Northwestern Journal of Technology and Intellectual Property

Volume 11 | Issue 5 Article 5

2013

Putting Innovation Incentives Back in the Patent-Antitrust Interface

Thomas Cheng
University of Hong Kong

N O R T H W E S T E R N
JOURNAL OF TECHNOLOGY
AND
INTELLECTUAL PROPERTY

Putting Innovation Incentives Back in the Patent-Antitrust Interface

Thomas Cheng

April 2013 VOL. 11, NO. 5

Northwestern Journal of Technology and Intellectual Property
Volume 11, Number 5 (April 2013)

Putting Innovation Incentives Back in the Patent-Antitrust Interface

By Thomas Cheng*

This Article proposes a new approach, the constrained maximization approach, to the patent-antitrust interface. It advocates a return to the utilitarian premise of the patent system, which posits that innovation incentives are preserved so long as the costs of innovation are recovered. While this premise is widely accepted, it is seldom applied by the courts in patent-antitrust cases. The result is that courts and commentators have been overly deferential to dynamic efficiency arguments in defense of patent exploitation practices, and have failed to scrutinize the extent to which patentee reward is genuinely essential to generating innovation incentives. Under the constrained maximization approach, the antitrust courts attempt to maximize the net social benefits of an innovation by adjusting the scope of patent exploitation, subject to the constraint that innovation costs are recouped. This approach will allow the courts to take into account two important considerations in the balance between static and dynamic efficiencies that have been largely overlooked: the contribution of cumulative innovation to social welfare and the variety of ways in which innovators recover their R&D investments in addition to patent protection. Incorporation of both of these considerations lends support to a more robust approach to the patent-antitrust interface.

* B.A. (Yale), J.D. (Harvard), B.C.L. (Oxon); Attorney & Counsellor, New York State; Associate Professor, Faculty of Law, The University of Hong Kong. The author wishes to thank Addiped Cheng, Henry Chau, and Keith Chan for their able research assistance.

INTRODUCTION

¶1 The patent-antitrust interface is probably the area of antitrust law that has received the most academic attention in recent years.[1] With its juxtaposition of patent policy with antitrust policy, it is generally perceived to be one of the most intricate areas of antitrust. In fact, to the best of this author's knowledge, it is the only area on which a separate treatise has been published.[2] To the extent that patent policy and antitrust policy conflict over the assessment of the legality of a patent exploitation practice, the courts must balance antitrust concerns about static efficiency and consumer welfare with the dynamic efficiency considerations of patent law. Any area of antitrust that requires balancing is, by nature, complicated. A balancing exercise that requires consideration of the policy of another body of law is yet more complex.

[1] For a sample of this vast body of literature, see William Baxter, *Legal Restrictions on Exploitation of the Patent Monopoly: An Economic Analysis*, 76 YALE L. J. 267 (1966); WARD BOWMAN, PATENT AND ANTITRUST LAW (1973); Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813 (1984); Willard K. Tom & Joshua A. Newberg, *Antitrust and Intellectual Property: From Separate Spheres to Unified Field*, 66 Antitrust L. J. 167 (1997); Michael A. Carrier, *Unraveling the Patent-Antitrust Paradox*, 150 U. PENN. L. REV. 761 (2002); ANTITRUST, PATENTS AND COPYRIGHT: EU AND US PERSPECTIVES (François Lévêque & Howard Shelanski eds., 2005); THE INTERFACE BETWEEN INTELLECTUAL PROPERTY RIGHTS AND COMPETITION POLICY (Steven D. Anderman ed., 2007); Daniel A. Crane, *Intellectual Liability*, 88 TEX. L. REV. 253, 254 (2009); MICHAEL A. CARRIER, INNOVATION FOR THE 21ST CENTURY (2009); Josef Drexl, *Real Knowledge is to Know the Extent of One's Own Ignorance: On the Consumer Harm Approach in Innovation-Related Competition Cases*, 76 ANTITRUST L. J. 677 (2010).

[2] HERBERT HOVENKAMP ET AL., IP AND ANTITRUST: AN ANALYSIS OF ANTITRUST PRINCIPLES APPLIED TO INTELLECTUAL PROPERTY LAW (2010).

¶2 The standard argument concerning the patent-antitrust interface is as follows: the patent system provides incentives for innovation by granting the innovator a period of exclusivity, during which he or she may be able to charge a supra-competitive price for its technology or for products incorporating this technology. The result of this exclusivity is that consumers will have to bear a higher price. It is this supernormal profit that is said to allow innovators to recover the costs of innovation. R&D costs are generally thought of as sunk costs. Without patent protection, competitors will be able to imitate the technology quickly and cheaply, driving the price to its marginal costs, which of course exclude sunk costs. Assuming that the marginal costs of production are constant,[3] the innovator will be denied recovery of its R&D costs. Innovation will hence be deterred. To remedy this unsatisfactory state of affairs, the patent system grants the innovator the right to exclude and a host of other patent rights. These rights and the patentee reward that they help to generate are crucial to attract innovators, supplying them sufficient financial incentives to pursue R&D. Since innovation is the engine of growth in the modern-day economy, innovation should be encouraged as much as possible. The common belief is that more innovation is better for society. Under this tenet, antitrust policy should defer to patent policy unless there is persuasive evidence of certain consumer harm, perhaps as in the case of price fixing.[4]

¶3 The courts often seem all too willing to accept this standard argument and uphold patent exploitation practices. This is in no small part due to a perception that dynamic efficiency arguments are difficult to verify. While the courts are ready to challenge claims of short-run consumer harm, they have largely withheld scrutiny from arguments asserting reduction of innovation incentives. When a defendant makes a credible claim that a certain patent exploitation practice is essential to generating sufficient patentee reward, which in turn is essential to securing adequate innovation incentives, the courts are generally reluctant to examine to what extent the exploitation practice at issue is genuinely essential to the recovery of innovation costs. They are equally reluctant, if not even more so, to question whether patentee reward in general is necessary for attracting innovation. Claims of the deterrence of future innovation are even less susceptible to refutation, perhaps because they are by and large empirically unverifiable. Dynamic efficiency arguments at times have come to be treated as a trump card in patent-antitrust cases.[5]

¶4 The goal of this Article is to show that for the courts to attain a more balanced view of the patent-antitrust interface, they must begin to scrutinize dynamic efficiency arguments. Claims about harm to innovation incentives should not be accepted at face value. In particular, this Article will attempt to refocus attention on the fundamental premise of the patent system that innovators will have sufficient incentives to innovate as

[3] This assumption is important because if marginal costs are rising, the innovator will be able to recover at least part of the R&D costs from the sale of the initial units. Since the equilibrium price is determined by the marginal costs of the last unit being sold, and the marginal costs of that unit will be higher than those of the infra-marginal ones under an assumption of rising marginal costs, the equilibrium price will necessarily exceed the marginal costs of the infra-marginal units, allowing the innovator to recover part of its R&D costs.

[4] In the 1926 *General Electric* case, the Supreme Court seemed to have suggested that antitrust policy yields even in the presence of such conduct. *See* United States v. Gen. Elec. Co., 272 U.S. 476, 47 S.Ct. 192 (1926).

[5] CSU v. Xerox Corp., 203 F.3d 1322 (Fed Cir. 2000); United States v. Microsoft Corp., 253 F.3d 34 (D.C. Cir. 2001).

long as their innovation costs are fully compensated. The reward offered by the patent system is not closely calibrated to match innovation costs. Hence the scope of rights permitted by the current patent system may significantly exceed that which is necessary to generate innovation.

¶5 The argument that patent rights must be respected to preserve innovation incentives is hence not universally valid. In fact, there are reasons to believe that it is not so in many instances. This is especially true once one takes into account the various alternative ways in which the innovator can appropriate the benefits of its creation without the assistance of patents. As will be explained subsequently, there are a myriad of ways that have been found to be even more effective than patent protection in securing private returns to innovation. Incorporating these observations about innovation incentives, this Article proposes a "constrained maximization" approach to the patent-antitrust interface, under which the courts attempt to maximize the net social benefits of an innovation subject to the constraint that the private benefits of innovation, which include patentee reward and returns from other means of appropriation, must cover innovation costs. Drawing on insights from this approach, this Article will then provide practical guidance on how to decide patent-antitrust cases by focusing on the various attributes of innovation.

¶6 This Article is divided into five sections. Section I provides an overview of the policy conflict underlying the patent-antitrust interface and a critical examination of some of the approaches that have been proposed to resolve it. Section II examines the theoretical justifications for patent protection and the implications they have for the patent-antitrust interface. In particular, it will be argued that the current assumption about the imperative of preserving patentee reward often leads to timid enforcement against harmful patent exploitation practices. Section III attempts to bring the theoretical discussion in Section II to a more practical level and offers concrete suggestions on how to apply the lessons drawn from the previous Section in actual cases. Section IV illustrates the application of the approach proposed in this Article by applying it to refusal to license cases. Section V concludes the Article.

I. CONFLICT BETWEEN PATENT LAW AND ANTITRUST—APPARENT OR REAL?

A. Short-Run Conflict Between Patent and Antitrust

¶7 The goal of antitrust is to promote consumer welfare by fostering competition on the merits and preventing firms from deploying anticompetitive practices to exclude rivals and inflict harm on consumers. Consumer welfare is enhanced when consumers are able to obtain the same good at a lower price or obtain a higher-quality good at the same price. Consumer welfare is also improved when consumer choice is widened. In economic parlance, antitrust is principally concerned with static efficiency—the allocation of goods and services over the short run. Dynamic efficiency, which refers to the ability of a market or an economy to produce innovation, is also important to antitrust. Meanwhile, patent law spurs invention and innovation[6] by providing creators of

[6] While invention and innovation may be often used interchangeably in common parlance, economists draw a clear distinction between the two terms. Invention refers to the creation of a new technology or product, or a substantial improvement of an existing one. Innovation refers to the development and

patentable inventions a period of exclusivity, during which the patentee has the right to exclude anyone from practicing and commercializing the invention.[7] This exclusivity is intended to allow the patentee to impose a supra-competitive price for its innovation so that it can recoup its R&D investment. Users of the innovation and the end consumers

---

commercialization of a new technology into a good or service that will redound value to consumers. *See* Stuart Macdonald, *Exploring the Hidden Costs of Patents*, *in* GLOBAL INTELLECTUAL PROPERTY RIGHTS: KNOWLEDGE, ACCESS AND DEVELOPMENT 13, 23 (Peter Drahos & Ruth Mayne eds., 2002) ("An invention is a discovery: an innovation is a product or service that is new to the market, or simply new to the adopter."); Keith E. Maskus, Sean M. Dougherty & Andrew Mertha, *Intellectual Property Rights and Economic Development in China*, *in* INTELLECTUAL PROPERTY AND DEVELOPMENT: LESSONS FROM RECENT ECONOMIC RESEARCH 295, 299 (Carsten Fink & Keith M. Maskus eds., 2005) ("*Invention* refers to the creation of new knowledge, and *innovation* (or *commercialization*) refers to the development of marketable products from that knowledge."). The differentiation of invention and innovation is said to date back to Joseph Schumpeter. RICHARD R. NELSON & SIDNEY G. WINTER, AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE 263 (1985). The two do not necessarily go hand in hand. As Schumpeter pointed out decades ago, it is possible to have innovations without a new invention, whereas invention need not be followed by innovation. JOSEPH A. SCHUMPETER, BUSINESS CYCLES : A THEORETICAL, HISTORICAL, AND STATISTICAL ANALYSIS OF THE CAPITALIST PROCESS 84-85 (1939).

One of the controversies regarding patent protection on the one hand, and invention and innovation on the other hand, is whether the patent system should aim to provide incentives only to invent, or also to innovate and commercialize inventions. Some believe that the patent system should only be concerned with inventions, while others insist that patents must provide incentives for both inventions and innovation. SUBCOMM. ON PATENTS, TRADEMARKS & COPYRIGHTS OF THE SENATE COMM. ON THE JUDICIARY, 85TH CONG. 2D SESS. AN ECONOMIC REVIEW OF THE PATENT SYSTEM 9 (Comm. Print 1958) (written by Professor Fritz Machlup) ("[i]t is invention rather than enterprising innovation which the patent system is supposed to encourage."); WILLIAM J. BAUMOL, THE FREE MARKET INNOVATION MACHINE 65-67 (2002) (using ancient China, ancient Rome and the former Soviet Union to illustrate that innovation is more important than invention in promoting economic growth); ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS: HOW OUR BROKEN PATENT SYSTEM IS ENDANGERING INNOVATION AND PROGRESS, AND WHAT TO DO ABOUT IT 43 (2004) (using Xerox's photocopying machine to illustrate that patent protection must aim to provide incentives both to invent and to innovate). Some have even argued that the patent system should only protect innovation. DIRECT PROTECTION OF INNOVATION 1-34 (William Kingston ed., 1987) (proposing property right system that comes into effect only when new product is introduced to market). There has been no definitive resolution of this debate. *See* F.M. Scherer, *The Economics of the Patent System*, *in* INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 439, 441 (F.M. Scherer ed. 1980) (hereinafter "*Economics of Patent System*").

This Article will make no attempt to resolve this debate. In keeping with the more contemporary views of commentators such as Baumol, Jaffe, and Lerner, it will proceed on the premise that patent protection provides incentives for both invention and innovation. These two terms, and the related terms of "inventors" and "innovators," will thus be used interchangeably in this Article as far as their relationship with patent. Efforts will be made to specify the sense in which the two terms are used where necessary.

[7] The right to exclude granted by the patent law to the patentee is sometimes known as patent monopoly. *See* Int'l Wood Processors v. Power Dry, Inc., 792 F.2d 416, 426 (4th Cir. 1986) ("Since patent law grants a patent holder the right to exclude others for a period of seventeen years, the property right thereby created is often referred to as either a limited or patent monopoly. Even though the patent statute does not describe a patent "monopoly," the exclusive rights granted to a patent holder permit exploitation of the patent free from competition for seventeen years and amount to no less however called."). While that term may be apt in a loose sense, in a strict antitrust sense, patents need not confer a monopoly. Whether patent exclusivity confers monopoly power crucially depends on the existence of alternatives to the patented technology, barriers of entry to the innovation market, and the availability of substitutes in the final product market. In fact, a patent does not confer a monopoly in most cases. Therefore, for the sake of accuracy, this Article will use the term "patent exclusivity" rather than "patent monopoly."

Posner and Landes similarly criticize this lack of precision in terminology, arguing that the facile characterization of patent exclusivity as monopoly "led judges to suppose that there is an inherent tension between intellectual property law, because it confers 'monopolies,' and antitrust law, which is dedicated to overthrowing monopolies. That was a mistake. At one level it is confusion of a property right with a monopoly." WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 374 (2003).

will thus have to pay a higher price for the technology in the short run. An economy's ability to generate innovations over time is known as dynamic efficiency. Under patent law, static efficiency, or short-run consumer welfare, is sacrificed for dynamic efficiency.

¶8 Based on this basic view of the goals of patent and antitrust laws, a number of courts have noted the tension between them, at least over the short term. In *SCM Corp. v. Xerox Corp.*, the Second Circuit observed that:

> The conflict between the antitrust and patent laws arises in the methods they embrace that were designed to achieve reciprocal goals. While the antitrust laws proscribe unreasonable restraints of competition, the patent laws reward the inventor with a temporary monopoly that insulates him from competitive exploitation of his patented art.[8]

In *International Wood Processors v. Power Dry, Inc.*, the Fourth Circuit expressed a similar view that "there may be conflict between the patent laws on the one hand, which encourage monopoly power by granting patent holders the right to exclude and be free from competition, and the antitrust laws, on the other hand, which generally proscribe monopoly and encourage competition."[9] Commentators have similarly remarked on the conflict between patent and antitrust. In his seminal article on the patent-antitrust interface, Professor Louis Kaplow characterizes this conflict as "even more deep-seated than is generally perceived."[10]

¶9 The conflict between these two bodies of law has also been described as overstated. It has been argued that they in fact share similar goals. Like patent law, antitrust is concerned with dynamic efficiency. This concern is motivated by the fact that in the long run, the greatest enhancement to consumer welfare comes not from lower prices obtained from static competition, but from the emergence of new technology and new products. There is a wealth of economic evidence that shows that innovation has been the single most significant source of economic growth and improvement in general welfare.[11] Meanwhile, patent law also considers consumer welfare and is cognizant of the harm that patent exclusivity inflicts on consumers. The tradeoff between static welfare loss from short-run supra-competitive pricing and dynamic efficiency gains from more abundant innovation over time is said to be implicitly struck by Congress when it set the length and breadth of patent rights.

¶10 This more conciliatory view of the patent-antitrust conflict has received support from a number of prominent scholars in antitrust, patent, and innovation economics. Professor Suzanne Scotchmer, a leading innovation economist, believes that the tension

---

[8] 645 F.2d 1195, 1203 (2d. Cir. 1981).

