supra* note 117. Propensity to patent is usually measured by a ratio of patents issued relative to firm size or, alternatively, patents issued relative to firm sales, which is found to vary inversely (propensity to patent declines as firm size or firm sales amount increases). Evidencing a similar point, research shows that smaller firms account for a disproportionately large number of patents relative to their share of total R&D expenditures. *See* PHILLIP AREEDA & LOUIS KAPLOW, ANTITRUST ANALYSIS: PROBLEMS, TEXTS, CASES 31 (5th ed. 1997) (noting that research indicates that firms with more than 5000 employees spend eighty-eight percent of private R&D dollars but receive only fifty percent of the patents issued); Scherer, *Firm Size*, *supra* note 117, at 1104-05 (finding that firms with sales of less than fifty-five million dollars in 1955 accounted for approximately 43.9% of patents issued to manufacturing corporations and noting other data finding that small firms accounted for only fourteen percent of U.S. R&D expenditures). In the semiconductor industry, entrants tend to patent more per R&D dollar invested than larger firms. *See* Hall & Ziedonis, *supra* note 86, at 107-08. The study defines entrants as start-ups and firms that are not publicly traded. Moreover, small firms in the semiconductor industry, which are generally responsible for designing (rather than manufacturing) new semiconductor chips, report that they patent generally for exclusionary reasons, while larger firms, which are generally responsible for manufacturing chips, report that they patent generally in order to enable the exchange of know-how and technological knowledge with their competitors. *See id.* at 110-11.

119 *See infra* note 120 and accompanying text.

120 *See* LANJOUW & SCHANKERMAN, *supra* note 63, at 4-5. The authors study the determinants of patent suits in the U.S. during 1978-99, using a database of 13,625 patent cases filed during this period. *But see* Lerner, *Patenting in the Shadow of Competitors*, *supra* note 69, at 471-72 (noting that a 1990 survey of 376 firms found that, relative to the larger firms in the sample, the smaller firms believed that their patents were infringed more frequently but were less likely to litigate these infringements due to the cost of doing so).

121 *See* LANJOUW & SCHANKERMAN, *supra* note 63, at 4-5; *see also* Scherer, *Firm Size*, *supra* note 117, at 1113 (noting that small firms may be more eager to patent in order to combat larger

suggest strongly that patents must often represent the principal valuable asset of a smaller firm and therefore, a smaller firm more often finds it worthwhile to take expensive legal action against a patent infringer.

To the extent that small and outsider firms lack access to many private appropriation mechanisms that are widely available to large-firm incumbents, patents may be necessary to induce these firms to contest incumbents' market position by investing funds in ambitious R&D projects. Suppose that an entrant has developed a product that offers superior quality or price (that is, cost) attributes and would substitute for the incumbent's product. In a world without patents, the incumbent firm could often adopt a "fast second" strategy—allowing the entrant to incur the high product development costs and bear the high risk of technical and market failure, then imitating the entrant's resulting product if it proves successful, and, finally, defeat the entrant through its superior brand image, customer base, and manufacturing and distribution efficiencies.[122] Given the possibility of this "fast second" response from the incumbent firm, no potential entrant will be willing to contest the incumbent's dominant position (except in the case of technologies that are especially difficult to reverse engineer) even if it could offer a product or product line superior to that of the incumbent in terms of price or performance. Where small and outsider firms seek entry into a concentrated market that is dependent on technological innovation, patent protection may be the most effective means of overcoming the powerful first-mover advantage enjoyed by the incumbent firm or group of firms. Hence small firms are eager to patent and to defend their patent holdings against infringers.

Although this line of reasoning employs some of the structure of the conventional incentives argument, it reaches a notably different result. The conventional incentives argument proposes that patent protection creates an artificial monopoly that is a necessary evil because the market fails to supply all or most innovators with sufficient exclusionary devices in order to recoup innovation proceeds. Without patents, no or almost no firm will be willing to invest resources into uncertain and costly research ventures. By contrast, the market-entry thesis proposes that exclusionary devices are abundant but unequally distributed between incumbents and entrants, and therefore only entrants lack significant innovation incentives in the absence of patent protection. By providing entrants with at least some appropriation capabilities, patent protection encourages such firms to contest incumbents' dominant market position through the development of novel technologies. Rather than curing an undersupply of innovative output by the market in general, patents correct an undersupply of

---

firms' competitive advantage in terms of economies of scale in distribution and production).

122 *See* Baker, *supra* note 112, at 638-39.

innovative output by small-firm and large-firm entrants in particular.

B. *The Normative Thesis*

The market-entry thesis naturally recalls Joseph Schumpeter's notion of "creative destruction." In his famous thesis, Schumpeter extolled the figure of the entrepreneur, whose new ideas and products contest and ultimately undermine the dominant position of established firms.[123] The entrepreneur then in turn takes a dominant market position, which will later become the target of further destruction by a subsequent aspiring outsider.[124] Schumpeter identified this process of continual turnover as the central driving force in economic development in a liberal capitalist economy.[125] The market-entry theory of patent protection offers a small revision to this picture by suggesting that the Schumpeterian hero requires an "invisible hand" (of the state and *not* the market) to help him or her along. If this is true as a positive account of patenting practice, we must then confront the normative question of whether extending this invisible hand (in the form of patent protection) to the entrepreneur enhances social welfare. That is, a normative defense of the patent system requires showing why the state should use patent protection in order to enable small firms to enter technology markets in which incumbents already have access to effective appropriation mechanisms.

In this section, I propose such a normative argument. I advance a number of theoretical and empirical grounds suggesting that patent protection, by encouraging small firms and outsider firms to make R&D investments in concentrated technology markets, can play an important role in improving the direction or quality of research outcomes. Specifically, by rendering contestable incumbent-dominated markets, patent protection may induce potential small-firm and large-firm entrants to attempt entry by undertaking risky product development projects. As a consequence, the aggregate industry R&D portfolio is tilted away from incremental research targeted at developing minor improvements that contribute minimal social value and toward fundamental research targeted at developing breakthrough innovations that contribute immense social value.

This argument relies on a distinction introduced by Richard Nelson and Sidney Winter between "routinized" and "entrepreneurial" regimes

---

123 *See* JOSEPH A. SCHUMPETER, THE THEORY OF ECONOMIC DEVELOPMENT 128-56 (Redvers Opie trans., Oxford Univ. Press, 1934).

124 *See id.*

125 *See id.*

of innovation investment.[126] Much of the innovation literature agrees that monopolists and incumbent firms holding dominant market positions tend to carry out routinized or incremental research, consisting principally of cost-reducing "debugging," incremental process modifications and minor product improvements that enhance brand image, all of which add relatively little to the existing knowledge base, but help to preserve an incumbent's market share.[127] By contrast, the innovation literature also tends to agree that small firms and other entrants carry out a disproportionate share of entrepreneurial research consisting of radical innovations that significantly expand the existing knowledge base and permit the innovator to penetrate a concentrated market dominated by a few incumbents.[128] Additionally, small firms may speed up technological advance by increasing "research diversity"—that is, by multiplying the number of approaches applied in order to resolve a given research problem.[129] As Schumpeter suggested in the "early" period of his career, it is small firms that are primarily the engine of breakthrough innovation and the resulting process of creative destruction.[130] Without breakthrough innovation, an industry's R&D

---

126 *See* NELSON & WINTER, *supra* note 8; Sidney G. Winter, *Schumpeterian Competition in Alternative Technological Regimes*, 5 J. ECON. BEHAV. & ORG. 287, 293-97 (1984). For a similar distinction between "routinized" and "adaptive" forms of innovation, see P.R. Beije & J. Groenewegen, *A Network Analysis of Markets*, 26 J. ECON. ISSUES 87, 101-02 (1992). On the related distinction between "radical" and "incremental" innovation, see Gerard Fairtlough, *Innovation and Organization*, *in* THE HANDBOOK OF INDUSTRIAL INNOVATION 325-26 (Mark Dodgson & Roy Rothwell eds., 1996).

127 *See* Beije & Groenewegen, *supra* note 126.

128 *See id.*; *see also* Richard J. Rosen, *Research and Development with Asymmetric Firm Sizes*, 22 RAND J. ECON. 411 (1991) (noting that the empirical literature "indicates that in general . . . smaller firms make a disproportionate share of major innovations").

129 *See* NELSON & WINTER, *supra* note 8; *see also* BESSEN & MASKIN, SEQUENTIAL INNOVATION, *supra* note 11 (arguing that, in industries characterized by sequential and complementary innovation, firms may welcome competition because competitors' parallel R&D and imitative and improvement efforts can expand purchaser demand by generating additional applications); Wesley M. Cohen & Steven Klepper, *Firm Size Versus Diversity in Achievement of Technological Advance*, *in* INNOVATION AND TECHNOLOGICAL CHANGE: AN INTERNATIONAL COMPARISON, *supra* note 118, at 947-48 [hereinafter Cohen & Klepper, *Firm Size Versus Diversity*] (stating that, even if large firms have attributes conducive to R&D, industry consolidation may not be desirable as an innovation policy due to the resulting reduction of research diversity); Merges & Nelson, *Complex Economics*, *supra* note 8, at 908-09 (arguing that technological advance in "cumulative technology" industries is accelerated by competition between multiple agents engaged in rivalrous research, rather than a single monopolist controlling a broad technological field).

130 Later in his life, Schumpeter abandoned the entrepreneurial figure and identified the corporate R&D laboratory as the prime engine of technological advance. This shift spawned what is generally known as the "Schumpeterian hypothesis", according to which innovation is best conducted in concentrated markets in which firms are shielded from short-term market pressures. This thesis is obviously the converse of Schumpeter's earlier "creative destruction" thesis. For the later Schumpeterian thesis, see JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY 131-34 (5th ed. 1976) [hereinafter SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY]. For a review of the ensuing academic debate concerning, and attempts to test, this thesis empirically, see P.A. GEROSKI, MARKET DYNAMICS AND ENTRY 214-29 (1991)

resources are likely to be concentrated in developing incremental improvements that add little to the existing knowledge base.

This thesis regarding the differential innovative capacities of large-firm incumbents and small-firm entrants is by no means uncontroversial or intended to be comprehensively true. It is easy to think of corporate giants that consistently introduce a fresh stream of innovative products. Notwithstanding such anecdotal evidence to the contrary, the industrial economics literature has devoted significant effort attempting to assess through econometric and case-study evidence whether any correlation exists between market concentration and innovative output and firm size and innovative output. Although there is noticeable disagreement on many issues[131] and some inter-industry variation in research results,[132] the preponderance of the evidence arguably suggests that large firms

---

[hereinafter GEROSKI, MARKET DYNAMICS]; MORTON I. KAMIEN & NANCY L. SCHWARTZ, MARKET STRUCTURE AND INNOVATION 22-48 (1982); Franco Malerba & Luigi Orsenigo, *Schumpeterian patterns of innovation*, *in* TECHNOLOGY, GLOBALISATION AND ECONOMIC PERFORMANCE 241-44 (Daniele Archibugi & Jonathan Michie eds., 1997); *see also* NELSON & WINTER, *supra* note 8, at 275-351; F. M. Scherer, *Changing Perspectives on Firm Size*, *in* INNOVATION AND TECHNOLOGICAL CHANGE: AN INTERNATIONAL COMPARISON, *supra* note 118, at 24-38 [hereinafter Scherer, *Changing Perspectives on Firm Size*]. For an interesting argument that Schumpeter's later view can be reconciled with his earlier emphasis on the entrepreneurial figure, see Richard J. McNulty, *On Firm Size and Innovation in the Schumpeterian System*, 8 J. ECON. ISSUES 627 (1974); Cohen & Klepper, *A Reprise of Size and R&D*, *supra* note 46.

131 Some studies have shown that firm size positively correlates with R&D investment, other studies have shown that R&D investment is at its highest levels among very small and very large firms and then falls among intermediate firms, and yet other studies controlling for other possible influencing factors found no significant correlation. The smaller literature that examines any possible correlation between market concentration and innovative output has similarly failed to reach a definitive consensus. While earlier studies tended to find a positive correlation between market concentration and R&D intensity, later studies have contested this conclusion on the ground that such earlier studies failed to (1) consider that R&D intensity may not be a good measure of innovative output; and (2) control for the variable degree of appropriability across industries. *See* GEROSKI, MARKET DYNAMICS, *supra* note 130, at 6-7; Baker, *supra* note 112, at nn. 88-89. For descriptions of this contradictory literature and, in most cases, presentation of independent findings, see GEROSKI, MARKET DYNAMICS, *supra* note 130, at 1-12; Acs & Audretsch, *Innovation and Technological Change*, *supra* note 118, at 13-15; Zoltan J. Acs & David B. Audretsch, *R&D, Firm Size and Innovative Activity*, *in* INNOVATION AND TECHNOLOGICAL CHANGE: AN INTERNATIONAL COMPARISON, *supra* note 118 [hereinafter Acs & Audretsch, *R&D, Firm Size and Innovative Activity*]; Zoltan J. Acs & David B. Audretsch, *Innovation and Size at the Firm Level*, 57 S. ECON. J. 739 (1991) [hereinafter Acs & Audretsch, *Innovation and Size*]; Cohen & Klepper, *A Reprise of Size and R&D*, *supra* note 46; Cohen & Klepper, *Firm Size Versus Diversity*, *supra* note 129; Wesley M. Cohen & Steven Klepper, *Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D*, 78 REV. ECON. & STAT. 232 (1996) [hereinafter Cohen & Klepper, *Firm Size and the Nature of Innovation*]; Wesley M. Cohen, et al., *Firm Size and R&D Intensity: A Re-Examination*, 35 J. IND. ECON. 543 (1987); Richard C. Levin et al., *R&D Appropriability, Opportunity and Market Structure: New Evidence on Some Schumpeterian Hypotheses*, 75 AMER. ECON. REV. 20 (1985).

132 *See* GEROSKI, MARKET DYNAMICS, *supra* note 130, at 220-22; *see also* K. Pavitt et al., *The Size Distribution of Innovating Firms in the UK: 1945-1983*, 35 J. IND. ECON. 297, 301-02 (1987) (finding that in certain industrial sectors in the U.K., smaller firms account for the majority of innovative output, while the opposite is the case in other industrial sectors).

tend to bias their portfolios toward incremental innovations and small firms tend to concentrate on more radical innovations. [133] These empirical results can be further grounded in theoretical arguments regarding the relative advantages and disadvantages of conducting R&D in small-firm and large-firm organizational settings. Together, these empirical data and theoretical arguments offer a compelling basis for believing that small firms are a crucial source of fundamental innovation investments and therefore, the patent system rests on a strong justification to the extent that it is an effective means for facilitating such investment.

1. Empirical Support

As stated above, a large amount of empirical literature has focused on testing the relative innovative competencies of large and small firms.[134] While disagreement persists,[135] there are certain well-supported trends showing that, at least in the U.S. market, small firms appear to be more inclined to conduct the most high-risk forms of fundamental product development and are therefore likely to be a central catalyst in realigning the aggregate industry R&D portfolio. Researchers have found that R&D investment does not increase more than proportionately with firm size[136] and that large firms do not engage in greater levels of R&D intensity (measured as the ratio of R&D expenditures to sales revenues) than small firms.[137] Moreover, small firms tend both to be more efficient innovators (as measured by the number of innovations per dollar of R&D or per employee)[138] and to

---

133 For a similar assessment of the literature, see Cohen & Klepper, *A Reprise of Size and R&D*, *supra* note 46, at 925.

134 *See supra* note 128 and accompanying text.

135 *See id.*

136 *See* GEROSKI, MARKET DYNAMICS, *supra* note 130, at 220-22; Cohen & Klepper, *A Reprise of Size and R&D*, *supra* note 46, at 928.

137 *See supra* note 134.

138 *See id.*; *see also* Acs & Audretsch, *R&D, Firm Size and Innovative Activity*, *supra* note 131 (studying sample of firms responsible for ninety-five percent of corporate R&D expenditures in the U.S. and finding that the productivity of R&D falls along with firm size); Mark Dodgson & Roy Rothwell, *Innovation and Size of Firm*, *in* THE HANDBOOK OF INDUSTRIAL INNOVATION, *supra* note 126, at 318 (noting that studies have shown that the relative R&D efficiency of small firms in the United Kingdom during 1969-80, as measured by the ratio between R&D expenditures and innovative output, is considerably higher than that of larger firms, and noting similar data in studies examining U.S. firms); Maria Brouwer, *Firm Size and Efficiency in Innovation: Comment on Van Dijk et al.*, 11 SMALL BUS. ECON. 391 (1998) (noting that empirical research shows generally that R&D does not increase more than proportionately with firm size and that small firms are much more innovative than large firms (as measured by innovation per employee), especially in highly innovative industries); Scherer, *Firm Size*, *supra* note 117 (finding that firms with sales of less than fifty-five million dollars in 1955 accounted for approximately 43.9% of patents issued to manufacturing corporations and noting that this result

account for a disproportionate share of innovations (especially, significant innovations)[139] in many industries.[140] Finally, numerous studies find that large firms bias their R&D portfolio toward minor or process innovations and rely on small firms for basic ideas that can serve as the basis for product improvements, while small firms tend to concentrate R&D investment on more drastic product innovations.[141] Interestingly, although the largest firms account for a disproportionately large share of the basic research investments in most industries, they tend to carry out a disproportionately small share of the high-risk, breakthrough R&D aimed at developing entirely new products and process innovations.[142]

---

suggests that smaller firms may conduct more cost-efficient R&D projects given that these smaller firms account for only fourteen percent of U.S. R&D expenditures). It should be further noted that estimates of small-firm innovative output are likely to understate such output since small firms tend to be underreported in official statistics. *See* GEROSKI, MARKET DYNAMICS, *supra* note 130, at 11.

139 *See* Pavitt et al., *supra* note 132, at 307-08 (finding that firms with fewer than 1000 employees accounted for only 3.3% of R&D expenditures in the U.K. in 1975, but 34.9% of the significant innovations in the U.K. between 1970 and 1979).

140 *See* Acs & Audretsch, *Innovation and Technological Change*, *supra* note 118, at 10-11; John H. Barton, *Antitrust Treatment of Oligopolies with Mutually Blocking Patent Portfolios*, 69 ANTITRUST L.J. 851, 864 (2002); Brouwer, *supra* note 138; GEROSKI, MARKET DYNAMICS, *supra* note 130, at 220-22; TAYLOR & SILBERSTON, *supra* note 74, at 324-26; *see also* Acs & Audretsch, *Innovation in Large and Small Firms*, *supra* note 77, at 678-88. The authors examine data collected by Small Business Administration on over 8000 innovations introduced to the U.S. market in 1982 (based on trade publication entries). The data show that (1) small firms (defined as firms employing less than 500 workers) produced almost half of all innovations in that period; (2) market concentration is *negatively* correlated with innovation activity; (3) large-firm composition is *positively* correlated with the total number of innovations; and (4) small firms represent a disproportionate amount of innovative activity even in markets dominated by large firms.

141 *See* GEROSKI, MARKET DYNAMICS, *supra* note 130, at 220-22; *see also* Acs & Audretsch, *Innovation as a Means of Entry*, *supra* note 112, at 234-35 (noting case-study evidence showing that in the video display terminal industry, large established firms tend to invest R&D resources in minor incremental innovations while small firms undertake most of the R&D investment directed at developing major product innovations); AREEDA & KAPLOW, *supra* note 118, at 32 (noting that "independent researchers, small businesses and new entrants of all types originate a disproportionate share of new ideas"); Cohen & Klepper, *Firm Size and the Nature of Innovation*, *supra* note 129, at 232-33 (using a dataset consisting of sales data for 587 business units in thirty-six manufacturing industries for the period 1974-76, finding that the share of process (rather than product) R&D undertaken by firms rises with firm size within the majority of industries, and arguing that large firms tend to allocate relatively more resources to process and more incremental R&D because of the larger output over which they can apply the results, and therefore spread the costs of their R&D); GEROSKI, MARKET DYNAMICS, *supra* note 130, at 21-22 (using data concerning 4,378 innovations in the U.K. during the period 1945-83 and finding that small firms and new entrants contribute a large percentage of major innovations). *But see* KAMIEN & SCHWARTZ, *supra* note 130, at 49-50 (noting some of the methodological difficulties that may undercut the results of studies showing large-firm bias toward incremental research).

142 *See* Edwin Mansfield, *Some Empirical Findings*, *in* R&D PATENTS & PRODUCTIVITY 128-29, 150-51 (Zvi Griliches ed., 1987). Mansfield reports findings based on data obtained from 108 firms accounting for roughly one-half of all industrial R&D expenditures in the U.S. As Mansfield notes, basic research is not necessarily directed at developing new products and processes. *See id.*; *see also* F. M. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC

2. Theoretical Support

These empirical findings are highly suggestive but certainly not dispositive. Given the absence of overwhelming empirical evidence, it is worthwhile considering the commonly posed theoretical argument (most famously set forth by the "later" Schumpeter)[143] that large firms offer the best setting in which to conduct R&D because large firms are able to offset the costs and risks of R&D against plentiful cash income from existing innovations, economies of scale, and reduced sensitivity to competitive pressures.[144] Re-applying and elaborating upon certain arguments I have presented elsewhere,[145] I present below rebuttals to each of these supposed virtues of large firms. I show why: (1) large-firm managers are unlikely to make use of their plentiful cash reserves to fund risky R&D projects; (2) large-firm managers are especially susceptible to competitive pressures (and, in particular in the case of a publicly traded firm, market pressures to maintain a stable stock price) in selecting R&D projects; and (3) the large-number organization of incumbents (responsible for their economies of scale) introduce bureaucratic constraints that are likely to stifle fundamental innovation projects. Together with the empirical evidence described above, these arguments supply strong ground for believing that policy interventions that encourage small firms to enter into technology markets dominated by incumbents can play an important role in rebalancing the aggregate industry R&D portfolio to include a significant amount of fundamental innovation projects, thereby positively contributing to social welfare in the form of accelerated technological advance.

a. Innovation Incentives

As Kenneth Arrow first suggested in positing the "replacement effect," a large firm that has a dominant market share may be reluctant to undertake or accelerate development projects that may generate radical innovations that could cannibalize the existing profit stream of

---

PERFORMANCE 437-38 (1980) (arguing that there is abundant case-study evidence suggesting that "potential new entrants play a crucial role in stimulating technical progress, both as direct sources of innovation and as spurs to existing industry members . . . new entrants contribute a disproportionately high share of all really revolutionary new industrial products and processes").