[9] *Int'l Wood Processors,* 792 F.2d at 426.

[10] Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813, 1815-16 (1984); *see also* Aaron Xavier Fellmeth, *Copyright Misuse and the Limits of the Intellectual Property Monopoly*, 6 J. INTELL. PROP. L. 1, 3 (1998) (recognizing conflict between two bodies of law); David McGowan, *Networks and Intention in Antitrust and Intellectual Property*, 24 J. CORP. L. 485, 485, n. 1 (1999) (same).

[11] Robert M. Solow, *A Contribution to the Theory of Economic Growth*, 70 Q. J. ECON. 65 (1956); Robert M. Solow, *Technical Change and the Aggregate Production Function*, 39 REV. ECON. STATS. 312 (1957); Paul M. Romer, *Increasing Returns and Long-Run Growth*, 94 J. POL. ECON. 1002 (1986); ELHANAN HELPMAN, THE MYSTERY OF ECONOMIC GROWTH 34-54 (2004).

between patent and antitrust only exists in the short run.[12] Professor Mark Lemley, a leading patent scholar, characterizes the view that antitrust focuses on static efficiency while patent law pursues dynamic efficiency as "oversimplified."[13] Professor Herbert Hovenkamp, the author of the leading antitrust treatise, calls the patent-antitrust conflict "readily exaggerated."[14] He notes that in most cases in which there is an appearance of a conflict, a deeper understanding of patent and antitrust policies will help one realize that none exists.[15] In the few cases that do present a genuine conflict, "a significant portion of it is explained by deep uncertainty about the optimal amount and scope of IP protection. As long as that uncertainty remains, there will always be tension between IP and antitrust."[16] This is a key insight into the patent-antitrust interface. The conflict between these two areas of law is the most acute when antitrust policy may undermine innovation incentives by limiting a patentee's ability to exploit its patent.

¶11 Between these two areas of law, antitrust is probably the one with better-defined goals and policies.[17] There is a clear consensus that antitrust protects consumer welfare. The major theoretical debates in the field focus on how this is achieved in different contexts, such as resale price maintenance and unilateral conduct by a monopolist, which remains highly controversial. While the calibration and implementation of antitrust policy in these contexts may seem open-ended, the goals are fairly well defined: the maximization of consumer welfare through competition, both in terms of product and innovation.[18] Antitrust encourages firms to compete in both the final product market and the technology market. As far as the patent-antitrust interface is concerned, the two paramount considerations for antitrust are the consumer harm resulting from restrictive patent exploitation practices and foreclosure of innovation opportunities by a dominant patentee against rival technology developers.

¶12 Patent law does not share the same clarity in policy.[19] The primary goal of patent law of course is to encourage innovation. Patent law, however, does not pursue innovation at all costs. Otherwise, it would have stipulated exclusivity of unlimited duration and a much more expansive scope of rights. As some commentators have recognized, there can be too much innovation. Innovation is excessive when the social costs of an innovation outweigh its social benefits. Society as a whole would be better off if the resources devoted to develop that innovation are channeled to alternative uses instead. Professors Michele Boldrin and David Levine formulate the test slightly

---

[12] SUZANNE SCOTCHMER, INNOVATION AND INCENTIVES 161 (2004) ("This is more a short-run tension than a long run tension, since in the long run intellectual property law leads to innovation, which improves the welfare of consumers. Since consumer welfare is the concern of competition law, there is no fundamental inconsistency.").

[13] Mark A. Lemley, *Industry-Specific Antitrust Policy for Innovation* 2 (2010), *available at* http://ssrn.com/abstract=1670197 [hereinafter *Industry-Specific Policy*].

[14] Herbert J. Hovenkamp, *Restraints on Innovation*, 29 Cardozo L. Rev. 247, 247 (2007) [hereinafter *Restraints on Innovation*].

[15] *Id*. at 247-48.

[16] Herbert J. Hovenkamp, *United States Antitrust Policy in an Age of IP Expansion*, *in* INTERNATIONAL ANTITRUST LAW & POLICY: FORDHAM CORPORATE LAW 225, 226 (Barry Hawk ed., 2004) [hereinafter *Age of IP Expansion*].

[17] Herbert J. Hovenkamp, *The Intellectual Property-Antitrust Interface*, *in* 3 ISSUES IN COMPETITION LAW AND POLICY 1979, 1982 (ABA Section of Antitrust Law, 2008) [hereinafter *Intellectual Property-Antitrust Interface*].

[18] *Id*. at 1982; *Age of IP Expansion*, *supra* note 16, at 226.

[19] *Intellectual Property-Antitrust Interface*, *supra* note 17, at 1983.

differently, asserting that an optimal level of innovation is attained when the value of an innovation as determined in a competitive market covers its R&D costs.[20] Beyond that, innovation becomes excessive. To take a systemic perspective, patent protection is optimal when the marginal social benefits of protection equals its marginal social costs.[21] It should be obvious that such a condition for optimality requires balancing of countervailing policy considerations and results in less determinacy in patent policy than the focus on consumer welfare found in antitrust law.

B. Long-Run Divergences Between Patent and Antitrust

¶13 The conciliatory view that the tension between patent and antitrust is confined to the short run is overly optimistic; the conflict in fact extends to the long run. While it is true that both antitrust and patent laws seek to encourage innovation, they seem to share contrary visions of what kind of market structure is most conducive to it. This harkens to the Schumpeter-Arrow debate that has been continuing in economics since Joseph Schumpeter first made the famous claims that a monopolist may have greater incentives than competitive firms to innovate and that markets progress through creative destruction whereby firms compete for the market through innovation, displacing an old monopoly with a new one.[22] Kenneth Arrow and others have since challenged these claims, asserting that innovation is more abundant in a competitive environment.[23] According to Arrow, monopolists tend to avoid drastic innovations that will displace its existing product in the market. A monopolist's competitors do not bear the cost of the loss of profit from its previous dominant product, and a monopolist may not stand to gain much from innovation. This is known as the replacement effect, which deters innovation by monopolists. This debate has spawned a large body of literature, both empirical and theoretical, that seeks to verify Schumpeter and Arrow's claims.[24] It remains largely unresolved.[25]

---

[20] MICHELE BOLDRIN & DAVID K. LEVINE, AGAINST INTELLECTUAL MONOPOLY 127 (2008). This, again, is an important insight. As will be explained subsequently, there is a tendency in the existing literature to assume that patentee reward should allow the patentee to capture the full value of its innovation. Allowing the patentee to do so would amount to a windfall. The patent system, in conjunction with antitrust law, provides sufficient incentives to potential inventors so long as they are allowed to recover their R&D expenditure, including the opportunity costs of innovation (what the inventor would have earned in its next best endeavor).

In fact, it has been widely noted in the economics literature that competitive patent races can result in substantial waste in resources. The R&D effort invested by firms to win the race is excessive. *See, e.g.*, Partha Dasgupta & Joseph Stiglitz, *Industrial Structure and the Nature of Innovative Activity*, 90 ECON. J. 266 (1980); Jennifer F. Reinganum, *A Dynamic Game of R and D: Patent Protection and Competitive Behavior*, 50 ECONOMETRICA 671 (1982); Pankaj Tandon, *Rivalry and the Excessive Allocation of Resources to Research* 14 BELL J. ECON. 152 (1983); Partha Dasgupta & Paul Stoneman, *The Economic Theory of Technology Policy: An Introduction*, *in* ECONOMIC POLICY AND TECHNOLOGICAL PERFORMANCE 18-21 (Partha Dasgupta & Paul Stoneman eds, 1987).

[21] Kaplow, *supra* note 1, at 1825-26.

[22] JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY 87-120 (2010).

[23] Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS 609, 619-20 (Nat'l Bureau of Econ. Research ed., 1962).

[24] Philippe Aghion et al., *Competition and Innovation: An Inverted-U Relationship*, 120 Q.J. ECON. 701 (2005) (positing inverted-U relationship between market concentration and innovation performance); Richard J. Gilbert, *Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate?*, *in* 6 INNOVATION POLICY AND THE ECONOMY 159 (Josh Lerner et al. eds., 2006) (same); Richard J. Gilbert &

¶14 While the academic debate still rages, policymakers and enforcers nonetheless must take an explicit or implicit view of the matter. Given the fundamental belief in antitrust that competition is inherently beneficial to society and is the engine of economic progress, it should come as no surprise that antitrust generally subscribes to the Arrowian view. For example, in the Department of Justice-Federal Trade Commission 1995 Guidelines for the Licensing of Intellectual Property, the two agencies identified as a potential harm of cross-licensing arrangements and patent pools the possibility of a reduction in innovation competition.[26] According to Scotchmer, this theory of harm is premised on "the economic hypothesis that competition leads to more R&D than cooperation."[27] In his advocacy for robust antitrust enforcement to foster innovation, Professor Jonathan Baker posits that firms innovate in an attempt to escape product market competition.[28] In the eyes of antitrust enforcers and commentators alike, competition promotes innovation.

¶15 Given the ideological underpinning of patent law and its expansionist tendencies over the last two decades, one may conclude that patent law emerges on the Schumpeterian end of the debate. A preoccupation with promoting technological competition is generally absent in patent doctrines. The increasingly relaxed standard for patentability and the expanding scope of patent rights mean that it is easier than before for an innovator to exclude rivals from competition. This enhanced ability to exclude in turn facilitates the acquisition of market power in the product market. If the central policy goal of patent law is to encourage innovation, the implicit assumption seems to be that dominant firms produce more innovation. Moreover, patent law has not been particularly accommodating to follow-on innovation, which may often compete with the original one. The Federal Circuit has interpreted the doctrine of reverse equivalents very narrowly, which allows a literally infringing invention to escape infringement if it represents a substantial improvement of the patented technology.[29] A more liberal application of this equitable doctrine would allow follow-on innovators to pursue substantial improvements without fear of infringement or strategic behavior by the initial innovator. Not only is the reverse equivalents doctrine rarely invoked, but there are even doubts as to its continual validity.[30] It thus seems that patent law does not actively encourage innovation through competition, evincing a Schumpeterian view of innovation.

Steven C. Sunshine, *The Use of Innovation Markets: A Reply to Hay, Rapp and Hoerner*, 64 ANTITRUST L.J. 75 (1995) (summarizing relevant literature).

[25] SCOTCHMER, *supra* note 12, at 173 (noting that whether competition delivers more innovation than concentrated markets depends on a range of factors); F.M. Scherer, Antitrust, Efficiency, and Progress, 62 N.Y.U. L. REV. 998, 1011 (1987) [hereinafter *Antitrust, Efficiency, and Progress*] ("When the general advance of knowledge opens up opportunities for technological improvements at a fast pace, innovation is likely to proceed most rapidly under relatively fragmented (but not atomistic) market structural conditions. When the relevant knowledge base advances slowly, monopolies or tightly-knit oligopolies are likely to innovate more rapidly than fragmented industries."). *Cf. Industry-Specific Policy*, *supra* note 13, at 2 ("In fact, however, there is substantial evidence suggesting competition itself may act as a greater spur to innovation than monopoly.").

[26] U.S. Department of Justice and Federal Trade Commission Antitrust Guidelines for the Licensing of Intellectual Property (1995) 28-29, *reprinted in* 4 Trade Reg. Rep. (CCH) ¶ 13,132.

[27] SCOTCHMER, *supra* note 12, at 172.

[28] Jonathan B. Baker, *Beyond Schumpeter vs. Arrow: How Antitrust Fosters Innovation* 7 (June 2007), *available at* http://ssrn.com/abstract=962261.

[29] Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1657 (2003).

[30] Tate Access Floors v. Boyden Power Brake Co., 279 F.3d 1357, 1368 (Fed. Cir. 2002) (suggesting

C. Theoretical Approaches to the Patent-Antitrust Interface

1. Michael Carrier's Innovation-Based Framework

¶16 Given the theoretical complexity of the patent-antitrust interface, it is no surprise that it has been the subject of much scholarly writing over the years. Commentators have suggested a variety of approaches to the patent-antitrust interface. Professor Michael Carrier argues that dynamic efficiency considerations should trump consumer welfare concerns in patent-antitrust cases.[31] He proposed a framework for resolving patent-antitrust cases in which a presumption of legality applies to a patent exploitation practice if the patentee can provide a plausible business justification that is only rebutted if a plaintiff can show that innovation is easy to create or difficult to imitate, contending that there are market-based incentives for innovation, and that innovation in the industry tends to be cumulative.[32] The defendant is then allowed to establish the sur-rebuttal by furnishing evidence of innovation in the industry.[33] There is no meaningful balancing of the conflicting static and dynamic efficiency considerations underlying the patent-antitrust interface, which he believes to be beyond judicial capability.[34]

¶17 Carrier does not seem to limit the sur-rebuttal to patentable innovations; he only requires evidence of innovation.[35] Given this broad understanding of innovation, it would take an unusually stagnant industry to deny the defendant evidence to establish the sur-rebuttal. The consequence is that patent exploitation practices will be almost always upheld.[36]

¶18 An anomaly in Carrier's innovation-based framework is that it attempts to resolve two areas of law that individually balance static and dynamic efficiencies without resorting to balancing. While Carrier's framework would certainly improve administrability—it effectively simplifies the entire patent-antitrust enterprise into one inquiry: the existence of innovation of any kind in an industry—it overlooks the reality that restrictive patent exploitation practices can and do harm consumers. Administrability is achieved by ignoring one side of the balance altogether. Unfortunately, the existence of innovation in an industry does not render irrelevant the potential consumer harm of these practices. Concerns about judicial competence do not warrant jettisoning the fundamental considerations of antitrust from the analysis altogether. Therefore, while one must be mindful of the limits of judicial competence in

---

that doctrine had no further relevance after passage of the 1952 Patent Act).

[31] Carrier, *supra* note 1, at 816-33.

[32] *Id*. at 817-19.

[33] Carrier seems to have implicitly accepted the view that consumer welfare has been adequately taken into account within patent law.

[34] *Id*. at 799-800.

[35] In fact, one may argue that given today's relaxed standard for patentability, even an insistence of patentable innovations would not have been much of a hurdle. JAFFE & LERNER, *supra* note 6, at 142-43 ("While there is a formal process of patent examination, in practice the system seems more akin to a registration system: in many cases it appears that a determined patentee can get almost any award he seeks.").

[36] If one examines closely the way Carrier's framework is constructed, the rebuttal, the most complex step in it, is largely superfluous. When confronted by a challenge to a patent exploitation practice, an antitrust court need only ask itself whether there is evidence of innovation in the industry. If so, the case is closed and the practice is upheld.

antitrust cases, it is equally important not to underestimate it and over-simplify the analysis in the name of administrability.

2. Louis Kaplow's Ratio Test

¶19 Of all the approaches that have been proposed for the patent-antitrust interface, Professor Louis Kaplow's framework remains the most sophisticated and nuanced.[37] The pivot of Kaplow's framework is the ratio test, which is defined as patentee reward divided by monopoly loss. Patentee reward and monopoly loss refer to the incremental reward and loss caused by a patent exploitation practice.[38] There is a ratio associated with every type of patent exploitation practice.[39] In order to determine whether antitrust law should permit a particular patent exploitation practice, the ratio associated with that practice is compared with a benchmark ratio of some kind. The benchmark suggested by Kaplow is the patentee reward-monopoly loss ratio implicit in the optimal patent life.[40] This ratio represents the most cost-effective way society can induce invention by adjusting patent life. If a ratio associated with a particular patent exploitation practice is lower than the optimal ratio, the practice should be prohibited. If it is higher than the optimal ratio, the practice should be allowed, subject to the requirement that patent life should be shortened accordingly.

¶20 Despite the theoretical elegance of his framework, Kaplow himself acknowledges that there are considerable obstacles to applying it in real-world contexts. As a second-best solution, Kaplow proposes a cost-effectiveness analysis, which requires the antitrust agency to derive the ratio for all possible patent exploitation practices and align them from the highest to the lowest. A comparison can then be made of the practices that are currently allowed and prohibited to ensure that the total reward is obtained from practices that have the highest ratios.[41] However, he believes that even this second-best formulation is probably too difficult to apply. In its place Kaplow suggests a number of factors which would facilitate the application of the test, including "the extent to which the reward is pure transfer, the portion of the reward that accrues to the patentee, and the degree to which the reward serves as an incentive."[42] When applying his framework to

---

[37] Kaplow, *supra* note 1.

[38] The reason that patentee reward as opposed to marginal social benefit is used is because the optimum patent life is set as given, which is determined by the marginal social benefit and marginal social cost of granting patent protection. Once the optimum patent life has been set, what the ratio test seeks to answer is "whether the total reward to the patentee implicit in the optimal patent life can be achieved at a lower cost." Kaplow, *supra* note 1, at 1831.

[39] Strictly speaking, there is a ratio associated with every patent exploitation practice for every patent, because patentee reward depends on "a number of factors, including the market value of the invention, the structure of the market involving the patented process or product, and the attributes of the patentee (such as marketing and production capacities) that determine its range of options within the market." *Id*. at 1823. However, for ease of application, it is assumed that there is a generalized ratio for every type of patent exploitation practice.

[40] When determining the optimal patent life by comparing the marginal social benefit and marginal social cost of granting patent protection, the policymaker will implicitly determine an optimal patentee reward-monopoly loss ratio. This is because marginal social benefit is dependent on incremental patentee reward, and monopoly loss is one very important component of the marginal social cost.

[41] This analysis is called the cost-effectiveness analysis because the goal of the exercise is to obtain the same total reward in the most cost-effective manner, *i.e.*, by incurring the least aggregate monopoly loss.

[42] *Id*. at 1842. The first of these factors requires some explanation. By a pure transfer, Kaplow refers to a situation in which a patent exploitation practice results in transfer of surplus from one group in society to

concrete examples of patent exploitation practices, Kaplow focuses on the effects of a practice on the ratio's denominator and numerator. Therefore, if a practice were likely to result in a substantial increase in patentee reward without a proportionate increase in monopoly loss, Kaplow would permit it. An example would be a practice that effectuates a pure transfer. If a practice has the opposite effects on the denominator and the numerator, Kaplow would prohibit it. An example would be a price fixing cartel disguised as a licensing arrangement with price restrictions.

¶21 There are two important omissions, however, in Kaplow's framework. First, there is insufficient consideration of the extent to which patentee reward provides incentives to innovate.[43] Kaplow posits a three-step causal link between patentee reward and social benefits. The first is that an increase in patent term or an expansion of the scope of permissible patent exploitation increases patentee reward, which in turn incentivizes innovation, which in turns redounds benefits to society. His ratio test takes as a given that all three steps in the link are valid, which allows him to focus on the relationship between antitrust restrictions on patent exploitation and patentee reward. While he does briefly discuss how innovators may perceive patentee reward from different types of restrictive practices differently, he fails to consider the more crucial question of the incentive effect of patentee reward.[44]

¶22 Kaplow acknowledges that "our knowledge of the functional relationships between the separate links in the chain connecting patent life to social benefits remains quite limited."[45] That was true in 1984. Our understanding of the relationship between patentee reward and innovation incentives has substantially improved since then. There is now a wealth of theoretical and empirical economic literature that questions the second step in Kaplow's causal chain. Once the causal chain breaks down, one can no longer focus on patentee reward, and must begin to consider difficult questions about whether the static efficiency loss caused by a patent exploitation practice serves any useful social purpose, or is merely a windfall to patentees. In other words, instead of focusing on patentee reward, the inquiry must directly address innovation incentives and the social benefits of patent protection. A focus on innovation incentives requires us to distinguish between different types of innovation.

¶23 As it turns out, the degree of dependency on patent protection for innovation incentives varies across types of innovation. For example, economists have found that, perhaps due to the greater difficulty for rivals to reverse engineer contemporary technology, process innovation tends to be less reliant than product innovation on patents for protection.[46] In fact, patentee reward was found to be the least important means by which an innovator reaps benefits from process innovation. The diversity of innovation

---

another, such as licensees to the patentee, without an attendant increase in deadweight loss. In the case of a pure transfer, the patentee reward may increase substantially without a corresponding increase in monopoly loss. Therefore, all else being equal, a patent exploitation practice that results in a pure transfer is to be preferred to one that does not.

[43] While he includes the relationship between patentee reward and innovation incentives as one of the factors in his third-best formulation of the ratio test, he does not provide very detailed guidance on how this factor should be applied in the analysis. This Article seeks to remedy that.

[44] *Id*. at 1838.

[45] *Id*. at 1824.

[46] Richard C. Levin et al., *Appropriating the Returns from Industrial Research and Development*, *in* 1987:3 BROOKINGS PAPERS ON ECONOMIC ACTIVITY 783, 794–95 (1987).

will be explored in greater detail in Section **Error! Reference source not found.**. Suffice it for now to note that Kaplow overlooked this very important issue.

¶24 Second, Kaplow's framework pays scant attention to the importance of cumulative innovation, and how patent exploitation practices may hamper such innovation. Given the dual focus of his framework on patentee reward and monopoly loss, cumulative innovation cannot be easily incorporated. Again, since his article, economists have substantially improved our understanding of innovation generally and cumulative innovation in particular. Cumulative innovation has been shown to be of great importance to technological progress, and patent exploitation practices may significantly affect the prospects of cumulative innovation. Therefore, cumulative innovation must be incorporated in the analysis. The loss of potential cumulative innovation represents a serious social cost of patent rights, and hence belongs to the denominator side of his ratio. Just as the numerator has been redefined to focus on innovation incentives and social benefits, the denominator also needs to take into account loss of cumulative innovation.

¶25 One further weakness in Kaplow's framework is its universalist approach to the optimization of the scope of patent protection and patent-antitrust rules. He attempts to obtain one optimal patent life and one patentee reward-monopoly ratio for each and every patent exploitation practice, regardless of the market environment in which it is pursued. He fails to consider the possibility that the optimization exercise may yield different results according to the industry, or even the type of innovation, at issue.[47] Economic literature suggests that the importance of patent protection as a source of innovation incentives differs widely by industry and type of innovation. A one-size-fits-all approach to the optimization exercise is hence inappropriate. It would perpetuate the crudeness of the tradeoff between static and dynamic efficiencies inhered in the patent system, which will be discussed below. While patent law, at least the statutory part of it, may be bound to such a universalist approach (unless Congress decided to take an industry-specific approach to patent protection), there is no reason that antitrust should not take advantage of the case-specific nature of its decision-making process to take full account of the characteristics of the industry and the innovation at issue. This is what the approach to the patent-antitrust interface proposed in the next Section endeavors to do.

D. Judicial Deference to Dynamic Efficiency Considerations

¶26 Before launching a theoretical examination of the patent-antitrust interface, it is important to examine the prevailing judicial attitude toward it. The patent-antitrust interface requires a balancing between static and dynamic efficiencies. Dynamic efficiency considerations are by nature long term. Innovation can take years to accomplish, especially for research- and resource-intensive R&D activities such as those found in the pharmaceutical industry.[48] To complicate matters further, innovation can be highly uncertain. R&D efforts obviously can succeed or fail. An assessment of dynamic efficiency considerations may require the antitrust courts to enter the perilous territory of

[47] This oversight again is excusable because awareness of the industry-specific variations in dependence on innovation incentives only came about since the publication of Kaplow's article. Kaplow's framework indeed reflects the state-of-the-art of our knowledge about the patent system in 1984.

[48] Burk & Lemley, *supra* note 29, at 1581 (noting that the entire development process for a new drug can take a decade).

estimating the likelihood of success for R&D projects. These myriad complications mean that the antitrust courts are generally loathe to second-guess dynamic efficiency arguments. It sometimes seems that if a dynamic efficiency argument passes the muster of minimum credibility, the courts defer to it.[49] It is aptly illustrated by the D.C. Circuit's treatment of Microsoft's arguments regarding its development of a Java Virtual Machine that is incompatible with its rival Sun Microsystem's version. The court was unwilling to scrutinize the veracity of Microsoft's innovation even though there was ample evidence that Microsoft had developed its own Java Virtual Machine with the express goal to create confusion in the market and to undermine Sun's product.[50] Similarly, in predatory product design cases, the courts have been reluctant to challenge innovation-based defenses to predatory product design claims unless there is clear evidence of predatory intent on the part of the defendant to use a technological feature to exclude rivals.[51] This deferential judicial attitude toward innovation is also evident in other important patent-antitrust cases.[52]

¶27 While a prudent approach to dynamic efficiency arguments is commendable, this author believes that the antitrust courts have shown too much deference to them. The tradeoff in patent law between static and dynamic efficiencies was made when Congress set the patent term and the scope of patent rights. The one-size-fits-all approach under the patent statute means that the balance struck is necessarily a crude one.[53] This balance thus should not be assumed to be universally optimal. Furthermore, while it is true that longer-term and more general policy judgments are usually left to the legislature and that innovation-based arguments are by nature long-term, it is not true that only Congress is equipped to make policy judgments concerning innovation. Even in patent law, the courts are actively involved in fashioning micro policy decisions through individual cases. Professors Dan Burk and Mark Lemley have argued that patent law provides the courts with a range of micro and macro policy levers to make sector-specific adjustments to the various parameters of patentability and patent rights.[54] These decisions may have long-term and broad implications for innovation policy. The Federal Circuit's decision to permit patents for business methods is a case on point.[55]

---

[49] In fact, according to Professor Michael Carrier, this is exactly what the courts should do. Carrier, *supra* note 1, at 762–64.

[50] United States v. Microsoft Corp., 253 F.3d 34, 55 (D.C. Cir. 2001). Admittedly, there was evidence that Microsoft's Java Virtual Machine was superior to Sun's in some ways. This case thus presented a particularly difficult set of facts to the court.

[51] *See, e.g.*, C.R. Bard, Inc. v. M3 Sys., Inc., 157 F.3d 1340 (Fed. Cir. 1998).

[52] *See* United States v. General Elec. Co., 272 U.S. 476 (1926); Standard Oil Co. v. United States, 283 U.S. 163 (1931); Rambus, Inc. v. FTC, 522 F.3d 456 (D.C. Cir. 2008); Schering-Plough Corp. v. FTC, 402 F.3d 1056 (11th Cir. 2005); Matsushita Elec. Indus. Co. v. Cinram, 299 F.Supp.2d 370 (D. Del. 2004); *In re* Indep. Serv. Orgs. Antitrust Litig. (*Xerox*), 203 F.3d 1322 (Fed. Cir. 2000); Filmtec Corp. v. Hydranautics, 67 F.3d 931 (Fed. Cir. 1995); Brunswick Corp. v. Riegel Textile Corp., 752 F.2d 261 (7th Cir. 1984); Foremost Pro Color, Inc. v. Eastman Kodak Co., 703 F.2d 534 (9th Cir. 1983); Berkey Photo, Inc. v. Eastman Kodak Co., 603 F.2d 263 (2d Cir. 1979).

[53] There are of course sector-specific statutes such as the Semiconductor Integrated-Circuit Layout-Design Act 2000 and the Hatch-Waxman Act that make adjustments to this uniform approach.

[54] *See* Burk & Lemley, *supra* note 29, at 1638–67. They note, however, that while judicial application of patent law is at times industry-specific, "there is no reason to believe that these [sector-specific] differences in the law represent a reasoned response to industry differences." *Id*. at 1577. In fact, they assert that with respect to biotechnology and software, the Federal Circuit "has gotten the policy precisely backwards." *Id*. at 1578.

¶28 Policy decisions of general application, such as the length of patent term (and arguably a dramatic extension of patentable subject matter such as that undertaken in *State Street Bank*), no doubt should be reserved for Congress. In contrast, micro policy decisions about the optimal scope of patent rights in a particular case, which antitrust courts make when determining the legality of patent exploitation practices, are rightfully the province of the judiciary. These decisions present no greater obstacles to the antitrust courts than do merger review cases. There is no reason to believe that forecasting how a market will be affected by a merger transaction and how competitors will interact with each other afterward is any more difficult than predicting how a patent exploitation practice will impact innovation. If the antitrust courts can handle merger review cases, applying meaningful scrutiny to innovation-based arguments should not elude them. The long-term nature of dynamic efficiency considerations is a reason for caution, not abdication of responsibility, by the antitrust courts.

II. A THEORETICAL APPROACH TO THE PATENT-ANTITRUST INTERFACE

A. An Overview

¶29 As suggested by Hovenkamp,[56] much of the confusion regarding the patent-antitrust interface stems from the lack of consensus about the optimal amount of patent protection. Antitrust restrictions on patent exploitation affect the scope and the exercise of patent rights, which in turn alters the size of patentee reward. As such, the patent-antitrust rules are an integral part in the determination of the optimal patent protection (*i.e.*, how many incentives to offer to induce innovation). At the same time, it is impossible to delineate the optimal scope of the patent-antitrust rules without placing these rules in the context of the overall determination of optimal patent protection. The patent-antitrust rules are but one of the many pieces of the puzzle.

¶30 There are four relevant parameters in the determination of optimal patent protection: (1) social benefits; (2) social costs; (3) private benefits; and, (4) private costs of patent protection.[57] Social benefits can be broken down to external benefits and private benefits. Likewise, social costs can be broken down to external costs and private costs. Private costs and benefits are those that directly affect the innovator in question. To simplify, they are the rewards received by the innovator from its innovation and the R&D costs incurred by the innovator to develop the technology. External benefits and costs are those that affect other firms in the industry and society as a result of the innovation.

¶31 The external benefits of patent protection include the direct benefits of the innovation for society. For example, if an innovation is a cost-reducing technology, its direct social benefit is the cost savings achieved by society from the adoption of the

[56] *Age of IP Expansion*, *supra* note 16, at 226.

[57] Here it is assumed that the social costs and benefits of patent protection encompass those of the underlying innovation that is being patented. The tacit assumption is that the innovation at issue would not have been created absent patent protection. Strictly speaking, the social benefits and costs of patent protection are only the value and costs of innovation that would not have been created without patent protection. Some, and perhaps a great amount of, innovation may still be created in the absence of a patent system. This distinction would be important for more systemic issues such as whether a patent system should be instituted. On the level of individual innovations, for the purpose of our inquiry, it is appropriate to assume that the social costs and benefits of patent protection and innovation are the same.

technology.[58] If an innovation is the development of a new product, the direct social benefit is the additional amount that consumers are willing to pay for this product, or the difference in consumers' willingness to pay for the improved as compared to the original products.[59] As the term "direct social benefits" suggests, social benefits can also be indirect. These encompass the beneficial impact of an innovation on the R&D programs conducted by other firms in the industry and on the economy in general,[60] and the possibility of follow-on innovation built on this present innovation. Indirect social benefits or spillovers of innovation can be substantial.[61] Professor William Baumol estimates that innovators are able to capture only 20% of the value created by their innovations; the remaining 80% of the value benefits the rest of society.[62] He argues that without these spillovers, a majority of the industrialized world would be condemned to a pre-Industrial Revolution standard of living.[63] Other studies have similarly ascertained the magnitude by which social benefits of an innovation exceed its private benefits.[64] What is particularly noteworthy about these indirect social benefits is the impact of cumulative innovation on overall social welfare. Many commentators have noted the importance of cumulative innovation to technological development.[65] Therefore, to the extent that patent-antitrust rules affect the pursuit of cumulative innovation by other firms, these effects must be taken into account.

¶32 There are two main types of social costs of patent protection. The first is one with which antitrust commentators are most familiar, the deadweight loss created by the supra-competitive pricing imposed by the patentee. From a social perspective, this cost should be minimized to the extent possible without substantially undermining innovation

[58] Suzanne Scotchmer, *Standing on the Shoulders of Giants: Research and the Patent Law*, 5 J. ECON. PERSP. 29, 31 (1991).

[59] *Id*.

[60] The beneficial impact of the new knowledge need not be confined to the industry itself, and may enhance the R&D efforts of firms in other industries. This is especially like with general-purpose technology, such as computers and the internet, whose beneficial impact on innovation clearly extends beyond the boundaries of their own respective industries.

[61] Baumol refers to the indirect external benefits of innovation as spillovers. BAUMOL, *supra* note 6, at 124.

[62] *Id*. at 121, 134–35. *But see* BOLDRIN & LEVINE, *supra* note 20, at 163–66 (disputing claims that spillovers of innovation are substantial).

[63] *Id*. at 125.

[64] *See* Jeffrey I. Berstein, *The Structure of Canadian Industry R&D Spillovers, and the Rates of Return to R&D*, 37 J. INDUS. ECON. 315, 315-28 (1989) (social value of innovation at least twice as high as private value for industries studied); Timothy F. Bresnahan, *Measuring the Spillovers from Technical Advance: Mainframe Computers in Financial Services*, 76 AM. ECON. REV. 742, 753 (1986) (social benefits from mainframe computers very substantial); Michael Kremer, *Patent Buyouts: A Mechanism for Encouraging Innovation*, 113 Q. J. ECON. 1137, 1141 (1998); Edwin Mansfield et al., *Social and Private Rates of Return from Industrial Innovations*, 91 Q. J. ECON. 221, 234 (1977).