143 *See* SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY, *supra* note 130.

144 On the debate concerning the possible efficiency virtues of R&D activity within large firms, see Scherer, *Changing Perspectives on Firm Size*, *supra* note 130; Cohen & Klepper, *Firm Size Versus Diversity*, *supra* note 129.

145 *See* Barnett, *supra* note 43, at 1021-28.

current products.[146] Instead, it will prefer to devote R&D resources to the less risky development of incremental innovations that complement its existing and profitable product line.[147] Relative to a potential entrant, a patent monopolist that already extracts monopoly profits from an existing product may have lower incentives to introduce a fundamental innovation that will partially or entirely displace that product. Whereas the potential entrant expects to gain the entire profit stream from the proposed innovation, the dominant firm expects to gain the new profit stream *minus* the displaced profit stream from certain existing products. This result could be made even stronger (and more realistic, incidentally) if we take into account the significant transition costs that only an incumbent must incur—for example, liquidating production facilities and discarding organizational routines dedicated to the old product line and constructing production facilities and formulating organizational routines dedicated to the new product line.[148]

---

[146] *See* Arrow, *supra* note 15; *see also* ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 240 (arguing that large firms may sometimes fail to develop significant discoveries, especially where such discoveries are perceived as "threatening a firm's core business"); Barton, *supra* note 140, at 864 (noting Arrow's hypothesis and stating that this thesis implies that the propensity of a firm to invest in research will arguably increase with the number of firms in the relevant market); GEROSKI, MARKET DYNAMICS, *supra* note 130, at 148-49 (stating that, while large incumbents have the resources to quickly exploit new innovations, "innovative activity is often rent-displacing and this dulls the incentives of such firms to innovate"); KAMIEN & SCHWARTZ, *supra* note 130, at 110 (stating that firm managers may face conflicting motives in contemplating developing a product innovation, since the "carrot" of extraordinary profits from the proposed new innovation conflicts with the "stick" of existing profits on the existing product line); Lerner & Merges, *Patent Scope*, *supra* note 7, at 17 (stating that a pharmaceutical manufacturer may not want to develop therapeutic products for a disease where it has an existing product for fear that the new, superior product will cannibalize existing sales); Jennifer F. Reinganum, *The Timing of Innovation: Research, Development, and Diffusion*, *in* THE HANDBOOK OF INDUSTRIAL ORGANIZATION, *supra* note 126, at 851 [hereinafter Reinganum, *Timing of Innovation*] (stating that "[w]hen innovation is uncertain, a firm which currently enjoys a large market share will invest at a lower rate than a potential entrant, for an innovation which promises the winner a larger share of the market"); Michael L. Katz & Carl Shapiro, *R&D Rivalry with Licensing or Imitation*, 77 AM. ECON. REV. 402 (1987) (arguing that post-development dissemination of an invention, either through licensing or imitation, will cause non-innovating firms to benefit from the invention, and that, consequently, dominant firms will tend to develop minor innovations and will only develop major innovations if imitation is difficult). *But see* Philippe Aghion & Jean Tirole, *The Management of Innovation*, 109 Q. J. ECON. 1185, 1204-05 n.26 (1994) (noting that, although Arrow's "replacement effect" suggests that the incumbent has reduced incentives to innovate, the incumbent may have increased incentives to the extent that the incumbent gains more from remaining a monopoly than an entrant from becoming a duopolist).

[147] *See* Rosen, *supra* note 128, at 422-23.

[148] On this point generally, see Baker, *supra* note 112, at 636-38. In an interesting case study, Baker argues that the "cannibalism" effect (and, in particular, high transition costs) could explain in part the decision by the Big Three U.S. automakers to accommodate entry by Japanese manufacturers into the U.S. compact car market in the 1970s (rather than seeking to deter entry by developing a competing, high-quality compact car product). A further illustration is found in the alleged reluctance of the Bell telephone monopoly in the 1920s and 1930s to adopt the combined handset and dial system and automatic toll boards until many years after these products were developed due to the preexisting investments in older and incompatible technologies. *See*

Given the foregoing, the incumbent firm may find that no or postponed development of new products is its profit-maximizing strategy, depending on the expected profits from the new innovation, the current profits from the existing innovation, and most importantly, the potential for preemptive development by known or unknown rivals. Placing great weight on the last condition, later commentators have correctly noted that this argument assumes (or is strongest if it assumes) that entry is entirely blocked. Where this assumption is relaxed (that is, entry is open to challengers), the incumbent faces the loss of its entire product line even if it selects a no-innovation strategy, since rivals may introduce a radical "replacement" innovation. Where this threat is credible, the incumbent firm may pursue an alternative strategy of "preemptive patenting" or "preemptive publication."[149] When potential challengers threaten to bring a drastic innovation to the market, a monopolist may elect to develop the drastic innovation prior to any challenger, patent it or publish it (so as to prevent its being patented by competitors), and then suppress or postpone the innovation's market introduction until the monopolist exhausts profits from the existing product line.[150] Of course, where competitors can invent around the

---

FLOYD L. VAUGHAN, THE UNITED STATES PATENT SYSTEM 233-35 (1956).

149 *See* Richard J. Gilbert & David M. G. Newberry, *Preemptive Patenting and the Persistence of Monopoly*, 72 AMER. ECON. REV. 514, 517-19 (1982); *see also* Reinganum, *supra* note 146, at 869-75 (assuming that entry into the market is open and still finding that incumbents should rationally invest less in producing drastic innovations than potential entrants, largely due to the fact that incumbents must deplete current cash flow to fund such innovation (in contrast to potential entrants, which have no profit stream to forfeit)). On preemptive publication in particular, see Douglas Lichtman et al., *Strategic Disclosure in the Patent System*, 53 VAND. L. REV. 2175 (2000); Gideon Parchomovsky, *Perish or Publish*, 98 MICH. L. REV. 926 (2000).

150 Admittedly, there are few definitively proven examples of such technological suppression by an incumbent. However, the history of innovation includes several instances where incumbent firms possibly sought to defend their monopoly share by suppressing new innovations that could displace existing products and trigger a competitive market. *See* Richard Dunford, *The Suppression of Technology as a Strategy for Controlling Resource Dependence*, 32 ADMIN. SCI. Q. 512, 514-16 (1987) (citing historical evidence that AT&T may have delayed the introduction of the automatic telephone because of "patent consolidation" concerns, and that General Electric delayed the introduction of fluorescent lighting because it wished first to saturate the market for incandescent lighting); VAUGHAN, *supra* note 148, at 233-36 (arguing that the Bell telephone monopoly controlled a large percentage of patents in the relevant market and failed to develop many of the patented technologies in order to protect its market share, and that General Electric cooperated with the electric utilities in order to delay the introduction of fluorescent lamps, which would have used less electricity than existing incandescent lamps). Another good illustration of possible technological suppression is found in the antitrust case law in the facts of *McDonald v. Johnson & Johnson*, 722 F.2d 1370 (8th Cir. 1983). The plaintiffs had developed an electronic pain control device and then sold it to Johnson & Johnson after having been given assurances that Johnson & Johnson would develop and promote the technology. *See id.* at 1372. The plaintiffs alleged that Johnson & Johnson effectively suppressed the acquired technology because it posed a threat to its over-the-counter and prescription drug business. *See id.* at 1372-73. Although the district court upheld the plaintiffs' claim, the Eighth Circuit Court of Appeals vacated the antitrust claims due to lack of standing, proximate causation, and lack of cognizable injury under the antitrust laws (largely because the plaintiffs had entered into the buy-out agreement). *See id.*

relevant patent (or, in the case of preemptive publication, can overcome the incumbent's first-mover advantage without the aid of patent protection) and/or the project development costs are especially high or uncertain, the expected net benefits of such preemptive product development may be meager or zero.[151]

Putting together the foregoing theoretical positions, a revised view might hold that a large dominant firm has lower incentives either to develop or, even if developed, put on the market, breakthrough innovations that may render obsolete an existing and profitable product line. By postponing the development or introduction of a radical product innovation, a dominant firm (especially if it faces few potential entry threats) maximizes the total profit stream it can gain from its existing products *plus* the new innovation (and *minus* the opportunity cost of delaying the new innovation).

b. Agency Costs

As has been demonstrated on innumerable occasions, incentive problems or "agency costs" arise from the separation of ownership and management within a large, publicly traded firm.[152] There are several reasons to believe that this laundry list of agency costs should include the tendency of risk-averse managers to prefer lower-return, lower-risk research projects over higher-return, higher-risk projects. Large-firm managers have relatively weak incentives to endorse a risky R&D project since they expect to share in a small portion of the monetary gains if the project succeeds but a large portion of the reputational cost, up to and including job termination, if the project fails.[153] Empirical evidence showing the extremely skewed distribution of returns from R&D projects, which generally consist of a handful of "big winner" projects and a multitude of net losses,[154] underscores the very

---

at 1373.

151 *See* Gilbert & Newberry, *supra* note 149, at 522-24.

152 The agency cost literature is enormous. For the classic sources, see Eugene M. Fama, *Agency Problems and the Theory of the Firm*, 88 J. POL. ECON. 288 (1980); Michael C. Jensen & William H. Meckling, *The Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305 (1976).

153 *See* SHARON M. OSTER, MODERN COMPETITION ANALYSIS 303 (2d ed. 1994); *see also* David Hirshleifer & Yoon Suh, *Risk, Managerial Efforts, and Project Choice*, 2 J. FIN. INTERMEDIATION 308, 308-09 (1992) (stating that "it is harder to motivate a manager to make the right decisions when projects differ in risk" and that "managers will avoid projects that are subject to early and conspicuous failure in order to maintain their reputations as good judges of project quality"); Rosen, *supra* note 128, at 421-22 (noting that there is evidence that in many large organizations, "managers are punished for bad outcomes more than they are rewarded for good outcomes").