[65] *See* Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 843–44, 870 (1990) (arguing that patent scope decisions should allow competitive environment for follow-on innovation); John H. Barton, *Patents and Antitrust: A Rethinking of Patent Breadth and Sequential Innovation*, 65 ANTITRUST L. J. 449, 453 (1997) (arguing that follow-on innovators deserve more protection); Joel Mokyr, *The Industrial Revolution and the New Economic History*, *in* THE ECONOMICS OF INDUSTRIAL REVOLUTION 1 (Joel Mokyr ed., 1989) ("The cumulative effect of small improvements made by mostly anonymous workers and technicians was often more important than most of the great inventions."); Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 Tex. L. Rev. 989, 997 (1997) [hereinafter *Economics of Improvement*"] ("And since 'improvements' may in many cases dwarf the original work in terms of their practical significance, dynamic market efficiency over different generations demands such access as well.").

incentives. Another social cost is less well-documented and perhaps less understood. This refers to the cost imposed by a new innovation on the producers of an existing product, and is sometimes known as the costs of creative destruction. To the extent that the new innovation renders existing products in the market obsolete, the machineries that these firms have installed, and the plants that these firms have constructed, to produce their now-obsolete products may become redundant.[66] Digital photography, which has practically rendered traditional film photography redundant, at least for the leisure photographers, serves as a fitting example. The machines and plants built to produce films and film cameras have largely lost their productive value. Firms certainly do not take into account these welfare losses suffered by their rivals when deciding whether to pursue an innovation. If anything, they may prefer to inflict these losses on rival firms. However, from a social perspective, these losses could be so high as to outweigh the benefits brought forth by the innovation.[67] It is possible for society to be better off without it.

¶33 Finally, another important social cost is the various administrative costs of the patent system. These include the operating costs of the Patent and Trademark Office and the courts, the fees and expenses incurred by patent applicants, and the legal costs incurred by parties in patent litigation. These costs are no doubt substantial.[68] They are, however, mostly relevant on a systemic level. They should have little relevance when determining optimal patent protection in individual cases.

¶34 Private benefits of innovation and private benefits of patent protection must be clearly distinguished. This distinction will have considerable relevance to the resolution of the patent-antitrust conflict. Private benefits of patent protection are estimated against the counterfactual of what would the reward to the patentee be absent patent protection,[69] or in other words, private benefits are the additional value patent protection creates for the innovator. Private benefits of innovation consist of all the revenue earned by the innovator from the exploitation of its innovation, including sale of the final product incorporating the patented technology, licensing of the technology, or even the sale of it in the form of an assignment.[70]

¶35 While licensing and assignment of technology depends on patent protection, product sale does not necessarily rely on it. As it turns out, firms have a variety of ways to capture the benefits of their innovation. These are known as "appropriation mechanisms."[71] First, firms can patent their innovation, which would allow them to exclude competitors from practicing or commercializing the technology. To the extent that the product incorporating the technology does not face close substitutes, the patentee will be able to raise the price of the product and profit from the innovation. The innovator can also resort to trade secrecy and conceal the technology from its competitors altogether. For example, if a firm is able to keep a cost-cutting technology secret and utilize it in its production process, it will acquire a cost advantage

[66] BAUMOL, *supra* note 6, at 136–37.

[67] JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK 98 (2008).

[68] ERICH KAUFER, THE ECONOMICS OF THE PATENT SYSTEM 42–43 (1989).

[69] SCOTCHMER, *supra* note 12, at 275.

[70] BESSEN & MEURER, *supra* note 67, at 97.

[71] ALEXANDRA ZABY, THE DECISION TO PATENT 30 (2010).

over its rivals and be able to undercut their prices. This will allow the firm to capture market share from its competitors, especially if the product is homogenous or if the competing firms' offerings are close substitutes. Trade secrecy is hence another way for the firm to profit from its innovation.

¶36 A firm may appropriate returns from its innovation even in the absence of patent protection and trade secrecy. This related set of appropriation mechanisms generally falls under the rubric of first-mover advantages. The first firm to come up with a new technology and commercialize it gains an advantage over its rivals. First, even if the technology is not protected by trade secrecy (assuming it is a product innovation that is apparent from the appearance of the product itself, such as a hammer) or patent protection, it will still take rivals time and money to reverse-engineer the technology. The technology may be very difficult to reverse-engineer, perhaps because replication of the technology requires substantial tacit know-how that is not discernible from the product itself. Even if the technology is not technically very difficult to reverse engineer, it may take rivals considerable time to do so, simply due to the laborious nature of the process. The amount of time that rivals need to enter the market through imitation is called an "imitation lag."[72] The longer is this lag, the greater is the innovator's first-mover advantage. The innovator will have more time to profit from supra-competitive pricing before rivals enter the market.

¶37 Through a number of empirical studies, Professor Edwin Mansfield has confirmed the time-consuming and resource-intensive nature of imitation.[73] In a survey published in 1981, Mansfield and his co-authors found that the ratio of imitation costs to innovation costs was about 0.65, and the ratio of imitation time to innovation time was about 0.70.[74] In other words, it takes an imitator 70% of the time and 65% of the cost expended by the innovator to replicate the technology. In fact, imitation costs were no smaller than innovation costs in one-seventh of the cases surveyed in their study.[75] Imitation is by no means costless or instantaneous, as is sometimes assumed. If it took an innovator four years to invent a new technology, the imitator on average would need almost three years to replicate it. The innovator will have a three-year window to establish its product in the market and gain consumer acceptance. The innovator will also be able to charge a higher price in the interim to recoup its R&D costs.

¶38 This first-mover advantage can be enhanced by branding and product differentiation strategies. These strategies will give the innovator greater customer loyalty and allow it to more effectively protect its market share after entry by rivals. More generally, sales and marketing efforts also constitute an appropriation mechanism.[76] The most intuitive and direct way to profit from one's innovation is to engage in sales and marketing to boost product sales. On the production side, learning curve advantages have been known to give the innovator an edge over its imitating rivals. By virtue of its early start, the innovator will be ahead of its rivals on the learning curve, and may

[72] Sidney G. Winter, *Appropriating the Gains from Innovation*, *in* WHARTON ON MANAGING EMERGING TECHNOLOGIES 250–51 (George S. Day et al. eds., 2000).

[73] *See* Edwin Mansfield et al, *Imitation Costs and Patents: An Empirical Study*, 91 ECON. J. 907, 907–17 (1981); Edwin Mansfield, *Patents and Innovation: An Empirical Study*, 32 MGMT. SCI. 173, 173–80 (1986).

[74] Mansfield et. al., *supra* note 73, at 909.

[75] *Id*. at 910.

[76] Levin et al., *supra* note 46, at 795.

continue to possess technical and know-how advantages.[77] These learning curve advantages will allow the innovator to produce technically superior products even after imitation has succeeded.

¶39 The variety of appropriation mechanisms means that private returns to innovation are likely to substantially exceed private returns to patenting.[78] This observation has been confirmed by a number of empirical studies.[79] Using a variety of methodologies, such as estimating private value of patents from firms' willingness to pay renewal fees for their patents, economists have tried to determine the private benefits of patent protection independent of the value of the underlying technology. Their general conclusions are that: (1) the private values of patents are highly dispersed; (2) the distribution of values is highly skewed, with most value accounted for by the top-earning patents; and, (3) the average value of patents is much lower than the average R&D costs.[80] While the last conclusion may cause concerns that the patent system generates inadequate innovation incentives and that innovators are systematically incurring losses from their R&D activities, the fact that innovation continues to be created in huge number suggests otherwise.

¶40 In an efficient market, firms could not have been pursuing a loss-making activity for years without suffering severe consequences. Therefore, the inescapable conclusion is that firms recover their R&D costs through other means. Patent protection is but one of the many ways through which R&D investments are recouped. Grilriches, Pakes and Hall found that the aggregate private value of patents amounts to roughly only 10% to 15% of the total national expenditures of R&D.[81] This led Landes and Posner to conclude that "incremental increases in patent protection are unlikely to influence inventive activity significantly and incremental reductions might actually enhance economic welfare."[82] Given that firms rely on a range of appropriation mechanisms to capture of the value of their innovations,[83] the aggregate private benefits of innovation must be considered when determining optimal patent protection, and by extension, the optimal patent-antitrust rules.

¶41 Private costs of innovation and patenting again could be different. Strictly speaking, private costs of innovation refer to the R&D costs incurred to develop an innovation. Private costs of patenting refer to the costs incurred by a firm to secure a patent. However, it seems that the two terms are used interchangeably by economists and

---

[77] *Id*.

[78] BESSEN & MEURER, *supra* note 67, at 98 ("It follows from such observations as the above that the value of the patent per se is not the same as the value of the technology to which it is attached. The private value of the technology is the value of the rents obtained by all means, so this is generally larger than the value of the relevant patent, often much larger.").

[79] Zvi Griliches, Ariel Pakes & Bronwyn H. Hall, *The Value of Patents as Indicators of Inventive Activity*, *in* ECONOMIC POLICY AND TECHNICAL PERFORMANCE 120 (Partha Dasgupta & Paul Stoneman eds., 1987); *see also* Mark Schankerman, *How Valuable is Patent Protection? Estimates by Technology Field*, 29 RAND J. ECON. 77 (1998).

[80] SCOTCHMER, *supra* note 12, at 275.

[81] Zvi Griliches, Ariel Pakes & Bronwyn H. Hall, *The Value of Patents as Indicators of Inventive Activity*, *in* ECONOMIC POLICY AND TECHNICAL PERFORMANCE 120 (Partha Dasgupta & Paul Stoneman eds., 1987); *see also* Schankerman, *supra* note 79 (patent values only recover 24% of private spending in R&D); BESSEN & MEURER, *supra* note 67, at 113 (putting same figure at 3% of private R&D investment).

[82] LANDES & POSNER, *supra* note 7, at 327.

[83] For a more detailed discussion of the different appropriation mechanisms for innovators to capture the value of their innovations see Part **Error! Reference source not found.**.

lawyers alike. When economists speak of and measure patent costs, they seem to have in mind the R&D costs, perhaps together with the patenting costs.[84] For most patents, there are good reasons to believe that R&D costs dwarf patenting costs.[85] The focus of our attention in the ensuing discussion will be R&D costs.

¶42 How do the patent-antitrust rules fit into the picture? How do the rules relate to all these parameters? The conventional wisdom is that within the patent system, consumer welfare loss is traded for patentee reward. In other words, private benefits of innovation are generated at the expense of external costs. The relationship, however, is more complex than that. Private benefits, in particular patentee reward, are related to both external costs and external benefits. Returns from innovation are derived from multiple sources. In addition to consumer welfare loss, these returns are also derived from the minimization of spillovers and the denial of follow-on innovation by rival innovators. In other words, an increase in patentee reward may cause a loss of surplus by other firms and beneficiaries of spillovers. The tradeoff is hence not just between patentee reward and consumer welfare loss. It is more accurately characterized as one between patentee reward and net social benefits.[86]

B. Determination of Patentee Reward

1. A Costs-Based Approach to Patentee Reward

¶43 From a social perspective, an innovation would be beneficial for society if its social benefits outweigh its social costs, or if it results in net social benefits. Society would be better off with the innovation than without it. In addition, society would like to maximize its net social benefits from each innovation so long as it is made. The decision to pursue the innovation, however, lies not with society but with the innovator. It is its private calculus of costs and benefits that determines whether an innovation is made. An innovator does not care about the net social benefits of its creation. It is only concerned

---

[84] *See* SCOTCHMER, *supra* note 12, at 273; *see also* Schankerman, *supra* note 79, at 94. Schankerman found that the top 1% of patents accounted for 12% of the total value of patent rights in the pharmaceutical industry, 14% in the chemicals industry, 21% for mechanical patents, and 24% for electronics patents. He further found that the top 5% of patents accounted for 34% of the total value of patent rights in the pharmaceuticals industry, 38% in the chemicals industry, 50% for mechanical inventions, and 55% for electronic inventions. Schankerman, *supra* note 79, at 94.

[85] Bronwyn H. Hall, Adam B. Jaffe & Manuel Trajtenberg, *Market Value and Patent Citations*, 36 RAND J. ECON. 16 (2005) (reporting R&D costs per patent of about $2.8 million in 1996 dollars); Jean O. Lanjouw & Mark Schankerman, *Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators*, 114 Econ. J. 441 (Apr. 2004) (reporting costs per patent in 1998 dollars to be $4 million for pharmaceutical patents, and around $2 million for chemical, mechanical, and electronics patents). Meanwhile, the costs of patenting itself are only in the neighborhood of tens of thousands of dollars. *See* Gene Quinn, *The Cost of Obtaining a Patent in the US*, IP WATCHDOG (Jan. 28, 2011, 1:14 PM), http://ipwatchdog.com/2011/01/28/the-cost-of-obtaining-patent/id=14668/ (estimating patent attorney fees to be $5,000 for an extremely simple technology to more than $15,000 for a highly complex one). According to the Patent and Trademark Office, the patent filing fee for three claims or less ranges from $220-330, the patent search fee from $100-$540, and the examination fee from $140-650. *See* United States Patent and Trademark Office, Fee Schedule (effective October 2, 2008), *available at* http://www.uspto.gov/web/offices/ac/qs/ope/fee2009september15.htm (last visited March 4, 2011). Adding all these fees together, the costs of obtaining a patent are still dwarfed by the R&D costs.

[86] There does not seem to be a tradeoff, at least not a strong one, between other sources of private benefits and the various components of net social benefits. Increased returns from sales and marketing may reduce surplus for other firms. However, this has nothing to do with the patent-antitrust interface.

with its profitability. The alternative to innovation is to spend the funds on another investment opportunity. A firm contemplating innovation is confronted with an array of investment options, such as building a new production plant to expand capacity, launching a marketing campaign to increase sales, improving its in-house IT system to improve its operational efficiency, or developing a cost-cutting technology that will allow it to produce its goods more cost-effectively. The firm will choose to pursue innovation if it promises the highest rate of return, or if returns from innovation cover its direct R&D costs plus opportunity cost, which would be the returns that the firm can obtain from its next best investment option. If the returns from innovation slightly exceed its R&D costs and opportunity cost, the firm will choose to develop the technology.

¶44 The implication is that from the perspective of the optimization of patent protection and patent-antitrust rules, we would first like to determine the desirability of an innovation by weighing its social benefits against its social costs. If the former exceeds the latter, society would benefit from the innovation. Once it is determined that an innovation creates net social benefits, society would prefer to obtain it at the lowest possible costs. Society would prefer to offer the innovator a reward just large enough to cover its private costs of innovation, which now include both direct R&D costs and opportunity cost. Any reward greater than what is strictly necessary to obtain the innovation will be socially wasteful and will constitute a windfall to the innovator. The task of maximizing social gains from an innovation while ensuring that innovation costs are covered can be achieved through a variety of means, such as adjusting the length of patent term, the scope of patent rights, or the amount of antitrust restrictions on patent exploitation. Given that this Article is concerned with the patent-antitrust interface, it will focus on the last means.[87]

¶45 In fact, if it was possible for the government to predict accurately *ex ante* the costs for developing every innovation that emerges in the economy, it could provide a lump-sum prize to reward innovators.[88] So long as the government can make a credible

[87] Given the composition of social benefits and social costs, it is obvious that their magnitude is to some extent dependent on the scope of patent rights and the patent-antitrust rules. When we determine the social desirability of an innovation, we must do so against a backdrop of patent and antitrust rules. But if our goal is to determine the optimal scope of patent-antitrust rules, we run the risk of circularity in this exercise. In economic parlance, there is a risk of endogeneity in the analysis. The problem with determining the social desirability of an innovation is that it must be done against some background legal rules that in turn affect its social benefits and costs. For example, legal rules that minimize spillovers and the possibility of cumulative innovation will foreclose a major source of social benefits. Likewise, legal rules that permit liberal compulsory licensing may introduce competition and help to drive the price of the product toward marginal cost. This will help to reduce deadweight loss and hence the social costs. In the ideal world, we would like to determine the social desirability of an innovation against a set of neutral patent and antitrust rules. Unfortunately, there are no such rules. Patent and antitrust rules alter the size of patentee reward one way or the other. The best we can do is first to ascertain the social desirability of an innovation, and then make adjustments accordingly to maximize its net social benefits. This maneuver is in some ways similar to Kaplow's assumption that the current length of patent reward is optimal from a social perspective, and the patentee reward-monopoly loss ratio for every patent exploitation practice is to be compared to the ratio implicit in the last year of patent term to see whether the same patentee reward can be obtained at a lower monopoly loss. Kaplow, *supra* note 1, at 1826–29.

[88] Ian Ayres & Paul Klemperer, *Limiting Patentees' Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive* Remedies, 97 MICH L. REV. 985, 1008 ("All problems of patent regulation ultimately devolve to a question of government information. If government had good information about the cost and/or consumer value of a particular information, it could pay efficient lump-sum bounties to the inventor and be done with the ex post deadweight loss altogether.").

commitment to paying the prize, or can be effectively prevented from reneging after the fact, this system would be superior to the patent system, which incurs a deadweight loss as a result of supra-competitive pricing. As Professors Ian Ayres and Paul Klemperer observed, "[t]he current patent system—which grants the patentee a monopoly for a limited number of years—is structured to minimize the government's information requirement."[89] The problem is that no one, including the government, possesses perfect information and no one can accurately predict innovation costs in advance. The patent system is our best option in light of this limitation.