154 *See* BRONWYN H. HALL ET AL., MARKET VALUE AND PATENT CITATIONS: A FIRST LOOK (Nat'l Bureau of Econ. Research, Working Paper No. 7741, 2000); Pakes, *Patents as Options*,

significant technical and commercial uncertainties that burden any long-term R&D project. Even if a large firm has access to plentiful capital reserves to fund fundamental research projects (which is thought to be one of the large-firm attributes conducive to R&D), its risk-averse managers may prefer incremental innovations that improve cash flow and share prices in the short term, rather than big-ticket projects that deplete cash flow and have a highly uncertain and deferred return.[155] This logic results in the "fast second" strategy described previously,[156] where the large-firm manager elects to have the potential entrant assume the high risks and costs of developing an uncertain product innovation and expects to be able to imitate the challenger's product if it turns out to be successful.[157] Moreover, if the firm is publicly traded and its managers cannot credibly signal to the market that reduced cash earnings are due to the diversion of cash to long-term R&D projects, rather than to mismanagement or other adverse events, then managers will have limited incentives to divert cash flow even to risky projects that have a higher net present value than safer projects with large near-

---

supra note 84, at 779; JEAN O. LANJOUW ET AL., HOW TO COUNT PATENTS AND VALUE INTELLECTUAL PROPERTY: USES OF PATENT RENEWAL AND APPLICATION DATA 2, 13-14 (Nat'l Bureau of Econ. Research, Working Paper No. 5741, 1996); Lanjouw, *Patent Protection in the Shadow of Infringement*, *supra* note 84, at 695. For detailed findings, see Schankerman, *supra* note 84, at 80, 93-94 (finding that, based on dataset covering nearly all patents applied for in France by French, German, U.K., Japanese, and U.S. applicants during the period 1969-82 in pharmaceuticals, chemicals, mechanical devices and electronics, the distribution of the private value of patent rights is highly skewed—e.g., the top one percent of patents accounts for twelve percent and fourteen percent of the total private value of patent rights in pharmaceuticals and chemicals, respectively, and twenty-one percent and twenty-four percent for mechanical and electronics patents, and the top five percent accounted for thirty-four percent and forty-eight percent of total value in pharmaceuticals and chemicals, and fifty percent and fifty-five percent for mechanical and electronics patents); Henry J. Grabowski & John Vernon, *A New Look at the Returns and Risks to Pharmaceutical R&D*, 36 MGMT. SCI. 804 (1990) (showing that, for approximately 100 new pharmaceutical products developed and introduced in the U.S. during the 1970s, the ten most profitable drugs contributed fifty-five percent of the "quasi-rents" from all ninety-nine new drugs and, in the aggregate, the "winners" yielded net returns just slightly above the losses generated by the "losers").

155 On the relationship between agency costs, the choice of R&D projects, and investment horizons, see Bengt Holmstrom, *Agency Costs and Innovation*, *in* THE MARKETS FOR INNOVATION, OWNERSHIP AND CONTROL 131, 131-53 (Richard H. Day et al. eds., 1993); *see also* Baker, *supra* note 112, at 638 (stating that risk-averse managers may reject R&D projects even if the proposed innovation has a greater expected return than alternative, safer investments); TAYLOR & SILBERSTON, *supra* note 74, at 30 (arguing that the available evidence suggests that large industrial firms tend to avoid R&D projects with a "pay-off period" of much more than five years and generally prefer a much more rapid pay-off period, and therefore, tend to prefer minor innovation projects having short development times). Note additionally that, relative to plant and equipment investments, R&D projects may be costlier to finance since secrecy considerations may require the use of exclusively internal financing, thus directly depleting the firm's current cash holdings. *See* Uri Ben-Zion, *R&D, Investment Decision, and the Firm's Market Value*, *in* R&D, PATENTS AND PRODUCTIVITY, at 301 (Zvi Griliches ed., 1984).

156 *See supra* note 122 and accompanying text.

157 *See* Baker, *supra* note 112, at 638-39.

term cash flows.[158] Where this signaling failure exists, then even entering into a risky long-term R&D project (not to mention visible failure of a research project after-the-fact) may injure the ability of the firm to access the capital markets for external financing (and the manager's compensation to the extent it is in the form of stock options).[159]

There are several reasons why small firms probably do not suffer substantially from these excessively risk-averse managerial tendencies. First, most smaller firms are not publicly traded and therefore not dependent on the equity markets for external financing and not subject to internal and external pressures from the trading community to maintain the firm share price.[160] Moreover, a small-firm manager's stock options will tend to be "out of the money" and, unlike large-firm managers, the only possible means of bringing the options "into the money" lies in undertaking long-term R&D projects with significant possible returns. More generally speaking, a small-firm manager recognizes that the firm has few growth opportunities other than through significant investment in developing fundamental innovations. Second, ownership and management are often not separated. Where this is the case, the owners/managers can expect to pocket a sizable portion of the gains from a successful innovation and thus have greater incentives to select high-risk/high-return R&D projects. Thus, unlike large-firm managers, small-firm managers are exposed to significant portions of not only the downside risk but also the upside gain of a research project. Third, small firms are often parties to financing agreements with venture capital firms, which must meet high "hurdle

158 *See* Clas Wihlborg, *On Business Myopia and Market Organization*, *in* THE MARKETS FOR INNOVATION, OWNERSHIP AND CONTROL, *supra* note 155, at 121-23. Note that this argument does not need to assume an informationally inefficient market or a market unable to distinguish between cash loss due to current business failures and cash loss due to long-term R&D investment. Even if the market can distinguish between high-quality and low-quality long-term R&D projects, this still may not correct the managerial bias toward short-term projects if managers' compensation is determined at least partially on the basis of earnings per quarter or even per year (or there is some other reason—other than sensitivity to share price—why managers prefer short-term over long-term projects), in which case the market will anticipate the resulting managerial short-term-bias and will therefore attribute any decrease in short-term cash flow to adverse performance rather than cash investments in long-term R&D. *See* Holmstrom, *supra* note 155, at 150-51; Giovanni Dosi, *Finance, Innovation and Industrial Change*, *in* THE MARKETS FOR INNOVATION, OWNERSHIP AND CONTROL, *supra* note 155, at 275-76. At the extreme, if the market cannot distinguish between reductions in current earnings due to adverse performance and those due to long-term R&D investment, then any large-scale diversion of cash resources to even a high-quality, long-term R&D project could depress the firm share price to such a low level as to invite takeover bids, thus threatening management's job security. *See id.* at 277.

159 *See* Holmstrom, *supra* note 155, at 132-33 (stating that concerns for reputation in the capital markets lead large firms to act cautiously in selecting investment projects, since poor performance can constrain the future availability of affordable capital).

160 *See id.* at 151.

rates" (target rates of return) for their external investors and therefore are likely to use their contractual control rights over the management of the corporation to counteract any risk-averse managerial tendencies in the selection of R&D projects.[161] Finally, smaller firms are likely to be less concerned about the reputational damage (and resulting limits on access to external financing) as the result of the failure of a risky R&D project since, unlike their large-firm counterparts, they have not yet accumulated significant reputational capital in the financing market and therefore, have little to lose in this respect.[162]

c. Organizational Limitations

The organizational theory literature suggests that large-firm entities may often lack the structural competence to develop fundamental product innovations. Some theories of organization view the firm as a collection of actors who rationally join together to maximize revenues by apportioning numerous tasks among agents with specialized skill sets.[163] Because each agent performs a single task repeatedly, it can execute that task more efficiently than an agent who operates alone and is responsible for performing several tasks. There is a tradeoff, however, between the returns to specialization achieved by cooperative

---

[161] *See* Anand & Galetovic, *supra* note 35, at 616-24. The authors make the interesting argument that venture capitalists can also more easily commit to sharing positive cash flow with researchers, thereby increasing researchers' expected upside percentage in the case of innovation success. By contrast, large-firm management cannot credibly commit to providing its researchers with any such cash-flow allocation to the extent that the firm either cannot easily keep track of, or can partially misrepresent, the costs and revenues of an R&D project. Although it is efficient for a large firm to integrate its accounts and operations, this means that it may not be able to treat the R&D project as a stand-alone vehicle for accounting purposes. This also means that the firm can at least partially (and, from the researcher's perspective, self-interestedly and unverifiably) shift costs and revenues over its many divisions. *See id.* at 616-17, 623-24; *see also* OLIVER E. WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM: FIRMS, MARKETS, RELATIONAL CONTRACTING 41-42 (1985) (stating that innovation incentives may be weak in a large firm because it is difficult to design a compensation scheme that reflects accurately responsibility for innovative output).

[162] *See* Wihlborg, *supra* note 158, at 125; Holmstrom, *supra* note 155, at 151; *see also* ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 244-45 (noting argument in literature that smaller organizations with fewer fixed assets at stake will be willing to bear greater risk than larger organizations with greater fixed assets at stake).

[163] *See* Patrick Bolton & Mathias Dewatripont, *The Firm as a Communication Network*, 109 Q. J. ECON. 809, 810 (1994). For other treatments of information processing capabilities, organizational structure and innovation competencies, see Stephen J. DeCanio & William E. Watkins, *Information Processing and Organizational Structure*, 36 J. ECON. BEHAV. & ORG. 275 (1998); David J. Teece, *Firm Organization, Industrial Structure, and Technological Innovation*, 31 J. ECON. BEHAV. & ORG. 193, 199 (1996). For a more general treatment of the relationship between production incentives and organizational choices, see Armen A. Alchian & Harold Demsetz, *Production, Information Costs, and Economic Organization*, 62 AMER. ECON. REV. 777 (1972).

action and the costs of communication within a large group.[164] As a firm grows in size, it must construct a bureaucracy that distributes information among a large number of agents and must institute a multidivisional hierarchy that monitors the performance of each of those agents.[165] Although large firms may enjoy impressive scale economies in distribution and manufacture (which, again, are thought to be a large-firm attribute conducive to R&D), the associated set of bureaucratic routines may hamper these firms from achieving creative insights or investing rapidly in radical product innovations.[166] A hierarchical structure that channels information and responsibilities through numerous agents according to a set of rigid procedures may diminish the speed with which the firm reacts to information about changed market conditions and contemplates new product concepts.[167] Even

---

164 *See* Bolton & Dewatripont, *supra* note 163, at 811.

165 *See* DeCanio & Watkins, *supra* note 163, at 287-88 (stating that a firm's choice of organizational structure often rests on a tradeoff between multiplying managerial levels to reduce informational overload and generating frictions that may result from an excess of hierarchical channels); *see also* WILLIAMSON, *supra* note 161, at 200-01 (stating that "if any one manager can deal directly with only a limited number of subordinates, then increasing firm size necessarily entails adding hierarchical levels"). For a more general discussion of the relation between innovation and relatively hierarchical or non-hierarchical forms of organization, see Mark Dodgson, *Technological Collaboration and Innovation*, *in* THE HANDBOOK OF INDUSTRIAL INNOVATION, *supra* note 126, at 285-92; Gerard Fairtlough, *Innovation and Organization*, *in* THE HANDBOOK OF INDUSTRIAL INNOVATION, *supra* note 126, at 325-35.

166 *See* ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 240 (noting literature arguing that large firms "with their established routines and structures" may be better suited "to making incremental improvements of existing technologies . . . than for making new discoveries"); WILLIAMSON, *supra* note 161, at 200-01 (stating that "since innovation . . . tends to be untidy, innovation—which is a poorly structured, high-risk activity—may not be an activity which the large, mature bureaucracy is constitutionally well-suited to handle"); *see also* Rebecca M. Henderson & Kim B. Clark, *Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms*, 9 ADMIN. SCI. Q. 9, 9-10, 12-13 (1990) (noting that incremental innovation reinforces the capabilities of established firms while radical innovation creates difficulties because it destroys the usefulness of existing procedures and forces firms to draw on new technical skills and adopt new approaches); Rebecca Henderson, *Underinvestment and Incompetence as Responses to Radical Innovation: Evidence from the Photolithographic Alignment Equipment Industry*, 24 RAND J. ECON. 248, 251-52, 260-62 (1993).

167 *See* ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 240 (arguing that "large firms may be particularly unsuited for radical or breakthrough innovations" and even when they make a significant discovery, they "may not recognize or nurture it adequately, especially if the discovery is not perceived as relating to the firm's core operations and markets"); DeCanio & Watkins, *supra* note 163, at 290 (stating that "an organization's size may constrain its speed in adopting innovation"); Henderson, *supra* note 166, at 251-52 (noting that the standardized informational routines adopted by larger firms in order to reduce routine information processing costs may prevent the firm from easily digesting unorthodox types of information, thereby impeding the firm's ability to undertake radical innovation projects); Thomas M. Jorde & David J. Teece, *Innovation and Cooperation: Implications for Competition and Antitrust*, 4 J. ECON. PERSP. 75, 84 (1990) (stating that "one property of large integrated structures is that they have the potential to become excessively hierarchical and less responsive to market needs"); Teece, *supra* note 163, at 201 (stating that innovation is often badly served by bureaucratic structures since the "new and the radical will almost always appear threatening to some constituents" of the

among those divisions of an incumbent firm devoted to product development, researchers' efforts may be governed by research protocols that tend to focus on improvements to existing products rather than on radical innovations possibly leading to entirely new products.[168]

Small firms do not face the large-numbers problem and therefore do not need to construct hierarchical communication and monitoring structures that may inhibit creative research and adaptive flexibility. Evidencing this claim, smaller, research-intensive firms tend to exhibit shallow hierarchies, low degrees of compartmentalization, and a higher concentration of decision-making authority in the founding individuals.[169] Further evidence is found in the fact that large firms, in order to emulate the more adaptive structure of small-firm organizations, have sometimes spun off their research divisions into separate entities or split themselves up into separate competing entities, each devoted to a different line of research.[170] Moreover, because small firms have not developed and adopted elaborate information processing and other bureaucratic structures, they do not need to incur the significant costs and risks, as well as overcoming internal political opposition, which attend the process of liquidating portions of these structures in order to undertake a radical innovation project.[171] For the foregoing reasons, smaller firms, relative to more established competitors, may be more inclined and better suited to undertake fundamental innovation activities.

---

representative structure); James M. Utterback, *The Dynamics of Product and Process Innovation in Industry*, *in* JAMES UTTERBACK, MASTERING THE DYNAMICS OF INNOVATION 54-55 (1979) (noting that established market leaders tend to have difficulty making the shift to major innovations that would render obsolete, and disrupt existing investments in, their existing product line).

168 *See* Henderson & Clark, *supra* note 166, at 9 (stating that "radical innovation often creates great difficulties for established firms"); *id.* at 14-15 (stating that established organizations focus on improving the components of a "dominant design" rather than in "architectural" innovations that could lead to an entirely different dominant design). Interestingly, historical experience shows that when small, entrepreneurial Silicon Valley firms have acquired large, vertically integrated firms, the organizational controls of the large target corporation have tended to undermine the entrepreneurial capacities of the smaller acquiror company. *See* Teece, *supra* note 163, at 212 n.32.

169 *See* Teece, *supra* note 163, at 213.

170 *See* ARORA ET AL., MARKETS FOR TECHNOLOGY, *supra* note 35, at 240-41, 244-45 (noting that large organizations sometimes seek to encourage "exploration" of new products and technologies by assigning such development to a separate, "spun off" entity); Scherer, *Changing Perspectives on Firm Size*, *supra* note 130, at 30 (noting decision of Cray Research Corporation to split itself into two separate and competing companies, each of which is devoted to a separate line of research).

171 *See* Henderson, *supra* note 166, at 252; Henderson & Clark, *supra* note 166, at 17-18; *see also* Michael L. Tushman & Phillip Anderson, *Technological Discontinuities and Organizational Environments*, 31 ADMIN. SCI. Q. 439, 444-46 (1986) (stating that radical innovations render obsolete existing technological competencies and enable new firms founded to exploit the new technology to gain market share at the expense of incumbents that, due to "traditions, sunk costs, and internal political constraints", are still bound to the outmoded technology).

C. *The Incumbent Patenting Objection*

Just as the conventional incentives thesis encountered an empirical obstacle, so too does the alternative market-entry thesis. Based on a large body of empirical research, the market-entry thesis proposes that patents probably only offer significant marginal value as an exclusionary device to small and outsider firms that lack access to more effective private appropriation mechanisms. As a positive account of patenting practice, this thesis would then predict that only or principally small and outsider firms would invest resources in obtaining patents. However, this is not at all the case. Large, incumbent firms routinely patent and are the leaders in terms of the number of patents issued annually.[172] The existence of widespread incumbent patenting raises doubts as to whether the market-entry thesis offers a plausible account of patenting behavior and, by implication, whether it offers a reliable basis for a normative discussion of patent policy. In response (and keeping in mind the demonstrated ineffectiveness of patent protection), I propose below several possible explanations for this anomalous result and explore the normative implications of each such explanation. As I argue below, the market-entry thesis more than survives both as a compelling positive and normative account of patent protection in light of the incumbent patenting phenomenon.

1. Entry Barriers

Incumbent patenting may reflect incumbents' desire to erect a barrier against potential entrants by assembling a wide patent portfolio that discourages these competitors from investing R&D resources in the

---

[172] *See* Sabra Chartrand, *I.B.M. Again Leads Businesses in Patents Awarded by U.S.*, N.Y. TIMES, Jan. 13, 2003 (stating that for the tenth year in a row, the U.S. PTO awarded more patents to inventors at I.B.M. than any other company); Long, *supra* note 11, at 642 n.50 (stating that the top ten patentees (by volume) in the U.S. in 2000 were I.B.M., NEC, Canon, Samsung, Lucent, Sony, Micron, Toshiba, Motorola and Fujitsu); Allison & Lemley, *supra* note 69, at 2128-29 (reviewing sample of 1,000 patents issued during 1996-98 and finding that 70.7% of the patents were issued to large entities and 29.3% to small entities, of which 17.5% were individuals, 1.1% were non-profits and 10.7% were small businesses). Further but more indirect evidence of the widespread extent of incumbent patenting is provided by the finding that patent counts, when weighted by subsequent citation frequency, are correlated with the market value of both larger and smaller firms in certain industries. For a review of the econometric literature that seeks to assess the informativeness of patent counts as an indicator of innovation investment, output and productivity, see HALL ET AL., *supra* note 154, at 6. If appropriately weighted patent counts are apparently a fairly reliable indicator of firm value (and assuming the market is not systematically overestimating or otherwise misconstruing the informational value of patent holdings), then the publicly traded firms in the industries for which this proposition holds must be patenting their innovative output on a fairly regular basis.

same or related technological fields.[173] There is evidence showing that this strategy has some effect: In the biotechnology industry, larger firms' patent portfolios apparently influence the research programs of smaller firms, such that smaller firms tend to pursue research programs in sectors that are not already "crowded" with other patented innovations.[174] This explanation suggests that much large-firm patenting is socially inefficient[175] since, given that incumbents already have access to ample non-patent appropriation mechanisms, an expanded patent portfolio reduces potential entrants' access to the existing knowledge base without any offsetting increase in ex ante innovation incentives. It is worth noting that this risk may be mitigated to the extent that the possibility of using patents to entrench an established market position may enhance incentives for small-firm entry by increasing the maximum expected reward obtainable if the potential entrant successfully displaces the existing incumbent.[176]

2. Arms Race or Knowledge Exchange

Incumbent patenting may indicate that incumbent firms are locked into an arms race in which a patent portfolio is necessary to compete and maintain research access against competitors who are similarly accumulating patents. From a social efficiency perspective, this scenario is normatively ambiguous. On the one hand, this explanation

---

[173] *See* Bessen, *Patent Thickets*, *supra* note 97, at 4. The antitrust case law contains some instances of antitrust litigations designed partly to disrupt such patent accumulation strategies. In the *General Electric* case, the district court held that General Electric's acquisition of the basic patents controlling the tungsten filament incandescent lamp, and its decision to condition licenses to other lampmakers on a grant-back clause relating to the development of new technology, constituted a violation of the ban on attempts to monopolize set forth in Section 2 of the Sherman Act. *See* U.S. v. General Electric, 82 F. Supp. 753, 901 (D.N.J. 1949). *But see* SCM Corp. v. Xerox Corp., 645 F.2d 1195, 1208-09 (2d Cir. 1981) (finding that Xerox's acquisition of patents covering plain paper copying technology, and its refusal to license the patents to competitors, did not constitute a violation of Section 2 of the Sherman Act because (1) Xerox's development of the technology preceded its acquisition of the patents; and (2) even a monopolist is permitted under the patent laws not to license its patents).

[174] *See* Lerner, *Patenting in the Shadow of Competitors*, *supra* note 69, at 471-72. To be precise, Lerner examines the patenting behavior of firms with low litigation costs and firms with high litigation costs. However, this low cost/high cost distinction generally tracks the small firm/large firm distinction since smaller, younger firms tend to have higher litigation costs, given that for purposes of the analysis Lerner assumes these costs to be inversely correlated to cash resources and previous litigation experience.

[175] James Bessen appears to be of this view, arguing that "much, if not most" patenting can best be explained as "strategic patenting"—that is, patents that are taken out not in order to protect an innovation against appropriation, but in order to create a "patent thicket" by which the holder of the patent portfolio can extract license payments from potential competitors, deter entry from potential competitors, or hold up innovation by potential competitors. *See* Bessen, *supra* note 34, at 4-6.