¶46 To circumvent this problem of imperfect information, the patent system eschews the task of estimating innovation costs *ex ante* altogether. It simply makes no attempt to match patentee reward to innovation costs.[90] What seems to have happened instead is that patentee reward is roughly calibrated to the social value of innovation. For example, under the principle of pioneering patents, groundbreaking innovations that open up a new field of research or commercial development, which obviously create greater social value, are entitled to broader protection than less significant innovations.[91] Under the doctrine of reverse equivalents, a technological improvement may escape literal infringement if it represents a substantial progress from existing art.[92] Even the basic patent doctrine of nonobviousness seems to reflect the value of the innovation as indicated by its improvement from prior art.[93] The extent of improvement from prior art is one indication of the social value of an innovation; the greater the social value, the more expansive the scope of protection. As Scotchmer noted, "[g]iven that the length and breadth of patent protection cannot depend on the expected costs of an R&D project, the only way to ensure that firms undertake every research project that is efficient is to let the firms collect as revenue all the social value they create."[94] Given that the basic parameters of patent protection, such as patent length and scope of patent rights, are set *ex ante*, this may be the only feasible solution to the problem of imperfect information within the patent system.[95]

---

[89] *Id*.

[90] LANDES & POSNER, *supra* note 7, at 300. Scotchmer, *supra* note 57, at 38. Scotchmer corroborates this statement, observing that "[t]he private value of patent protection is linked to the social value of the technology through market demand, but is not linked to firms' research costs."

[91] Burk & Lemley, *supra* note 29, at 1656.

[92] *Id*. at 1657-58. Burk and Lemley note that both doctrines have not been invoked by the Federal Circuit, leading some to question their continual validity.

[93] *Id*. at 1648-52; *see also* Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 HIGH TECH. L.J. 1, 34–36 (1992).

[94] Scotchmer, *supra* note 58, at 31. This of course does not mean that the patent system successfully allow firms to appropriate the full social value of their innovations. As suggested previously, innovation does create substantial external benefits that are not captured by its creators. What is suggested here is that patentee reward is intended to capture the social value of innovation, not that it necessarily succeeds in doing so.

[95] Even if one were to accept that the lack of perfect information requires patentee reward to be benchmarked against social benefits of innovation, the correct benchmark should be net social benefits, and not gross social benefits. An innovation is only beneficial to society if it redounds net social benefits. Therefore, patentee reward must take the full social costs into account. The innovator will naturally consider private costs of innovation, but it will pay little heed to the external costs, such as consumer welfare loss. In a system in which patentee reward is calibrated according to social benefits, some mechanism must be devised to deduct the external costs from patentee reward. Only then will patentee reward provide the socially efficient level of incentives.

2. Critique of an Innovation Value-Based Approach

¶47 The focus of our inquiry is not the patent system as such, but how antitrust law may adjust patentee reward by tightening or relaxing restrictions on patent exploitation.[96] The question is whether the antitrust courts use private costs or social value of innovation as the basis of their decisions. At first glance, it seems to be the latter. The prevailing judicial approach seems to be that so long as a patent exploitation practice does not exceed the permissible scope of patent rights granted by the statute, it would generally be upheld.[97] If social optimality requires patentee reward to be benchmarked against innovation costs instead of social value, and the patent system is prevented from doing so because of imperfect information *ex ante*, the crux of our inquiry then becomes whether the antitrust courts should perpetuate the suboptimal decision-making of the patent system or embrace a different standard instead.

¶48 Two arguments can be made in defense of using social value of innovation as a benchmark under the patent-antitrust rules. The first argument is the familiar one of imperfect information. Just as the patent system is unable to determine innovation costs *ex ante*, the antitrust courts will be equally incapable of doing so *ex post*. Therefore, using social value of innovation as a benchmark is our only feasible solution. While this imperfect information argument may apply to the patent system,[98] it is less persuasive in the antitrust context.[99] Antitrust decisions are made in *ex post* proceedings that take place after the innovation has been created and exploited. Innovation costs can be ascertained with a reasonable degree of accuracy in *ex post* judicial proceedings.[100] Once an innovation has been completed, the R&D costs are ascertainable from the firm's internal accounting documents. Moreover, the antitrust courts have the capability to handle complex calculations. They regularly do so with the sophisticated statistical and econometric analyses presented in market definition and quantification of consumer harm. Therefore, the information problem is considerably more surmountable *ex post* in an antitrust proceeding than *ex ante* under the patent system.

¶49 In comparison, an accurate assessment of social benefits will be beset with tremendous difficulty. The value of an innovation to a firm cannot be determined independent of the market structure in which it operates. The amount of profit a firm can

[96] While what ultimately rewards the innovator is not just the patentee reward, but the full private benefits of innovation emanating from the variety of appropriation mechanisms, antitrust law exerts the most direct impact on patentee reward. It has relatively less, or no, impact on the other means of appropriation. Therefore, the discussion that follows will refer to patentee reward, ignoring the impact of other private benefits of innovation for the time being.

[97] *See* Brulotte v. Thys Co., 379 U.S. 29 (1964); *see also* CSU v. Xerox Corp., 203 F.3d 1322 (Fed. Cir. 2000).

[98] In fact, this argument is not entirely apt even when applied to patent law itself. Although the general parameters of patent protection are set by the statute *ex ante*, the courts can and do make adjustments to the scope of protection and rights *ex post*. This is the essence of what Burk and Lemley call "policy levers" in patent law. Burk & Lemley, *supra* note 29, at 1638-67. Even the standard of patentability is adjusted *ex post* by the courts. Merges has argued that the doctrine of nonobviousness should be adjusted *ex post* to take into account the R&D costs of innovation. Merges, *supra* note 93, at 34–54.

[99] A related, but different, objection is that even if private costs of innovation could be accurately calculated *ex post*, what matters to the decision to invest in an innovation is not these costs which are ascertained *ex post*, but the expected costs of innovation arrived at by the firm *ex ante*. This argument will be addressed subsequently.

[100] Merges, *supra* note 93, at 55.

make from an innovation depends on the existence of substitutes and competitors and entry barriers, among other things. For example, in Baumol's model of contestable, oligopolistic markets involving routinized innovation, the value of innovation is said to be equal to its marginal cost of production plus a depreciation charge that will allow the firms to recover their R&D outlays.[101] In addition to private benefits, social benefits of an innovation include the benefits from follow-on innovations made possible by the initial innovation. The full benefits of cumulative innovation will be difficult to ascertain with a reasonable degree of accuracy. Social benefits become practically impossible to measure accurately once one incorporates the spillover effects. Therefore, setting patentee reward against the social value of innovation is unlikely to be feasible.

¶50 This leads one to wonder how the existing patent system undertakes this complex exercise. The answer is that it does not. While the ostensible basis for patentee reward is social value of innovation, the patent system makes only limited effort to adjust the scope of patent protection to the social value of an innovation. As mentioned previously, some patent doctrines offer greater protection to groundbreaking innovations. This is in some way recognition of the importance of an innovation. The focus under these doctrines, however, is not the innovation's social value, but its technical merit. The standard is usually by how much the innovation improves upon the prior art. While technically meritorious innovations may have a greater chance of being social valuable, this correlation is by no means certain. One can imagine a highly sophisticated egg-cracking device that leaves no mess behind. It may be a substantial improvement from existing egg-cracking technology, but its social value is likely to be limited.

¶51 The lack of a meaningful attempt by the patent system to ascertain the social value of an innovation leaves open a distinct possibility that patentee reward may exceed its social value. This would amount to overcompensation of the innovators and is indefensible from a social welfare perspective. This is a further reason why antitrust cannot rely on the scope of patent rights as a basis for deciding patent-antitrust cases. The use of social value as a benchmark is beset with problems. The lack of a real attempt by patent law to match the scope of protection to social value means that scope of patent rights is a highly inaccurate benchmark for antitrust cases. Deciding patent-antitrust cases with a view to preserving patentee reward will unlikely lead to socially optimal outcome.

¶52 The second argument, which more generally applies to both the patent-antitrust rules and the patent system, is that an innovator should be allowed to reap the full benefits of her innovation. That is, the innovator should enjoy the benefits of what she creates. This argument has great intuitive appeal. What can be more fair than reaping the full value of one's own labor? This argument also resonates with the Lockean labor theory of property.[102] Unfortunately, Lockean theory only applies to physical property and there is no strong theoretical justification for allowing intellectual property owners to appropriate the full value of their creations. Intellectual property law generally, and patent law in particular, have always been justified on utilitarian grounds.[103] Intellectual property protection is offered to generate incentives, not to protect the value of one's

[101] BAUMOL, *supra* note 6, at 187.
[102] JOHN LOCKE, LOCKE'S TWO TREATISES OF CIVIL GOVERNMENT 314–15 (Peter Laslett ed., 1967).
[103] Burk & Lemley, *supra* note 29, at 1580.

creation.[104] This is clear from the language of the Constitution.[105] Numerous commentators have endorsed this incentive theory for patent protection. Professor Edmund Kitch argues that "a patent should not be granted for an innovation unless the innovation would have been unlikely to have been developed absent protection of a patent."[106] Professor Robert Merges expresses a similar view that "[t]he conventional ideal standard of patentability is that patents should only be awarded to those inventions that would not have been made without the availability of the patent."[107] Burk and Lemley observe that "[t]o a greater extent than any other area of intellectual property, courts and commentators widely agree that the basic purpose of patent law is utilitarian: We grant patents in order to encourage invention."[108] The notion that patentee reward should reflect the social value of an innovation because an innovator should gain the full value of its creation can only be justified by a just desert or reward theory of patent law, which has been roundly rejected.[109] If the goal of patent law is genuinely utilitarian, patentee reward must be calibrated to the private costs of innovation.

¶53 The idea that patentee reward should only aim to cover the private costs of innovation has also received considerable support. With respect to the determination of optimal patentee reward, Boldrin and Levine note that "producers must be compensated for their costs, thereby providing them with the economic incentive of doing what they are best at doing. But they do not need to be compensated more than this."[110] Professor Daniel Crane observes that "statutory innovation, legal doctrine, and judicial, executive, and administrative practice have begun to cast intellectual property as a right to recover the risk-adjusted costs of invention but not necessarily to exclude others from the invention,"[111] and that "[t]he optimal solution is the inclusion of those rights that grant just enough reward to induce the incentive or creative activity at the lowest social cost possible."[112] Merges' proposed reformulation of the doctrine of nonobviousness—which determines the breadth of patent rights—based on R&D cost consideration is also an implicit recognition that patentee reward should be concerned with recoupment of costs.[113] In light of the overwhelming consensus on the utilitarian justification of the patent system, and widespread recognition that patentee reward should only aim to cover private costs of innovation, it is surprising that most patent-antitrust cases focus on the preservation of the scope of patent rights and make no attempt to consider innovation costs.

---

[104] Crane, *supra* note 1, at 254 ("Intellectual property is incrementally moving away from the conventional right of the landowner to fence out trespassers and toward a right to collect royalties from constructive licensees."); *see also* Bessen & Meurer, *supra* note 67, at 33–34 (noting doctrinal differences between patent rights and real property rights).

[105] U.S. CONST. art. 1, § 8, cl. 8 ("To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.").

[106] Edmund W. Kitch, Graham v. John Deere, *New Standards for Patents*, 1966 SUP. CT. REV. 293, 301

[107] Merges, *supra* note 93, at 19.

[108] Burk & Lemley, *supra* note 29, at 1597.

[109] *Id*. at 1597. Some commentators, however, have advocated a reward-based theory for patent protection. *See, e.g.*, A. Samuel Oddi, *Un-Unified Economic Theories of Patents—The Not-Quite-Holy Grail*, 71 NOTRE DAME L. REV. 267, 275–77 (1996).

[110] BOLDRIN & LEVINE, *supra* note 20, at 128.

[111] Crane, *supra* note 1, at 254.

[112] *Id*. at 299.

[113] Merges, *supra* note 93, at 4.

¶54 Setting patentee reward higher than private costs of innovation not only overcompensates the innovators,[114] but it also attracts rent-seeking behavior.[115] Potential innovators will expend considerable costs to obtain patent protection, as demonstrated by patent races. Landes and Posner believe that the high incidence of failed drug development projects is evidence of rent-seeking behavior by the large pharmaceutical firms.[116] The existence of substantial economic rent from patent protection has been confirmed by Professors James Bessen and Michael Meurer. Economic rents are supernormal or excess profits earned by patentees above the level of competitive returns. They found that in 1999, the pharmaceuticals and chemicals industries earned $15.2 billion and other industries earned about $3.2 billion in patent rents in 1992 dollars.[117] They further found that the annual rent per patent in the pharmaceuticals and chemicals industries in 1999 was about $531,000, and that for other industries roughly $39,000 per patent.[118]

¶55 The disproportionately high patent rent for the pharmaceuticals and chemicals industries lends support to Landes and Posner's intuition that at least some of the failed R&D projects in the pharmaceuticals industry constitute rent-seeking behavior. This is further confirmed by Professor F.M. Scherer's observation that pharmaceutical firms tend to increase their R&D expenditure following a rise in revenue. Their R&D expenditure seems not to be determined by the number of promising technological opportunities, but by the amount of cash that they have to spare. He argued that a competitive rent-seeking model best describes the R&D agenda of these firms.[119] Rent-seeking behavior is particularly harmful in a differentiated market with competing patented products, which characterizes many pharmaceutical markets.[120] When there are numerous firms competing for economic rent, and so long as competition does not completely eliminate

---

[114] This windfall problem would be alleviated if we could somehow internalize all the social costs of patent protection so that the innovator will take into both the social benefits and costs of its innovation when making an innovation investment decision. However, there seems to be no effective means of doing this. The inability to internalize social costs into the innovator's investment calculus will most likely result in excessive innovation.

[115] LANDES & POSNER, *supra* note 7, at 300.

[116] *Id*. at 315–16 ("The prospect of large profits enabled by patent protection provides a lure for investment in research, yet the resources devoted to that research conceivably might be socially more productive in an industry in which innovation is not rewarded with a monopoly.").

[117] BESSEN & MEURER, *supra* note 67, at 114.

[118] *Id*. at 117. However, one may argue that this is not a meaningful comparison because the correspondence between patent and product, or the patent to final product ratio, in the pharmaceuticals and chemicals industries is probably lower than in other industries. One patent may already constitute one product in the former two, whereas in other industries such as semiconductors, automobile or complex machinery, thousands of patents may be needed to build one final product. *Id*. at 71 (semiconductors require thousands of patents). These other industries are more patent-intensive in the sense that the creation of one product requires the incorporation of a greater number of patented technologies. In a way, one may even question how we should draw the meaningful boundary of a technology in different industries, and whether the scope of a single patent is the correct one. A novel suspension system in an automobile is what most of us would think of as one discreet technology, so is a new protease inhibitor for treating HIV. The number of patents required to constitute each technology can be grossly different. A suspension system is arguably not even a meaningful economic product in the eye of end consumers. Instead of counting technology by the number of patents, there is a good argument that it should be done by the number of discrete, independently functional technology.

[119] F.M. Scherer, *A Note on Global Welfare in Pharmaceutical Patenting*, 27 WORLD ECON. 1127, 1136 (2004).

[120] BOLDRIN & LEVINE, *supra* note 20, at 10.

this rent, the rival oligopolists will produce too many similar items and this over-production will offset much of the social surplus from innovation.

C. A Closer Look at the Private Costs of Innovation

1. Unrecoverable Sunk Costs vs. Fixed Costs

¶56 A number of issues pertaining to R&D costs remain to be addressed. The first one is whether R&D costs are sunk costs that should be disregarded altogether in patent-antitrust cases or whether they are fixed costs of production that need to be recouped in order to preserve dynamic efficiency. Boldrin and Levine have argued that in a competitive market, no extra legal protection is needed to allow a firm to recover its R&D costs. They characterize R&D costs as the costs for producing the first unit, asserting that "[s]ince only ideas embodied in people or products matter, the cost of creation is the cost of producing the first unit. Such a 'sunk cost' is very ordinary in economics, and poses no particular threat to perfect competition."[121] Baumol agrees that if R&D costs were indeed sunk, the nonrivalrous nature of the consumption of innovation means that any license fee that would prevent some users from adopting the technology would be inefficient.[122] This, however, is where their agreement ends. According to Baumol, while the R&D costs for serendipitous innovation, which he understands to be a one-time innovation that arises more out of luck than persistent effort, are sunk, the R&D costs for what he calls routinized innovation are recurrent and should be taken into account when determining the optimal patentee reward.[123]

¶57 The better view is that R&D costs need to be recouped. First and foremost, the argument that sunk costs can be disregarded ignores the dynamic efficiency aspect of the problem. Patent protection affects not only the present innovator, but also potential innovators who will be deterred from pursuing R&D if they knew that they would be denied recovery of their R&D costs. Second, even Boldrin and Levine admit that there are situations in which ordinary competitive rent will be insufficient to induce innovative activity.[124] They believe that the subset of innovations to which this applies is small. There are reasons, however, to believe that they may have underestimated the size of the subset. Third, one may argue that patentee reward should aim to permit recoupment of the R&D costs of even serendipitous innovation. This is not to suggest that they are likely to be in the same magnitude as the costs of routinized innovation, yet it remains true that serendipitous innovation often emerges in the context of a routinized innovation project.[125] A research team may have accidentally discovered a chemical that turns out to be highly valuable.[126] The team probably has incurred considerable costs to put itself in a

---

[121] BOLDRIN & LEVINE, *supra* note 20, at 2–3.