[176] I am grateful to Ben Zipursky for this observation.

suggests that much large-firm patenting is socially inefficient insofar as large firms appear to suffer from a collective action problem in which each firm is worse off because no firm can credibly agree not to patent. If patents offer no appropriability advantage and each firm is patenting for purely or primarily strategic reasons (e.g., to block competitors or to prevent being blocked by competitors), then large-firm patenting comes close to a case of pure rent-seeking with no net social benefits. On the other hand, patents may facilitate inter-firm knowledge-sharing arrangements insofar as many competitors may only be willing to enter into such cooperative arrangements to the extent that their technological contributions to the joint venture are covered by reasonably secure legal entitlements.[177] The semiconductor industry offers a good example. Reflecting the cumulative and systemic nature of semiconductor technology, much product development in this industry has historically taken place pursuant to broad cross-licenses that cover whole portfolios of patents related to an entire technical field and, in some cases, covering future as well as existing patents.[178] Patents may make all firms in the semiconductor industry better off to the extent that (1) each participant in a cross-licensing arrangement requires a significant patent portfolio to trade against that of its competitors, and (2) without patent protection, no potential participant would be prepared to reveal valuable intellectual goods, which is a precondition to negotiating and entering into such mutually beneficial knowledge exchanges.[179] Put differently, in the absence of patent protection, competing firms in the semiconductor industry would be less likely to exchange information

---

[177] On this point, see NAOMI R. LAMOREAUX & KENNETH L. SOKOLOFF, INVENTIVE ACTIVITY AND THE MARKET FOR TECHNOLOGY IN THE UNITED STATES, 1840-1920 (Nat'l Bureau of Econ. Research, Working Paper No. 7107, 1999) (arguing that the surge in patenting and inventive activity in the mid to late-nineteenth century can be attributed in large part to the fact that patent system enabled inventors to extract value by trading their patented technologies, thereby providing an incentive to invest resources in invention in the first place).

[178] For a description of cross-licensing practices in the semiconductor industry, see Hall & Ziedonis, *supra* note 86, at 110; Grindley & Teece, *supra* note 102.

[179] Several studies attest that larger, more established participants in the semiconductor industry view patents principally as a bargaining chip to be used in cross-licensing negotiations. *See* ASHISH ET AL., *supra* note 35; Hall & Ziedonis, *supra* note 86, at 110; Grindley & Teece, *supra* note 102. For an interesting theoretical challenge to the notion that, in the absence of intellectual property rights, inventors will be reluctant to enter into negotiations with manufacturers (or, by extension, we might say, co-venturers) that may be able to produce and market the invention on a large scale, see James J. Anton & Dennis A. Yao, *Expropriation and Inventions: Appropriable Rents in the Absence of Property Rights*, 84 AMER. ECON. REV. 190 (1994). The authors argue that, even where intellectual property rights are nonexistent, an innovator could either (1) not reveal the innovation and then enter into a contingent contract with a potential buyer of the innovation that would enable the innovator to appropriate a significant portion of an innovation's value (if the innovation proved successful); or (2) reveal the innovation prior to contracting and still appropriate a significant portion of the invention's value by making a credible threat to reveal the innovation to a competing firm. *Id.* There is no empirical evidence to suggest that any such contracts exist in practice.

through cooperative ventures and product development would slow as a result.

The obvious retort to this argument is that patent portfolios should be unnecessary for this purpose given the fact that most potential co-venturers holding a large patent portfolio are established firms and therefore, following this Article's thesis, should have access to a large set of informal appropriation mechanisms. That is, even if patent protection were not available, these firms would still enter into similar knowledge-sharing cooperative ventures and product development rates would be roughly comparable. There are two reasons why this may not be the case. First, while extralegal devices may prevent competitors from replicating the tacit manufacturing, marketing and other knowledge that lies behind an innovator's product, no such protection is available where an innovator reveals such intimate technical knowledge through a cooperative venture. Second, a non-trivial but less than overwhelming level of patent protection may be necessary not only to facilitate such knowledge-exchange ventures but to induce their formation. Whereas a zero level of patent protection will prevent the formation of such ventures due to appropriability hazards, a strong level of patent protection will similarly prevent their formation by eliminating patent holders' incentives to incur the disclosure costs attendant upon entering into such cooperative projects.[180] Where a firm holds a patent securely controlling a broad technological territory, it will face a much lower risk of unauthorized third-party appropriation and therefore will have significantly weaker incentives to enter into cooperative ventures with, and thereby reveal proprietary information to, potential competitors. Such conduct may be privately beneficial for the patent holder but socially undesirable for the market as a whole. In many cumulative technology industries, such as biotechnology, semiconductors, and information technology, cooperative knowledge exchanges are crucial to, and the central form of organization for, technological advance.[181] Thus, to the extent that patents are an important means of both facilitating and encouraging the formation of these information-exchange vehicles, they likely speed up the pace of technological advance and almost certainly make an important contribution to social welfare.

---

[180] For an application of this argument to patent policy in the biotechnology sector, see Barnett, *supra* note 43, at 1011-12, 1028-31.

[181] With respect to semiconductors, see Hall & Ziedonis, *supra* note 86; Grindley & Teece, *supra* note 102. With respect to biotechnology, see *supra* notes 47 and 59 and accompanying text. With respect to information technology, see BESSEN & MASKIN, SEQUENTIAL INNOVATION, *supra* note 11, at 16-17.

3. Insiders/Outsiders

Large firms may sometimes act as "outsiders" when entering a new market where they will not have a first-mover advantage or any of the related appropriation capabilities and therefore require patent protection in order to compete with the incumbent. This will be true where the large-firm's reputational capital and technical expertise have little applicability to the new market into which the large firm wishes to enter. Just as a patent enables a small entrant to enter into a concentrated market dominated by incumbents, a patent should enable a large entrant to do so where the entrant's existing reputational capital and distribution and manufacturing efficiencies in other markets have limited or no applicability in the target market. This explanation suggests that some large-firm patenting may be socially efficient to the extent that it enables firms that are established in a particular market to challenge the dominant position of incumbent firms in other markets.

IV. IMPLICATIONS AND EXTENSIONS

In this Part, I consider some of the preliminary policy implications of the market-entry theory of patent protection and explore possible extensions of this theory to copyright in particular, and to intellectual property theory in general.

A. *Implications for Patent Issuance*

The market-entry thesis proposes that, without patents, entrants may have difficulty overcoming the natural competitive advantages enjoyed by first-mover firms, which have access to a wide range of effective non-patent appropriation devices. As its most direct policy implication, this thesis would appear to recommend that the patent office only grant patents to small and outsider firms. Taking the market-entry thesis to its logical conclusion, any firm that applied for patent protection and fell into the incumbent category probably would not require patent protection in order to recoup costs on its R&D project and therefore, granting patents to such a firm probably would not pass a social cost-benefit test.

Several compelling concerns of overinclusion and underinclusion counsel against such a proposal. First, it is not clear that it would be feasible to designate a standard that could distinguish between entrants and incumbents. As the antitrust literature can amply attest, a firm that

has captured less than a five-percent market-share may or may not be well on its way to penetrating a concentrated market and therefore may or may not have little need for patent protection. As a result, any market share threshold would almost inevitably be set arbitrarily.[182] Second, it is not clear that it would be feasible to formulate a standard that could distinguish between large "insider" firms attempting to enter a related market and large "outsider" firms attempting to enter an unrelated market. Only in the latter instance is the firm deserving of patent protection since in the former case it is reasonable to presume that the insider firm could use its existing reputational capital, manufacturing know-how and distribution contacts in order to penetrate related markets. But, as the antitrust literature can again attest,[183] a market definition standard would again most likely be set arbitrarily. Either a market share standard or market definition standard that is arbitrarily set risks granting patent protection to firms that *do not* require such protection for market-entry or denying patent protection to firms that *do* require such protection for market-entry. Third, given the likely arbitrary manner in which any discriminatory standard would be applied with respect to patent issuance, any attempt to exclude larger firms as ineligible applicants for patent protection could strongly discourage smaller firms from relying on patent protection (since they may very quickly be caught in the net of ineligible larger firms) in determining whether or not to contest an incumbent's market share. Fourth, even if the discriminatory standard could be applied consistently and confidently, it is not clear that patent protection should be denied to more established firms that have access to alternative appropriation mechanisms. As discussed above, even incumbent firms may require some level of patent protection in order to be able and willing to enter into mutually beneficial knowledge-sharing arrangements with their competitors, which are often crucial to market-entry and/or technological advance in some of the most fertile technology markets.[184] Thus, even if it were possible to distinguish between entrants and incumbents, and insider and outsider large firms, it still may not be desirable to deny patent protection to incumbents and insider large

---

182 On some of the difficulties encountered in the antitrust context with respect to market share definition and the determination of supracompetitive pricing power, see AREEDA & KAPLOW, *supra* note 118, at 571 (noting that "[i]t cannot be emphasized too strongly that market definition and the defendant's market share give, at best, only a suggestion of defendant's market power" and that the "boundaries of any product and geographic market are necessarily imprecise").

183 *See id.*

184 *See supra* notes 177-78 and accompanying text. Additionally, as noted elsewhere in this Article, even where incumbent firms patent primarily in order to inhibit entry by third parties, such patenting behavior could have some offsetting efficiency benefits by increasing the expected reward for entrants that successfully unseat an incumbent (in which case, the entrants could similarly seek to inhibit entry through patent accumulation), thereby increasing entrants' ex ante innovation incentives. *See supra* note 173 and accompanying text.

firms, which are likely to take part in these cooperative ventures.

Finally, a broad denial of patent protection to larger firms (even if feasible) could provide perverse incentives to opportunistic individuals and small entities to accumulate patent rights providing little in the way of significant technological knowledge. These individuals and entities could then demand buyout or royalty payments from larger firms unable to avail themselves of the bargaining leverage afforded such individuals and small entities by a differentially applied patent statute.[185]

B. *Implications for Antitrust Policy*

There is an obvious role for antitrust regulators to play in restraining the ability of incumbents to accumulate broad patent portfolios solely or primarily for the purpose of excluding potential entrants. Judicial precedent exists for bringing antitrust actions against firms that use patent portfolios or patent pools to prevent entry.[186] However, regulators must walk a fine line between ensuring that large firms do not assemble broad patent portfolios that preclude entry (or at least do not use those portfolios to preclude entry) and facilitating the formation of inter-firm ventures that spread the cost and uncertainty of R&D projects. As I have discussed in greater detail elsewhere,[187] recent antitrust policy (or more specifically, antitrust policy during the Clinton Administration) has aimed to trace that fine line by pursuing two divergent policies. On the one hand, Congress,[188] the agencies,[189] and

---

185 I am grateful to Mark Lemley for raising this issue.

186 *See, e.g.*, Zenith Corp. v. Hazeltine, 395 U.S. 100 (1969) (finding a Sherman Act § 1 violation by a group of firms that participated in a patent pool used to ensure that manufacturing was carried out in specific foreign nations); U.S. v. Singer Mfg., 374 U.S. 174 (1963) (finding that a cross-licensing agreement between two holders of blocking patents in various countries for sewing machine designs concealed an illegal conspiracy to acquire a broad patent portfolio so as to institute infringement litigation against the remaining global competitor); Hartford Empire Co. v. U.S., 323 U.S. 386 (1945) (upholding district court finding that the defendant, the leading U.S. manufacturer of glass-making machinery, conspired with smaller such manufacturers to acquire thousands of patents in order to enforce cross-licensing restrictions, production quotas, and territorial allocations that discouraged entry); Kobe, Inc. v. Dempsey Pump Co., 198 F.2d 416 (10th Cir. 1952) (finding that aggressive strategy of acquiring all patents for hydraulic pumps for oil wells and instituting frivolous infringement litigation against new entrants constituted an illegal attempt to monopolize). For an extensive discussion of the possibility of finding an antitrust violation when a monopolist denies or oligopolists deny entrants access to proprietary knowledge, see Barton, *supra* note 140, at 875-82.

187 *See* Barnett, *supra* note 43, at 1033-54.

188 In 1984, Congress passed the National Cooperative Research Act, which offered special treatment to registered research joint ventures, and in 1993, it enacted the National Cooperative Research and Production Act, which extended relief granted under the 1984 Act to registered production joint ventures.

189 For the enforcement agency guidelines regarding joint ventures, see U.S. Dep't of Justice, Antitrust Enforcement Guidelines for International Operations § 3.4 (1988), *reprinted in* 4 TRADE

the courts[190] have provided relaxed antitrust treatment for research and production joint ventures.[191] This generous stance encourages holders of complementary patents, blocking patents, or efficient distribution capacities to merge research, development, production, or diffusion processes to spread the cost of undertaking fundamental innovation projects. On the other hand, the Federal Trade Commission has applied unconventional compulsory licensing and asset divestiture remedies as an approval condition for several mergers during the 1990s in high-technology (and especially, pharmaceutical) markets.[192] These remedies have been grounded in a novel enforcement theory of protecting competition (and, consequently, research diversity) in "innovation markets"[193] and have successfully sustained multiple research lines and access to the industrial knowledge base in various technology-dependent industries.

C. *Extension to Copyright*

It is certainly apposite to consider whether this Article's critique of the traditional incentive theory of patent protection, and the proposed market-entry theory of patent protection, could have any application in copyright theory and policy. Possibly even more so than innovators of

---

REG. REP. (CCH) 13, 109.

[190] It is true that the Supreme Court has never overruled its widely criticized decision in *United States v. Topco Associates*, 405 U.S. 596 (1972), which applied a rule of *per se* illegality to a territorial restraint imposed by a joint venture that lacked dominant market share. But the Court's later decision in *Arizona v. Maricopa County Medical Society*, 457 U.S. 332 (1982), as well as numerous lower courts' decisions, suggest that the *per se* rule no longer applies to joint ventures that show significant degrees of economic integration and exhibit any plausible welfare-enhancing effects.

[191] For discussions of antitrust treatment of joint ventures, see Howard H. Chang, David S. Evans & Richard Schmalensee, *Some Economic Principles for Guiding Antitrust Policy Towards Joint Ventures*, 1998 COLUM. BUS. L. REV. 223 (1998); Herbert Hovenkamp, *Exclusive Joint Ventures and Antitrust Policy*, 1995 COLUM. BUS. L. REV. 1, 13 (1995); Joseph Kattan, *Antitrust Analysis of Technology Joint Ventures: Allocative Efficiency and the Rewards of Innovation*, 61 ANTITRUST L.J. 937 (1993); Gregory J. Werden, *Antitrust Analysis of Joint Ventures: An Overview*, 66 ANTITRUST L.J. 701 (1998).

[192] For a detailed discussion, see Barnett, *supra* note 43, at 1048-51.

[193] This method of analysis is presented in the 1995 DOJ Intellectual Property Guidelines. *See* U.S. Dep't of Justice & fed. Trade Comm'n Antitrust Enforcement Guidelines for the Licensing of Intellectual Property §§ 3.2.2, 3.2.3 (1995). The Guidelines define an innovation market as "the research and development directed to particular new or improved goods or processes, and the close substitutes for that research and development." *See id.* at § 3.2.3. For a general presentation of the innovation market concept by two former DOJ staff members who originally conceived this approach, see Richard J. Gilbert & Steven C. Sunshine, *Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets*, 63 ANTITRUST L.J. 569 (1995). For a full discussion (both "for" and "against") of the innovation-markets approach, see *Symposium: A Critical Appraisal of the "Innovation Market" Approach*, 64 ANTITRUST L.J. 1 (1995).

patentable goods, producers and distributors of copyrightable goods have easy access to many informal appropriation mechanisms, which obviously casts doubt on the traditional incentive theory behind copyright protection. In a well-known article, Professor and now-Justice Stephen Breyer argued that the existence of private appropriation mechanisms, such as various lead-time advantages, updating services and advance subscription methods, raised the possibility that copyright protection may not even be necessary to induce private investment in the "heartland" of copyright—namely, book publishing.[194] Whether or not this argument is correct with respect to book publishing,[195] its logic gains much greater currency with the advent of sophisticated encryption and other "digital rights management" technologies that permit the online delivery of literary and other expressive materials governed by highly tailored distribution formats and fee schedules.[196] Despite this fact, it remains unclear whether net exclusion costs for copyright distributors have fallen in light of technological advances that greatly lower reproduction costs for copyright consumers such that unauthorized, perfect reproductions can be easily made by users lacking anything but the most basic technical knowledge. It remains similarly unclear whether such alternative appropriation technologies are equally available to the full range of market participants, as we have seen is generally not the case among industries eligible for patent protection. Although there has been a

---

[194] Specifically, Breyer argued that copyright protection may not be necessary to induce investments in book publishing because (1) publishers could enter into advance subscription agreements with consumers and retailers (and publishers could promise updates to dedicated customers); (2) publishers could adopt secrecy practices that would enable them to enjoy a "lead time" in which no competing copies would be on the market; and (3) publishers could retaliate commercially against piracy by releasing "fighting editions" at below-cost prices. *See* Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281 (1970) [hereinafter Breyer, *Uneasy Case*]. Breyer's skepticism regarding the necessity of copyright protection to induce creative expression has recently manifested itself in his dissent in the Supreme Court's seven to two decision to uphold Congress' decision to extend the copyright term. *See* Eldred v. Ashcroft, 537 U.S. 186 (2003) (Breyer, J., dissenting). In an excellent but far-less-discussed article, Robert Hurt similarly contests (without entirely rejecting) the incentive rationale for copyright protection, arguing that (1) authors often invest resources in creating expressive works for non-monetary reasons; (2) private patronage could provide authors with monetary support without incurring the social costs attendant to granting copyrights; and (3) publishers have devices other than copyright by which to appropriate returns, such as lead-time advantages, the threat of retributive action against copiers, prepublication orders and special editions "authorized" by the writer. *See* Robert M. Hurt & Robert M. Schuchman, *The Economic Rationale of Copyright*, 56 AMER. ECON. REV. 421 (1966).

[195] For a critique of Breyer's argument and Breyer's rejoinder, see Barry W. Tyerman, *The Economic Rationale for Copyright for Published Books: A Reply to Professor Breyer*, 18 UCLA L. REV. 1100 (1971); Stephen Breyer, *Copyright: A Rejoinder*, 20 UCLA L. REV. 75 (1972).

[196] For a recent expansion of Breyer's thesis, which takes into account later technological developments, see MARK S. NADEL, QUESTIONING THE ECONOMIC JUSTIFICATION FOR COPYRIGHT (AEI-Brookings Joint Ctr. for Regulatory Studies Recent Publication 03-1, Aug. 2003).

profusion of scholarly commentary regarding digital rights management technologies and the claimed "death" of copyright,[197] it remains to my knowledge an unresolved and largely unexplored empirical question whether any significant percentage of producers and distributors of literary, visual or artistic material relies on copyright protection in electing to invest resources in either the production or distribution of such material. This is obviously a ripe subject for further study, the results of which would significantly increase the confidence with which policy recommendations can be made in this sector.

D. *Integrating Intellectual Property Analysis*

In a recent article, Gideon Parchomovsky and Peter Siegelman argued that intellectual property theory should be more completely "integrated" by viewing the effects of a single body of intellectual property law with other related bodies of intellectual property law (for example, by viewing patent protection in conjunction with the effects of trademark and copyright protection).[198] That Article's thesis suggests that intellectual property scholarship should also be more completely "integrated" in a vertical sense—that is, by viewing the effects of formal intellectual property protection in conjunction with the effects of informal protection for the relevant intellectual goods (for example, the effects of patent protection in conjunction with the effects of existing extralegal appropriation mechanisms). At the same time, this Article also integrates horizontally to some extent insofar as the first-mover advantage and, in particular, branding strategies, are dependent on trademark protection and, probably to a lesser extent, trade secret protection. An even more completely integrated analysis of formal intellectual property protection would seek to assess the effectiveness and cost of formal and informal intellectual property protection relative to alternative governmental actions, such as tax subsidies, grants, and government laboratories, and relative to private patronage mechanisms.

Any "integration analysis" of intellectual property protection—whether an analysis that integrates other bodies of intellectual property law, an analysis that integrates formal and informal forms of appropriation devices or an analysis that integrates government and charitable subsidies, grants and direct investment—greatly complicates the policymaker's task. Where such an analysis is feasible to any meaningful extent, however, these additional policymaking costs should be worth incurring in order to improve the probability of successful

197 *See supra* note 4.

198 *See* Parchomovsky, *supra* note 149.

policy prescriptions. An example might illustrate this point. Failing to undertake a vertical integration analysis and therefore failing to recognize the existence of market-generated alternatives to legally supplied forms of intellectual property protection could lead to recommendations for levels of patent protection that are either excessive or entirely unnecessary from a social cost-benefit analysis. To take one of many possible real-world examples (given the recent expansion of the category of patentable subject matter),[199] and without expressing any definitive view on the matter, calls on U.S. lawmakers to imitate the European Union's generous intellectual property protections for electronic databases may be misguided given that the U.S. database industry apparently flourishes even though formal intellectual property protection of database products is highly uncertain.[200] Given the apparent ease with which database products could be copied, conventional incentives theory would predict that supplying or strengthening formal property rights in such databases should provide otherwise absent incentives for investment in generating these products. But the existence of a flourishing database market in the absence of any secure, formal protection suggests that private appropriation mechanisms already exist whereby firms are able to capture significant innovation proceeds. If this is so, and if it is also true that private appropriation mechanisms are widely available to both incumbents and entrants,[201] then supplying or strengthening formal protections for database products may likely fail to generate social benefits in excess of the significant administrative, litigation and other social costs that

---

199 *See supra* note 90 and accompanying text.