[122] BAUMOL, *supra* note 6, at 223–24.

[123] *Id*. at 38-39 (Under routinized innovation, "tomorrow's sunk costs are still entirely variable today, and the fate of sunk outlays of the past is likely to affect any rational decision on the resources to be devoted to innovation in the future."). Routinized innovation refers to R&D activities that are "a regular and even ordinary component of the activities of the firm." *Id*. at 2. For more discussion of routinized innovation, see Section **Error! Reference source not found.**.

[124] BOLDRIN & LEVINE, *supra* note 20, at 3.

[125] Merges & Nelson, *supra* note 65, at 839.

[126] Viagra is a prime example. *See* Ian H. Osterloh, *The Discovery and Development of Viagra (sildenafil citrate)*, *in* SILDENAFIL 1, 3 (U. Duzendorfer ed., Burkhauser Verlag 2004).

position to make the accidental discovery. At least part of these costs should be considered when determining the size of patentee reward.

¶58 Finally, if patentee reward is to be based on innovation costs, and yet R&D costs are deemed to be sunk and unrecoverable, as argued by Boldrin and Levine, the necessary implication would be the abolition of the patent system, which is the central thesis of their book *Against Intellectual Property*. This is probably a step too far. What this Article advocates is only a return to the utilitarian roots of the patent system and the use of innovation costs as a reference point for the patent-antitrust rules when delineating the permissible scope of patent exploitation.

2. Taking Risks of Innovation into Account

¶59 It is often argued that patentee reward needs to take into account the risks of innovation. Innovation is a risky venture. Many R&D projects fail to produce any useful product for the firm while incurring considerable costs. Prime examples are failed drug development projects.[127] The argument that patentee reward needs to take risks into account usually comes in two forms: first, that the risk of a particular project needs to be reflected in its reward so that expected returns, and not actual returns, from the project are sufficient to cover its costs, and second, that patentee reward needs to allow firms to cover the costs incurred in both successful and failed projects. If firms only make a competitive return from successful projects while incurring wasted R&D costs from failed projects, firms would be running at a loss and innovation would be deterred.

¶60 These arguments are certainly valid. There is no denying that the risks of innovation must be considered when setting optimal patentee protection. However, it cannot be true that patentee reward must both reflect the risks of individual projects and cover the wasted R&D costs from failed projects. Doing so would amount to an insurance against all failures in R&D, which would surely lead firms to take on excess R&D risks and encourage the kind of rent-seeking behavior identified by Landes, Posner, and Scherer. As valuable as innovation is to society, society cannot provide a safety net for all failed R&D projects. Firms must bear the consequences of their failures in order for the R&D decision-making process to remain efficient, at least from a private perspective.[128] A better approach to incorporating risks of innovation is to adjust the patentee reward by the *ex ante* estimation of success rate. If an R&D project was estimated to have a 50% success rate prior to its commencement, and private costs of innovation turn out to be $1 million, the patentee reward should be raised to $2 million to ensure that the innovator's *ex ante* risks are reflected.[129]

---

[127] Studies of pharmaceutical R&D estimate that if one were to screen 5,000 to 10,000 chemical compounds for possible therapeutic value, roughly 250 of them will show sufficient promise for further testing. Out of these 250 molecules, about five of them will be put through clinical testing. And 80% of the drugs that are clinically tested end up in failure. JAFFE & LERNER, *supra* note 6, at 42. For more information on the economics of R&D in drug development, see William S. Comanor, *The Economics of Research and Development in the Pharmaceutical Industry* in PHARMACEUTICAL INNOVATION: INCENTIVES, COMPETITION, AND COST-BENEFIT ANALYSIS IN INT'L PERSPECTIVE 54 (FRANK A. SLOAN & CHEE-RUEY HSIEH EDS., 2007).

[128] This process probably cannot be socially efficient because firms will not capture the full social benefits of innovation, nor will they internalize the full social costs.

[129] Boldrin & Levine, *supra* note 6, at 144–45 (taking risk into account does not fundamentally change how innovation investment decision is made, but merely requires probability to be taken into account).

¶61 A caveat is in order. It is easy to overstate the extent of failure in innovation. Projects that do not produce readily commercializable products may generate useful knowledge that turns out to be valuable in subsequent projects. Unintended outcome or products emerge from R&D projects regularly. If the risks of failure are to be reflected in patentee reward, the unexpected or incidental benefits of these projects should be counted as well. These benefits should be deducted from the expected patentee reward.

D. A Proposed Approach to the Patent-Antitrust Interface

¶62 So what lessons does the foregoing discussion hold for resolving the patent-antitrust conflict? An innovation would be beneficial to society if its social benefits outweigh its social costs. If the contrary were true, society would be better off without the innovation. This may suggest that the antitrust courts should attempt to determine the social desirability of innovation. If an innovation fails to create social benefits, the court should try to limit the scope of patent exploitation to such an extent that private benefits of innovation are insufficient to cover its private costs. This would signal to future innovators the need to undertake a careful evaluation of the net social benefits of their R&D projects. The antitrust courts would help to restore the proper function of the patent system and to ensure that only net socially beneficial innovations are pursued.

¶63 While that may be theoretically sound, it would be immensely difficult in practice for the antitrust courts to arrive at a confident estimation of the social benefits of an innovation, even if the exercise is done after the fact. The difficulty with ascertaining social benefits has been alluded to earlier. Nascent technology compounds this difficulty. Possible uses of the technology may not have been discovered at the time of the suit. Therefore, even if the courts were able to derive accurate estimations of the total social benefits and costs of an innovation at the time of the suit, they would still fail the exercise by excluding possible future costs and, more importantly, benefits. If the courts systematically underestimate the social benefits of nascent technologies, litigants challenging a patent exploitation practice under antitrust law will be induced to bring suits early in the life of a technology.

¶64 Moreover, to allow the antitrust courts to determine the social desirability of an innovation and to roll back the incentives provided by the patent system would amount to a usurpation of the patent system. Even if the courts were able to come to a correct estimation in every case, it would introduce substantial uncertainty to corporate business planning. A firm would stand to suffer great loss if it turned out to have been wrong about the social costs and benefits of its innovation. The risks of innovation would be magnified. And if the courts were to make mistakes in their judgment, which they are bound to do, the costs of false positives would be enormous. In light of these serious detriments of overreaching, the antitrust courts should assume innovations that have been created as socially desirable and focus on calibrating patentee reward that would allow the innovator to recover its innovation costs while maximizing net social benefits.

¶65 In patent-antitrust cases, antitrust law should strive to maximize net social benefits while ensuring that private benefits of innovation cover innovation costs.[130] One way to

[130] Strictly speaking, the concern of the antitrust courts is that private benefits of innovation cover innovation costs. However, the component of private benefits over which the courts exert the most direct control is patentee reward. While legal rules certainly have impact on the other appropriation mechanisms,

think of this task is as a problem of constrained maximization. The entity to be maximized is net social benefits of an innovation, subject to the constraint that private benefits exceed innovation costs. This approach merely represents a return to the utilitarian roots of the patent system, which require patentee reward to be benchmarked against innovation costs. The goal of patent protection is to provide innovation incentives by ensuring that innovation costs are recovered. Since innovation costs include both direct R&D costs and the opportunity cost of innovation, if patentee reward slightly exceeds those costs, innovators should have sufficient incentives to innovate. Hence the purpose of the constraint is to ensure that socially beneficial innovations continue to be created.

¶66 Subject to the constraint that innovation costs are covered, the antitrust courts should attempt to maximize net social benefits by reducing social costs and raising social benefits.[131] Patentee reward should be traded off for increase in net social benefits. To the extent that patentee reward substantially exceeds innovation costs on a systemic basis, there will be room for antitrust to trade patentee reward for greater social benefits without undermining innovation incentives. This is where the maximization comes into the picture. The handsome patent rents earned in some industries suggest that this is a realistic prospect. The discussion thus far has focused on patentee reward because it is the kind of private benefit that is most directly affected by the patent-antitrust rules. It is important to recall the empirical evidence that the private benefits of innovation significantly exceed patentee reward. This is because of the range of appropriation mechanisms, such as first mover advantage and sales and marketing, at an innovator's disposal. In order to ensure that innovators are not overcompensated, the constraint under the constrained maximization approach should be reformulated to refer to private benefits as opposed to patentee reward. As long as the various kinds of private benefits in aggregate cover innovation costs, innovation incentives will be preserved.

¶67 Theoretically, the constrained maximization approach could work in both directions. It could either reduce or augment private benefits of innovation by adjusting the scope of patent exploitation. The constraint prong only requires private benefits of innovation to cover innovation costs; adjustments to patentee reward are not confined to downward ones. If innovation costs are found to have exceeded aggregate private

---

the impact is usually more indirect and difficult to predict. Therefore, this Article will assume that the principal focus of the courts on the private benefit side of the trade-off is patentee reward, while largely taking the remaining private benefits as given.

[131] In a way, this is similar to Kaplow's model, in which he recognizes the need simultaneously to solve the questions of the optimal scope of the patent-antitrust doctrine and the patent life. *See* Kaplow, *supra* note 1, at 1839–42. When Kaplow poses the question of the optimal patent life, what he seeks to answer is the optimal scope of patent protection. Under the constrained maximization approach, the optimal scope of patent protection, which encompasses both patent term and scope of patent rights, is taken to be that amount of protection which results in a patentee reward slightly exceeding innovation costs minus returns from other appropriation mechanism. While Kaplow trades monopoly loss for patentee reward after a systemically optimal patent life has been determined in advance, the constrained maximization approach attempts to maximize net social benefits and determine the optimal patentee reward at the same time. The reason for the difference in approach is that while Kaplow believes that it is possible to determine a universally optimal amount of patent protection, and then use his patentee reward-monopoly loss ratio as a guidepost to trade more freedom of patent exploitation for shorter patent term and vice versa, this Article assumes that there is no universally optimal amount of protection. Given the diversity of industry and technology covered by the patent system, it is not possible to adopt a one-size-fits-all approach and hope that one optimal ratio will work in every industry and for every technology. The optimization exercise must be done in every case.

benefits, the innovator could conceivably request the court to relax the patent-antitrust rules to allow it to recover innovation costs. There may even be concerns that patentees with failed R&D projects may try to recoup its losses through opportunistic use of antitrust suits.

¶68 When evaluating such a claim by the innovator, the courts should be mindful that private benefits are not measured only up to the time of suit, but over the lifetime of the technology. The court should only relax the scope of patent exploitation under the constrained maximization approach if there is convincing evidence that the innovator will not break even eventually. Moreover, while this may be a theoretical concern, it is unlikely to be a serious one in reality. Past experience shows that patents that raise significant antitrust issues are likely to be highly valuable. Private benefits for these innovations are likely to be high and to exceed their innovation costs. Meanwhile, patents for failed innovations are likely to be of low commercial value and are unlikely to raise antitrust concerns. Therefore, the opportunistic use of antitrust suits to make up for loss-making R&D projects should not be a matter of grave concern.

¶69 One advantage of the constrained maximization approach is that it avoids the pitfall of the prevalent judicial approach that tends to examine a patent exploitation practice in isolation. What the courts tend to do is to decide whether a particular exploitation practice is within the scope of patent rights, however the rights are defined. It does not matter how else the patentee has chosen to exploit is patent. A patentee may have extended the term of royalty payment beyond the patent term, which was deemed illegal by the Supreme Court in *Brulotte v. Thys Co.*,[132] but at the same time may have refrained from engaging in price fixing with its licensee-competitor under the well-known but much-maligned *GE* exception.[133] It may have exceeded the permissible scope with one exploitation practice, but may have stayed clear of the limit with respect to another practice. In the aggregate, the patentee may have reaped no more benefits than that which is necessary to recover its innovation costs. Under the current judicial approach, the extension of royalty payment beyond the patent term is deemed illegal, regardless of the totality of circumstances. This approach overlooks the fact that the size of the patentee reward is not determined by one patent exploitation practice, but by the overall mix of exploitation practices employed by a patentee. As Crane observed, the legality of a patent exploitation practice must be examined in light of the "bundle of rights" that a patentee has chosen to exercise.[134] By taking a holistic look at the social calculus, the constrained maximization approach avoids this pitfall.

¶70 The holistic analysis required by the constrained maximization approach will need to be undertaken within a Rule of Reason framework. It will be difficult to state *a priori* whether a particular patent exploitation practice is pro-competitive or not. This should be unproblematic because few commentators and courts advocate a *per se* approach to the patent-antitrust interface. What may be unfamiliar to some within the antitrust circle is the idea that the legality of one practice depends on the range of exploitation practices adopted by the patentee. This unfamiliarity, or perhaps even unease, is in some ways understandable because in few other areas of antitrust does the legality of one conduct depend on the firm's market behavior as a whole. This is almost the same as saying that

[132] 379 U.S. 29 (1964).
[133] United States v. Gen. Elec. Co., 272 U.S. 476 (1926).
[134] *See* Crane, *supra* note 1, at 256.

resale price maintenance is illegal because the firm has pursued predatory pricing in the relevant market, or perhaps even a different market. However, this interdependence is the necessary consequence of the fact that the resolution of patent-antitrust issues is essentially an optimization process. This is true in both Kaplow's model and the constrained maximization approach proposed in this Article. Optimization is necessary because calibration of the appropriate scope of the patent-antitrust rules requires the ascertainment of the optimal tradeoff between static and dynamic efficiencies. This optimization rationale does not exist in other areas of antitrust law. Antitrust does not strive to discover the optimal tradeoff between price and output for firms. Nor does it attempt to strike a balance between producer and consumer surplus.

III. PRACTICAL APPLICATION OF THE PROPOSED APPROACH

A. A Simplification of the Constrained Maximized Approach

¶71 While a full-fledged constrained maximization approach would be the most theoretically precise and comprehensive, it is unlikely that the courts would be prepared equipped to consider such a wide range of interrelated and difficult-to-quantify factors in their decision. Therefore, it is worth simplifying the analytical framework and focus on the most important components to improve the administrability of the approach.

¶72 The external benefits of innovation include the direct benefits of a new technology for society, the spillover effects, and the possibility of follow-on innovation built upon this technology.[135] The direct social benefits of an innovation are a key component of net social benefits and should clearly be incorporated in the framework. These benefits depend on two factors: the inherent technical contribution of the innovation in satisfying a previously unmet consumer demand and the availability of the technology. The former is unlikely to be affected, at least not directly, by the patent-antitrust rules.[136] The latter will be highly sensitive to the patent-antitrust rules. Stricter patent-antitrust rules will improve the availability of the technology to consumers, which implies that it is likely to be sold at a price closer to the competitive level.

¶73 The treatment of spillover benefits is more difficult. While spillover effects can be substantial, restricting patent protection in order to increase them and facilitate imitation would be a direct affront to the patent system. The quintessential patent right is the right to exclude. Imitation can only be facilitated and spillovers augmented if this right is dramatically curtailed. While consumers would no doubt be better off if more firms could commercialize the technology and offer the product at lower prices, which is a tradeoff that is best left to the patent system to make.

---

[135] The social benefits of innovation, of course, consist of both external and private benefits. Similarly, social costs of innovation consist of both external and private costs of innovation. The private benefits and costs of innovation, however, need not be reconsidered here because under the constrained maximization approach, private benefits are meant to be just sufficient to cover private costs. They are roughly equivalent. We can thus take them both out of the consideration and focus on external benefits and costs.

[136] It is conceivable that patent-antitrust rules may have such a direct and significant impact on patentee reward, and assuming that patentee reward accounts for a large proportion of the private benefits of innovation, that firms will make substantial adjustments to their R&D budgets accordingly. If we further assume that the significant changes in R&D investments will affect the technical merits of innovation, then the patent-antitrust rules may have indirect impact on the inherent technical contribution of an innovation.

¶74 The situation is different for cumulative innovation. Here, the improver does not simply replicate the innovator's technology and undercut its profit margin. The improver actually builds on the technology and creates new technologies or products out of it. Given the substantial benefits of cumulative innovation, it should be retained in the constrained maximization exercise.