200 *See* JUSTIN HUGHES, POLITICAL ECONOMIES OF HARMONIZATION: DATABASE PROTECTION AND INFORMATION PATENTS 5 (Cardozo L. School, Research Paper Series No. 47, Aug. 2002), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=318486# PaperDownload (last visited Mar. 20, 2003). Much of the uncertainty in intellectual property protections for databases derives from the *Feist* decision, which held that databases and directories were generally not eligible for copyright protection due to lack of originality. *See* Feist Publ'ns v. Rural Tel. Serv. Co., Inc., 499 U.S. 340 (1991) (ruling that unless it featured an original selection or arrangement of the factual data, the telephone directory could not be protected by copyright law because it lacks a minimal degree of creativity). Although *Feist* held that the creative arrangement and selection of facts in a directory or database could be eligible for copyright protection, this qualification has had little meaningful effect as courts since *Feist* have consistently refused to recognize copyright infringement even where third parties have copied all or large portions of a database, absent a claim that creative selection and arrangement were copied. *See* Hughes, *supra*, at 18.

201 It is not clear that this condition is fulfilled in the database market, given that one of the primary opponents of extending patent protection to this market is the Bloomberg Corporation, one of the leading incumbents. *See* HUGHES, *supra* note 200, at 41. Consistent with the market-entry thesis proposed in this Article, an incumbent such as Bloomberg should oppose the introduction of patent protection since patent protection will improve entrants' ability to contest incumbent market positions founded upon the use of non-patent appropriation devices. Certainly, there are other plausible (but not necessarily incompatible) explanations for Bloomberg's opposition to strengthening intellectual property protection for databases—chiefly, Bloomberg's reliance as a content provider on access to other directories and databases.

would be incurred by introducing such protection.[202]

CONCLUSION

The twin assumptions that the market suffers from a lack of incentives to invest in innovation projects and that patent rights play a principal role in supplying such missing incentives underlie much of governmental policy and academic discussion concerning the patent system. These assumptions presumably underlie the widely noted strengthening, and increased availability, of patent protection, which has arisen as a result of several significant legislative and judicial steps: (1) the creation of the Federal Circuit in 1982, the tendency of the Federal Circuit to uphold a greater percentage of litigated patents relative to previous courts (largely by more liberally applying the doctrine of equivalents),[203] and the judicial extension of patent protection to business methods[204] and certain genetically engineered organisms,[205] (2)

---

[202] It is worth noting an important objection that may be raised to this proposition. It may be argued that, even where (1) the market apparently supplies adequate devices for protecting intellectual goods in a particular industry (e.g., information databases); and (2) such devices are available to both incumbents and entrants, the state should still offer protection for such intellectual goods, and even encourage private parties to abandon private protection devices for such state-provided protection, to the extent that public protection generates lower net social costs than its private equivalent. In particular, patent protection may generate lower net social costs to the extent that it either (1) generates lower administrative and litigation costs relative to the various "business" and other costs that must be incurred by firms in establishing non-patent barriers to unauthorized imitation; and (2) encourages disclosure of existing innovations and therefore improves the dissemination of scientific and industrial knowledge. While the former possibility seems fairly implausible, the latter does not and raises a difficult question that certainly requires further study. However, even in the absence of any reasonably confident empirical basis on which to act, it would appear to be more prudent for the state to refrain from subjecting an industry to the enormous litigation and administrative costs of the patent system if the relevant market (1) already encourages significant innovation investment among its existing participants; and (2) does not erect serious barriers to entry (or at least, to making credible entry threats) by potential additional participants. Moreover, it is worth noting that it is unclear whether the patent system offers significantly greater disclosure than exists in an industry that is not eligible for patent protection. This is because first, it is unclear whether patent disclosures generally meet the statutory requirement under 35 U.S.C. § 112 of describing the "best mode" of making the patented invention and providing sufficiently detailed written description so as to "enable" other persons in the relevant field to make and use the invention without "undue experimentation." *See* Lemley, *Rational Ignorance*, *supra* note 69. Second, there is evidence that in many industries it is generally difficult to prevent information leakage regarding product innovations, whether patented or unpatented (which explains why most innovations, whether patented or unpatented, are imitated within several years of their market introduction, see *supra* notes 81-82 and accompanying text).

[203] The doctrine of equivalents extends the effective scope of the litigated patent to trivial improvements to the patented innovation and thereby seeks to prevent third parties from easily detouring around a patent. For the historically leading case, see *Graver Tank & Mfg. Co. v. Linde Air Products Co.*, 339 U.S. 605 (1950). For a more recent leading case, see *Warner-Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17 (1997).

[204] *See* State St. Bank v. Signature Fin. Group, 149 F.3d 1368 (Fed. Cir. 1998).

the legislative extension of the patent term for pharmaceutical and chemical products in order to adjust for the delays caused by the FDA approval process, pursuant to the Patent Term Extension Act of 1984,[206] and (3) the Bayh-Dole Act of 1980,[207] which permits and encourages universities and other institutions to patent discoveries that were made using federal research funds. These assumptions also underlie the U.S. government's position in recent international trade negotiations, which places a strong emphasis on inducing developing countries, and has successfully persuaded many such countries, to increase their intellectual property protection to levels closer to that of the U.S. system.[208] Finally, these assumptions underlie much scholarly discussion in the economic and legal literature, which often impossibly assumes either a "winner-take-all" patent race model in which patents confer perfect appropriability[209] or that R&D-intensive firms are finely responsive to subtle changes in the extent of patent protection or in the design of the patent prosecution or litigation process.[210]

This Article, and the extensive body of empirical evidence upon which it relies, cast serious doubt on these assumptions, which in turn suggests that current approaches to innovation policy in complex technology markets may be excessively patent-centric. The market often does not suffer from a lack of incentives to invest in innovation activities and, when it does, patents generally are of little help. Ronald

---

205 *See* Diamond v. Chakrabarty, 447 U.S. 303 (1980).

206 *See* Drug Price Competition and Patent Term Restoration Act of 1984, Pub. L. No. 98-417, 98 Stat. 1585 (1984) (codified at 21 U.S.C. §§ 156, 355 (1994 & Supp. IV 1998)).

207 *See* Bayh-Dole Act of 1980, Pub. L. No. 96-517, 94 Stat. 3015-3028 (codified as amended at 35 U.S.C. §§ 301-307 (1994 & Supp. IV 1998)).

208 For a description of these developments with respect to pharmaceuticals in particular, see JEAN O. LANJOUW & IAIN COCKBURN, DO PATENTS MATTER? EMPIRICAL EVIDENCE AFTER GATT (Nat'l Bureau of Econ. Research, Working Paper No. 7495, 2000).

209 *See, e.g.*, Edmund W. Kitch, *Elementary and Persistent Errors in the Economic Analysis of Intellectual Property*, 53 VAND. L. REV. 1727, 1729-38 (2000) (arguing that it is commonly and mistakenly believed that patents are monopolies when this is rarely the case since a patent's claims usually do not cover all of an economically relevant market); Partha Dasgupta, *Patents, Priority and Imitation or, the Economics of Races and Waiting Games*, 98 ECON. J. 66, 70 (1988) (stating that patents transform the competitive process among researchers into a race the aim of which is to award the entire reward to the winning research team). *See also* KAMIEN & SCHWARTZ, *supra* note 130, at 105 (stating that "[i]n the modern theories of market structure and innovation, competition among potential innovators is thought of as a race to be first"); John H. Barton, *Patents and Antitrust: A Rethinking in Light of Patent Breadth and Sequential Innovation*, 65 ANTITRUST L.J. 449, 454-55 (1997) (noting substantial discussion in the economics literature of patent races); Robert P. Merges & Richard R. Nelson, *On Limiting or Encouraging Rivalry in Technical Progress: The Effect of Patent Scope Decisions*, 25 J. ECON. BEHAVIOR & ORG. 1, 8-9 (1994) (stating that economists sometimes use the race model to analyze invention incentives); Manfredi La Manna et al., *The Case for Permissive Patents*, 33 EUR. ECON. REV. 1427, 1427-28 (1989) (stating that "formal models of patent races typically assume that the winner takes all").

210 *See* Mariko Sakakibara & Lee Branstetter, *Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms*, 31 RAND J. ECON. 77 (2001).

Coase showed that the textbook example of a non-excludable public good, the lighthouse, was actually privately funded in nineteenth-century England.[211] Similarly, available evidence shows that, with the exception of a few significant industries, much innovative output is closer to a private good and does not suffer from any significant excludability problem. Given this crucial fact, there is good reason for policymakers and government decision makers to pay close attention to the effectiveness and availability of non-patent and informal appropriation mechanisms. This is not to impugn the necessity on incentive grounds for fairly robust levels of patent protection in industries such as pharmaceuticals and chemicals in which they appear to offer effective and apparently crucial protection.[212] Setting aside these select technology settings in which product development tends to be highly discrete (rather than highly cumulative, as is generally the case in today's most technologically fertile sectors),[213] the traditional incentives theory cannot supply either a positive account of patenting behavior or a normative basis upon which to engage in patent policy discussions. In its place, this Article proposes that patents operate principally as an important means by which small and outsider firms can overcome the potent, informal appropriation mechanisms available principally to incumbents. Secondarily, in cumulative technology markets, patents probably provide a useful currency with which established firms, and sometimes less-established firms, can enter into cost-sharing and knowledge-sharing ventures that accelerate product development and improve informational dissemination. This market-entry theory not only works well as a descriptive account of the patent

---

211 *See* Ronald H. Coase, *The Lighthouse in Economics*, 17 J.L. & ECON. 357 (1974). Scholars have identified a handful of other goods or services that would seem to be "obvious" public goods but turn out, on closer examination, to exhibit many attributes of a private good. *See, e.g.*, Michael R. Montgomery & Richard Bean, *Market Failure, Government Failure, and the Private Supply of Public Goods: The Case of Climate-controlled Walkway Networks*, 99 PUB. CHOICE 403 (1999) (arguing that the existence of several "walkway networks" (underground and aboveground bridge and tunnel networks for pedestrian traffics to avoid extreme heat and cold) in major U.S. and Canadian cities that are primarily or exclusively funded and managed by private entities suggests that private arrangements can be an effective means of supplying what are commonly thought to be public goods); *see also* Bruce L. Benson, *Are Public Goods Really Common Pools? Considerations of the Evolution of Policing and Highways in England*, 32 ECON. INQUIRY 249 (1994) (presenting historical evidence showing that, prior to the consolidation of power in the monarchy, policing and the maintenance of highways in Anglo-Saxon England were funded and effected primarily by private parties with no significant free-rider problem). For a more general argument that scholars working in the Pigovian tradition have tended either to exaggerate the prevalence of market failures or to assume the plausibility of market failures without empirical support, see STEVEN N. S. CHEUNG, THE MYTH OF SOCIAL COST 38-49 (1980).

212 Ironically and reflecting an almost certain policy error, currently (or until fairly recently) in many developing countries (such as India), chemicals and pharmaceuticals are or were precisely the industries in which patents are or were *not* permitted.

213 *See supra* note 99 and accompanying text.

system but, on a normative level, shows that patents play a crucial role in redirecting the tilt of the aggregate industry R&D portfolio towards fundamental innovation investments that are likely to generate the most significant contributions of social value.