¶75 As for the external costs of innovation, they include consumer welfare loss from supra-competitive pricing and the costs of creative destruction. The costs of creative destruction should be left out. While it is true that they impose real costs on society, they seem to be an inevitable price to pay for technological progress. Assuming that these costs were internalized by a monopolist producer in a market, which chooses not to produce a new product because of the potential cannibalization of its revenue from existing products, few would believe that this is a desirable state of affairs.[137] This is what Kenneth Arrow has described as the replacement effect, and is the main reason why, according to him, competition is more conducive to innovation than monopoly.[138] By incorporating the costs of creative destruction in the analysis, we may inadvertently slow down competition in innovation.

¶76 Consumer welfare loss should be the main focus of the constrained maximization approach, however simplified. After all, consumer welfare is the main policy concern of antitrust law. Pared down to a more manageable set of variables, a simplified constrained maximization should focus on maximizing net social benefits with due focus on consumer welfare loss, the direct social benefits of innovation, and the benefits of cumulative innovation.

B. A Further Simplification of the Approach

¶77 Even with the proposed simplification, the antitrust courts may still feel uneasy about the suggested approach. The courts may feel ill-equipped to arrive at estimates of the direct social benefits of innovation or the benefits of cumulative innovation. The courts may feel uncomfortable with the idea of a direct trade-off between patentee reward and the net social benefits of innovation. The courts may further feel unable to arrive at an accurate estimate of the private costs of innovation, which include both R&D costs and opportunity costs. Admittedly, the quantitative terms in which the constrained maximization approach has been formulated thus far may be so far removed from the structure of conventional legal doctrines that the courts may hesitate to adopt it.

¶78 This concern can be addressed if the constrained maximization approach were reformulated in qualitative terms instead. The insights from the approach can be distilled into a number of qualitative guidelines. For example, should the various private benefits of innovation be too difficult to quantify, the courts can reinterpret the inquiry as how important patentee reward is to this innovator's ability to recoup its innovation costs. If it is found that patentee reward is not particularly important to an innovator, the courts will have greater leeway to minimize consumer harm and facilitate cumulative innovation.

---

[137] Frederic M. Scherer, *Technological Innovation and Monopolization* 42 (Am. Antitrust Inst. Working Paper No. 05-07, 2005) at 18–34 (arguing that AT&T withheld certain important innovations for fear of cannibalization effect).

[138] Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention, in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY 609, 619–20 (Nat'l Bureau of Econ. Research ed., 1962).

Similarly, the courts can qualitatively determine in which industries and under what circumstances cumulative innovation is likely to be especially beneficial to society. Fortunately, empirical studies by economists have demonstrated trends by innovation-type and industry with respect to reliance on patent protection for innovation incentives, importance of cumulative innovation, and other inquiries that will prove useful to the courts in applying the constrained maximization approach.

¶79 What follows is a host of guidelines to help the courts apply the insights from the constrained maximization approach. These guidelines focus on the nature, breadth, depth, and source of the technology; the relationship between the technology and the final product; the relationship between the technology and other related technologies; and the industrial environment in which the technology is created. It should be clear that after being reduced to these qualitative guidelines, the approach loses considerable precision. Nonetheless, these guidelines should provide the courts with some helpful guidance on how to apply an appropriate degree of scrutiny to verify innovation incentive arguments.

1. Types of Innovation

¶80 Economists have classified technologies along a number of different dimensions: incremental innovation vs. radical innovation; general-purpose technology vs. specialized technology; serendipitous innovation vs. routinized innovation; and, product innovation vs. process innovation. These classifications focus on the different attributes of a technology. As it turns out, these classifications provide helpful guidance on the qualitative application of the constrained maximization approach.

i) Product Innovation vs. Process Innovation

¶81 The distinction of product vs. process innovation pertains to the nature of the technology. A product innovation results in a new product that substantially improves on existing products, drawing new demand from consumers. A process innovation improves the production process of an existing product. Baumol defines a product innovation as "one that shifts the demand curve for the affected final product to the right, while a *process innovation* is one that shifts the pertinent cost curves downward."[139] The most readily measurable benefit of a process innovation is the reduction of production costs. The social benefit of product innovation is reflected by the additional amount that consumers are willing to pay for the new product. Economists have repeatedly found that product and process innovations exhibit varying degrees of dependence on patent protection as an appropriation mechanism. Specifically, it has been shown in a number of empirical studies that patent protection is a more important method of protecting product innovation than it is for process innovation. In one of the most exhaustive studies on the relative importance of different appropriation mechanisms, then-Professor Richard Levin and his co-authors (the "Yale survey") discovered that patents were generally rated as the least effective appropriation mechanism for process innovations, and that lead time, learning curve advantages, and secrecy were all deemed to be more effective by R&D managers of major corporations.[140] As for product innovations, patents

[139] BAUMOL, *supra* note 6, at 154.

[140] Levin et. al., *supra* note 46, at 794–95. This is corroborated by a recent example. The March 7, 2011

were more effective than secrecy as an appropriation mechanism, but were substantially less so than lead time, learning curve advantages, and sales efforts.[141] 80% of the sampled businesses rated sales and service effort, as an appropriation mechanism, at 5.0 out of a 7-point effectiveness scale. Only 20% of the sampled businesses gave patents the same effectiveness rating.[142]

¶82 These results have been corroborated by subsequent studies. In a 2000 update of the Yale survey, Professor Wesley Cohen and his co-authors (the "Carnegie-Mellon survey") reached largely similar results as those obtained in the Yale survey.[143] Revealingly, the Carnegie-Mellon survey found that direct profit is no longer the predominant motive for patenting. The prevention of rivals from patenting related inventions, the use of patents in negotiations, and the preemption of future infringement suits are now more important motivations for doing so.[144] Industry perceptions of the relative effectiveness of patents as a means of protecting product and process innovations have manifested themselves in patenting behavior. It has been estimated that large European firms applied for patents on only 36% of product innovations and 25% of process innovations.[145]

¶83 If the courts are able to arrive at fairly reliable estimations of the weight of patentee reward in the aggregate private benefits of innovation, they would be able to determine by how much patentee reward can be adjusted without undermining innovation incentives. However, in the likely event that this proves elusive, the courts can determine the relative importance of patentee reward as an appropriation mechanism based on the type of innovation at issue. If the case involves a process innovation, the court can be less concerned that reducing patentee reward would undermine innovation incentives. If the court is confronted with a product innovation, as seems more frequently to be the case from a quick survey of the leading patent-antitrust cases, it will need to proceed more cautiously.[146]

ii) Incremental Innovation vs. Radical Innovation

¶84 The distinction between incremental and radical innovations reflects the depth of the impact of the innovation. An innovation that has a deep impact on future

edition of *The Economist* contains an article about recent advancement in the use of carbon fiber materials in automobiles. *See* Charles Babbage, *A high-fibre diet*, THE ECONOMIST ONLINE, Mar. 7, 2011, *available at* http://www.economist.com/blogs/babbage/2011/03/carbon-fibre_composites&fsrc=nwl. The article describes in detail the innovation in the manufacturing process that has allowed McLaren to mass-produce sports cars that are much lighter and more maneuverable than their peers. This technology may revolutionize how sports cars are made. Even McLaren believed that its competitors could imitate the production technique from the knowledge they can glean from the article, it would be unlikely that it would have disclosed its process innovation in such detail to the reporter.

[141] *Id*. at 795.

[142] *Id*. at 796.

[143] *See generally* Wesley M. Cohen, Richard R. Nelson & John P. Walsh, *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)*, NAT'L BUREAU ECON. RESEARCH WORKING PAPER No. 7552 (2000), *available at* http://www.nber.org/papers/w7552.

[144] *Id*. at 4.

[145] BESSEN & MEURER, *supra* note 67, at 89.

[146] However, even that may no longer be true. Cohen and his co-authors found that since the Yale survey, secrecy has become the most important appropriation mechanism even for product innovations. Cohen et. al., *supra* note 143, at 3.

technological development and society in general would be considered a radical innovation. Otherwise, it would be characterized as an incremental innovation. Examples of radical innovation include the steam engine, penicillin, the airplane, and the Cohen-Boyer patents on recombinant DNA technology that have essentially opened the field of genetic engineering. Needless to say, most innovations are incremental. However, the dichotomy between the two may not be as clear-cut as it would initially seem. The depth of the impact of an innovation most likely varies along a spectrum. Moreover, what seems like a radical innovation often consists of a number of incremental innovations.[147] A prime example is the steam engine, which has been said to consist of numerous technical innovations by various inventors prior to James Watt, such as Thomas Newcomen, who first created a functional steam engine in 1712.[148]

¶85 It should be obvious that radical innovations redound substantially greater social benefits on society, both in terms of direct benefits and the possibility of follow-on innovation. Without attempting to quantify the precise magnitude of the direct benefits, the courts should be able to conclude that a patent exploitation practice that restricts consumer access to a radical innovation will inflict considerable welfare loss on society. Similarly, a patent exploitation practice that limits other innovators' access to the technology for possible follow-on innovations will deny society the substantial benefits of improved utilization of the radical innovation. This is especially significant from a social welfare perspective because many radical innovations require subsequent incremental ones to be successfully commercialized.[149]

¶86 The obvious objection is that by facilitating follow-on innovation and commercialization by other innovators, the courts will undermine the incentives of the original innovator to do the same. This is reminiscent of Kitch's prospect theory on patent protection, the key underlying assumption of which is that the original innovator possesses the best capability and the greatest incentives to undertake the follow-on innovations. This assumption will be examined at greater length later. Suffice it for now to say that a number of leading commentators of intellectual property have challenged this assumption.[150] It would be a mistake for the antitrust courts to sacrifice possible follow-on innovation by rival innovators in the name of a theory of questionable validity.

iii) General-Purpose Technology vs. Specialized Technology

¶87 Related to the incremental vs. radical innovation distinction is the classification of general-purpose technology ("GPT") and specialized or non-general-purpose

[147] Bart Verspagen, *Innovation and Economic Growth*, *in* THE OXFORD HANDBOOK OF INNOVATION 486, 494 (Jan Fagerberg et al., eds., 2005).

[148] BOLDRIN & LEVINE, *supra* note 20, at 1–5. *Contra* George Selgin & John Turner, *James Watt As Intellectual Monopolist: Comment on Boldrin and Levine*, 47 INT'L ECON. REV. 1341 (2006) (defending Watt's contributions to invention of steam engine).

[149] BAUMOL, *supra* note 6, at 57 ("It is more plausible that this division of the work of technical progress will continue, with the independent entrepreneur providing many if not most of the more revolutionary and heterodox contributions, while the routine innovation activities of the oligopoly corporations take those contributions and improve and extend them.").

[150] SCOTCHMER, *supra* note 12, at 150–51 (propounding model of innovation in which ideas are endogenous and may come from many sources); *Economics of Improvement*, *supra* note 68, at 1050 (calling it unrealistic to expect one party to identify all possible uses of new invention).

technology.[151] Examples of GPT include electricity and computer.[152] A recent addition of GPT is three-dimensional printing, which allows one to produce a three-dimensional object from scratch. According to *The Economist*, it is likely to revolutionize the manufacturing industries by allowing just about anyone to create a complex object at home. This emerging technology is expected to "reset the economics of manufacturing. Some believe it will decentralise the business completely, reversing the urbanisation that accompanies industrialisation."[153] Whether a technology is general-purpose or not is related to its breadth, in particular, its range of applications. A GPT is a technology that has a wide range of possible applications. A GPT is likely to be radical and redounds considerable benefits for society. The above discussion concerning how the antitrust courts should take into account the radical nature of an innovation thus equally applies to GPT. However, a radical innovation need not be general-purpose. AZT, the first drug that was effective in retarding the replication of HIV, was no doubt a radical innovation. Its range of application, however, is rather limited.

¶88 While much of the discussion about GPT in the academic literature has focused on its contribution to economic growth, the breadth of a technology may also have implications for the patent-antitrust interface. Here, it is necessary to invoke the intuition behind Frank Ramsey's theory on optimal commodity sales tax.[154] Ramsey argues that the optimal way to achieve a certain sales tax revenue target is to levy it on many goods as opposed to concentrating the tax incidence on one good. Here, the idea is that with respect to a GPT, society would be better off if innovation costs are recouped from a wide range of applications as opposed to a few of them. The welfare loss would be considerably smaller. Restricting the patentee reward of the innovator from one application would not undermine innovation incentives so long as the patentee could recover the shortfall from other applications. And the aggregate social welfare loss would be smaller than if the patentee attempted to recover its innovation costs from one application alone. Multiple-channel recoupment is only possible if the technology has a broad range of applications. That would be the case for GPT. Therefore, when facing an antitrust case involving a GPT, the courts should display greater readiness to trade patentee reward for gains in social welfare. In fact, it would serve social good for the courts to steer the patentee to diversify its recovery channels by restricting its patent exploitation practices within each application.

iv) Serendipitous Innovation vs. Routinized Innovation

¶89 This distinction pertains to the source and the process of innovation. It refers to the source of inspiration for innovation and how the process of innovation is managed. Ideas

[151] The term "general purpose technology" was coined in 1995 by Professors Bresnahan and Trajtenberg. Timothy Bresnahan & Manuel Trajtenberg, *General Purpose Technologies: 'Engines of Growth'*, 65 J. ECONOMETRICS 83 (1995).

[152] HELPMAN, *supra* note 11, at 51.

[153] Technology, *Print me a Stradivarius*, THE ECONOMIST, Feb 10, 2011, *available at* http://www.economist.com/node/18114327 (last visited Dec. 1, 2012).

[154] Frank P. Ramsey, *A Contribution to the Theory of Taxation*, 37 ECON. J. 47 (1927). For another application of the Ramsey theory to a patent-related problem, see Ayres & Klemperer, *supra* note 88**Error! Bookmark not defined.**.

for serendipitous innovation come from unexpected and unpredictable sources.[155] Serendipitous inventions may happen accidentally or follow a flash of genius. They do not come about simply because a firm devotes all the resources at its disposal to a technical challenge. They may be produced by innovators of all shapes and sizes, from individuals to small firms to multinational corporations. Meanwhile, routinized innovation is largely the province of large corporations. The routinization of innovation is the standardization and regularization of the innovative process.[156] It transforms innovation from "a fitful and uncertain discovery process into something closer to a routine internal matter governed by the bureaucratic and managerial procedures that also control many of the other activities of the large corporation."[157] Serendipitous innovation has been responsible for some of the most important inventions and discoveries,[158] while routinized innovation focuses on incremental improvement in product quality, reliability, and user-friendliness.

¶90 For our purpose, the most relevant characteristic of routinized innovation as opposed to serendipitous innovation is that the former is much less uncertain. By regularizing innovation activities, and by performing R&D on a regular basis and in a sufficient amount, routinization "increases the likelihood that *some* successful innovations will emerge at reasonably regular intervals."[159] The relatively low uncertainty for this type of innovation means that the adjustment that needs to be made to private benefits to reflect the risk of innovation will be smaller. Another conclusion of Baumol's concerning routinized innovation also has important implications for the patent-antitrust interface. He argues that in a contestable, oligopolistic market with low barriers to entry, firms will not make any more than normal profit plus depreciation contributions that will allow them to recover their R&D investment outlays.[160] This would imply that for these firms, patentee reward is barely sufficient to cover innovation costs. Assuming that other sources of private benefits are small, there will be little room for the courts to tinker with patentee reward without adversely affecting innovation incentives.

2. Correspondence Between Technology and Product Boundaries

¶91 The correspondence between technology and product boundaries focuses on the relationship between an innovation and its eventual commercial manifestation. It refers to the fact that while some products, such as drugs, consist of one or a handful of technologies,[161] other products, such as most consumer electronics, electrical appliances, complex machinery, automobile, etc. are made up of hundreds or even thousands of technologies. To put it slightly differently, the technology and product boundaries for innovations such as drugs coincide, while the two boundaries for consumer electronics and machinery do not. As Burk and Lemley noted, "[m]uch of the conventional wisdom

[155] ANDREW HARRISON, MAKING AND THINKING: A STUDY OF INTELLIGENT ACTIVITIES 67–68 (1978).
[156] BAUMOL, *supra* note 6, at 2–3.
[157] *Id*. at 30.
[158] For instance, four of the best-selling drugs worldwide in 2005, Cisplatin, Librium, Taxol, and Thorazin, were discovered completely by chance. BOLDRIN & LEVINE, *supra* note 20, at 230.
[159] BAUMOL, *supra* note 6, at 32.
[160] BAUMOL, *supra* note 6, at 31.
[161] Lemley, *Industry-Specific Policy*, *supra* note 13, at 6.

in the patent system is built on the unstated assumption of such a one-on-one correspondence."[162] The presumed direct correspondence between innovation and social welfare applies to discrete innovations such as pharmaceuticals. For technologies that demonstrate considerable complementarity, such a direct relationship can no longer be presumed. These technologies would only be socially beneficial if they could be utilized together with their complements.

¶92 The most important implication of technological complementarity for the patent-antitrust interface is that the ability of an innovation to create value for society crucially depends on the availability of other complementary technologies. Should the owner of one of the technologies decides to withhold its own contribution, many other technologies will be rendered useless and society will stand to lose tremendous value. This was vividly illustrated in the Blackberry saga when NTP, Inc., the owner of one of the patented technologies in Blackberry, threatened to shut down the whole Blackberry service.[163] It was eventually able to extract more than $600 million of damages from RIM, the developer of Blackberry.[164] This saga was also a painful reminder of the possibility of holdup for complementary technologies. When dealing with these technologies, the antitrust courts must be mindful of the enormous loss of social benefits that may result if one patentee is allowed to hijack the entire product by opportunistic behavior.

3. Cumulative Innovation

¶93 Cumulative innovation concerns the relationship between the initial innovation and future technological development, or more specifically, the follow-on innovation that may be developed based on the initial innovation. Cumulative innovation would not be an issue at all if the initial innovation were a dead-end technology that could not be further improved upon. In reality, however, very few technologies belong to this category. Paper clips, which were at one point patented, come to mind, although even they have been improved upon by encasement of the metal with plastic.[165] Most technologies, meanwhile, carry considerable potential for substantial improvement. Bicycles, which when first invented in the nineteenth century were fairly simple machines, have been substantially enhanced in the ensuing years through the deployment of new materials such as carbon fiber and advancement in mechanics.[166] The importance of cumulative innovation for technological progress has been affirmed by numerous scholars.[167] Industries for which cumulative innovation is particularly important include biotechnology, computer software, and computer hardware, some of the pivotal industries

---

[162] Burk & Lemley, *supra* note 20, at 1590.

[163] BESSEN & MEURER, *supra* note 67, at 47–50.

[164] *Id*. at 49.

[165] HENRY PETROSKI, INVENTION BY DESIGN: HOW ENGINEERS GET FROM THOUGHT TO THING 8–42 (1996).

[166] *See generally* DAVID V. HERLIHY, BICYCLE (2004).

[167] *See, e.g.,* Edmund Kitch, *The Nature and Function of the Patent System,* 20 J.L. & ECON. 265 (1977) [hereinafter *Nature and Function*]; Merges & Nelson, *supra* note 65, at 870; *Economics of Improvement*, *supra* note 65, at 997; Scotchmer, *supra* note 59, at 29; Mokyr, *supra* note 68, at 1; BAUMOL, *supra* note 6, at 51; JAFFE & LERNER, *supra* note 6, at 18; BOLDRIN & LEVINE, *supra* note 20, at 147, n.7; Landes & Posner, *supra* note 7, at 318.

to the modern-day economy.[168] Scholars have also noted that overly expansive patent protection retards cumulative innovation.[169] The challenge for the patent-antitrust interface is to facilitate the development of subsequent improvements while ensuring that the initial innovators are adequately compensated for their innovation costs.

¶94 Before delving into this policy challenge, the first question we need to answer is whether the initial innovators or the rival innovators are best positioned to come up with new improvements. The policy challenge would be much less severe if the initial innovators were in most cases better equipped to improve the initial technology, or at least to locate innovators who are able to do so. The challenge would become considerably more complex if the sources of improvement were often unexpected and unpredictable, defying coordination by the initial innovator. From a social perspective, the policy challenge is two-fold: first, society would want to secure the improvements so long as they redound net social benefits, and second, society would want to minimize the costs of creating these improvements. Specifically, society would want to avoid duplicative R&D by competing innovators pursuing the improvement. The detriment of duplicative R&D is exacerbated by the lack of an independent invention defense under patent law. Once an improver has secured a patent for an improvement, its rivals will be barred from practicing that improvement even if they had independently created it.

¶95 Views on this dual policy challenge have largely fallen into two camps. Edmund Kitch, the leading proponent of the prospect theory, believes that the initial innovator is best positioned either to come up with further improvements or to identify innovators able to do so. To Kitch, the main purpose of the patent system is not to provide incentives for innovation, but to encourage further commercialization of existing ideas and the exploration of improvements by granting prospective property rights over these as yet undiscovered ideas.[170] The initial innovator is hence tasked with coordinating commercialization of the existing innovation and the search for further improvements, with a view toward avoiding duplicative investments. If this view was correct, antitrust law should refrain from interfering with patent exploitation, at least as pertaining to cumulative innovation, since whatever licensing decisions taken by the patentee can be presumed efficient. Patentees would be presumed to know best how to commercialize and pursue further improvements. The prospect theory would hence leave a very limited role for antitrust to police patent exploitation practices, at least in relation to licensing.

¶96 Unsurprisingly, subsequent commentators have challenged Kitch's rosy view of the efficiency, if not omniscience and good faith, of the initial innovator. Critique of the theory has come from a variety of angles. First, it has been questioned whether the innovator is truly better positioned to pursue improvements than his rivals. Even if the innovator and its rivals were equally capable of producing improvements, the odds still hugely favor the rivals due to their numerical advantage.[171] For the incumbent innovator to out-innovate all its rivals, its technical edge over them would have to be enormous. Second, from a practical perspective, it has been argued that even if the initial innovator possessed perfect information to identify the most appropriate improver, the initial

[168] *See* BOLDRIN & LEVINE, *supra* note 20, at 17–21 (explaining how cumulative, collaborative innovation has allowed emergence of open-source software).

[169] JAFFE & LERNER, *supra* note 6, at 48–49; *Economics of Improvement*, *supra* note 65, at 998.

[170] *See Nature and Function*, *supra* note 181, at 275–80.

[171] SCOTCHMER, *supra* note 12, at 146–47.

innovator would still incur considerable transaction costs in order to reach an agreement with the improver.[172] Negotiating a licensing agreement can be enormously time-consuming and costly for the innovator. And transaction costs are not confined to the innovator's identification and negotiation with potential improvers. It is entirely possible for an improver to come up with a significant improvement to the initial innovation not knowing that the innovation is patented. The improver will then need to identify the patentee and negotiate with it. This task will be even more arduous if the improvement implicates more than one patent, as is often the case.

¶97 Transaction costs are compounded by the problem of uncertainty. Negotiating an agreement in light of the considerable uncertainty in technological improvement will necessitate higher transaction costs. The uncertainty would be reduced, if not eliminated altogether, if the innovator and the improver begin negotiations after the improvement has been made. There would be no more uncertainty about the success of the development of the improvement. However, *ex post* negotiations expose the improver to potential holdup by the innovator, especially if the improvement falls squarely within the scope of the initial patent. Commentators have noted that efficient negotiations over technological development must take place *ex ante* to avoid strategic behavior on the part of the initial innovator, which will result in under-compensation of the improver and a socially sub-optimal level of improvement.[173]

¶98 The more fundamental critique of the prospect theory is the unrealistic assumption of perfect information on the part of the innovator. Many commentators have argued that new ideas are likely to spring up from various, often unexpected, sources. It is almost impossible for a central planner lacking perfect information to identify all the possibilities. Under the evolutionary model of innovation, innovation is "best understood as a quasi-Darwinian process—a process almost of trial and error in which the market selects from among diverse approaches whose relative promise cannot be assessed in advance. This approach implies that a multiplicity of sources of inventive activity is superior to a centralized process directed by the patentee."[174] If one subscribes to this view of innovation, it seems obvious that the prospect theory will lead to a sub-optimal level of improvement. Merges and Nelson have argued that it is to be expected that "a single rightholder . . . *under*develop—or even ignore totally—many of the potential improvements encompassed by their broad property right."[175]

¶99 The obvious response to the evolutionary model is that theoretically speaking, there is no difference between having ten firms vying for innovation in an industry as opposed to unifying all these ten firms into one and having them operate as separate research divisions. Unfortunately, there are reasons to believe that the two scenarios are not equivalent and that inter-firm diversity would be more conducive to innovation than intra-firm diversity.[176] Managerial diseconomies impose a limit on how large a firm can grow before its management can no longer keep abreast of its internal operations and

[172] *Economics of Improvement*, *supra* note 68, at 1053–55.

[173] SCOTCHMER, *supra* note 12, at 146-47; Daniel G. Swanson & William J. Baumol, *Reasonable and Nondiscriminatory (RAND) Royalties, Standards Selection, and Control of Market Power*, 73 ANTITRUST L.J. 1, 7–10 (2005).

[174] LANDES & POSNER, *supra* note 7, at 318.

[175] Merges & Nelson, *supra* note 65, at 873–74.

[176] LANDES & POSNER, *supra* note 7, at 318–19.

further internalization of outside productive activities becomes inefficient. A firm that encompasses the innovative capacity of an entire industry is likely to have gone beyond that point.

¶100 The prospect theory presumes rational behavior on the part of the innovator. Lemley has questioned this presumption, arguing that the initial innovator may irrationally eschew licensing for non-economic reasons.[177] It may be due to a bitterly fierce rivalry between two competing firms. It may be due to distrust of a rival's good faith to abide by the licensing agreement. It may be due to fear, rational or irrational, that the licensed technology will somehow assist rivals in ways that are unforeseen by the innovator. For example, the Wright brothers steadfastly refused to license their patents on aircraft technology to anyone, and actively brought litigation against other aircraft developers.[178] In fact, it took a war (the First World War) and intervention by the Secretary of the Navy to break this gridlock and bring all the related patents together so that their complementary technical potential could be fully exploited. Meanwhile, there was no government intervention to overcome the personal animosity between Guglielmo Marconi, the patent holder for the diode technology—a crucial component of the early radio—in the U.S., and Lee de Forest, the inventor of the triode, which was a significant improvement over the diode.[179] Their failure to settle their differences withheld from the world the benefits of this significant improvement for years.[180] These anecdotes lend support to Lemley's assertion that proponents of the prospect theory have a misplaced faith in the knowledge, rationality, and good faith of the innovator to serve the social good.[181]

¶101 All these arguments lead to the inevitable rejection of the prospect theory. In a way, the debate between the proponents and the opponents of the prospect theory is reminiscent of the long-run conflict between patent and antitrust. It may be recalled that while the fundamental belief of antitrust is that competition generates more innovation over the long run, patent law seems to hold the contrary Schumpeterian view. It was argued then that the abiding faith of antitrust in competition means it prefers rivalry over exclusivity in innovation. Here, we are facing the same dilemma. The aforementioned theoretical arguments notwithstanding, whether the initial innovator or potential improvers are best positioned to generate improvements is ultimately an empirical question, the answer to which will perhaps remain elusive for quite some time. In the end, absent overwhelming evidence and arguments negating the merits of competition, antitrust should favor competition over exclusivity. And the evidence and arguments in support of the prospect theory are far from overwhelming. Finally, Kitch himself admits that prospect rights may create substantial market power and may result in supra-competitive pricing.[182] While he identified the problem, he offered no solution for it.[183] Given the focus of antitrust on consumer welfare, this probable consumer harm further militates against the prospect theory.

---

[177] *Economics of Improvement*, *supra* note 68, at 1059–61.
[178] JAFFE & LERNER, *supra* note 6, at 50.
[179] *Id*. at 51.
[180] *Id*.
[181] *See* Mark A. Lemley, *Ex Ante versus Ex Post Justifications for Intellectual Property*, 71 U. CHI. L. REV. 129, 148 (2004) [hereinafter "*Ex Ante Ex Post*"].
[182] *Nature and Function*, *supra* note 181, at 274.
[183] *Economics of Improvement*, *supra* note 68, at 1047.

¶102 A rejection of the prospect theory means that competition among potential improvers takes on added importance. Follow-on innovators must have sufficient technological access to allow them to develop their improvements. At the same time, it remains important to preserve the incentives of the initial innovator. The dual policy challenge highlighted earlier presents an inevitable internal conflict. The question is how should the balance be struck between these conflicting considerations. This question is particularly important for the patent-antitrust interface because of the considerable social benefits of cumulative innovation.

¶103 While antitrust needs to take an active role in securing the benefits of cumulative innovation, especially in situations where patent law does not provide the appropriate solution, it remains to be seen how it should be done. Economists distinguish three scenarios of cumulative innovation: first, where an initial innovation spawns numerous follow-on innovations; second, where many initial innovations, mostly in the form of research tools, are needed to create a single second-generation innovation; and third, where the subsequent innovation represents an improvement of the initial one and yet the two compete with each other.[184] The second and the third scenarios have been called, respectively, the "anticommons" and the "quality ladder" models of cumulative innovation.[185] This author will call the first scenario the "trunk-branch" model for lack of a better name in the literature. The three scenarios present disparate challenges for the patent-antitrust interface.

¶104 Of all three models, the quality ladder model probably presents the least pressing need for antitrust intervention. If the improvement is a radical one, it may benefit from the doctrine of reverse equivalents and escape infringement. To the extent that an innovation issue can be resolved internally within patent law, that solution is to be preferred. Even if that doctrine or other patent internal solutions are unavailable, antitrust still must proceed cautiously. There are a number of reasons for this. First, in most quality ladder scenarios, subsequent improvements directly compete with the initial innovation. This competition will allow the improver to erode the innovator's profit, which may be essential to the innovator's recoupment of its innovation costs. In fact, the erosion of profit might be so severe that future innovators will lose their incentives to innovate. Second, it would seem that the denial of the improvement would not deprive society of a product altogether; it would only leave us with a less advanced version of it. This scenario is different from the other two scenarios where denied access to the innovator's technology may preclude the creation of subsequent products altogether. Given that the first-generation technologies in those two models are often basic science research or research tools such as gene fragments, from which society derives no direct benefit, the development of second-generation products is doubly important to society. Moreover, as previously pointed out, the welfare effects of a product innovation are ambiguous. By choosing to interfere to facilitate the improver's access to the innovator's technology under the quality ladder scenario, it is conceivable that antitrust may be responsible for lowering the private benefits of innovation without producing any net social benefits.

[184] SCOTCHMER, *supra* note 12, at 132-33.

[185] *Id*. at 146. And for ease of reference, the first scenario will be referred to as the "trunk-branch" model of cumulative innovation.

¶105 Under the trunk-branch model, the beneficial social effects of the widespread utilization of the "trunk" technology are significant. So long as the initial innovator is compensated for its innovation costs, antitrust should facilitate technology diffusion. This is particularly so because the technology has a myriad of applications. However, given that many of the trunk technologies belong to basic scientific research, which have no independent commercial value, licensing to second-generation users would be the primary means of recouping the innovator's costs. Here, the Ramsey intuition is again relevant. From a social welfare perspective, society would be better off if the innovator distributed its cost recovery among a wide range of uses, while minimizing the distortionary effect within each use. The patent-antitrust rules should reflect this by limiting the scope of patent exploitation, so that the innovator can be steered to diversify its recovery channels.

¶106 The term "anticommons" was originally coined by Professors Michael Heller and Rebecca Eisenberg in 1998 as a play on the term "commons."[186] Anticommons present a particularly tricky situation because of the significant possibility of holdup by any one of the owners of the research tools. Private parties often resolve the situation themselves through patent pools or packaged licensing. This is another version of the complementarity problem that was mentioned previously. There it involved a number of components going into a final product and here a number of complementary research tools needed for further research and production of second-generation products. It was argued then that antitrust has a robust role to play to minimize instances of holdup. The same conclusion applies here.

4. Appropriability Conditions in the Industry

¶107 Appropriability refers to the conditions of an industry that allow the innovator to reap the economic benefits of its innovation. It "summarizes the possibilities of protecting innovations from imitation and of reaping profits from innovative activities."[187] As discussed in Section **Error! Reference source not found.**, there is a whole range of appropriation mechanisms at the innovator's disposal to allow it to profit from its innovation. To the extent that an innovator relies on other appropriation mechanisms, patentee reward will become relatively less important. In fact, Mansfield and Levin conclude that in many industries, patent protection was not even necessary to attract innovation.[188]

¶108 Economists have discovered that firms in different industries rely on patent protection for innovation incentives to varying extents. Levin and his co-authors classified the one hundred manufacturing industries surveyed in their study into three clusters.[189] The first cluster, consisting of food products and metal products among others, does not seem to rely on any particular appropriation mechanism except for sales

[186] Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCI. 698, 698–99 (1998).

[187] Franco Malerba, *Sectoral Systems: How and Why Innovation Differs Across Sectors*, in THE OXFORD HANDBOOK OF INNOVATION 380, 382 (Jan Fagerberg et al. eds., 2005).

[188] Mansfield et. al., *supra* note 76, at 915; Mansfield, *supra* note 76, at 174; Levin et. al., *supra* note 46, at 796–97.

[189] Levin et. al., *supra* note 46, at 801.

and marketing.[190] The second cluster depends on a wider range of appropriation mechanisms, such as first-mover advantages, learning curve advantages, and secrecy. The third cluster, which includes pharmaceuticals and chemicals, is the only group that utilizes patent protection extensively. These results are consistent with Mansfield's findings, which show that patent protection plays a much more important role in appropriating returns in the pharmaceuticals and chemicals industries than in electronics and machinery.[191]

¶109 These findings can be used as a basis for the courts to determine how much patentee reward can be traded off for net social benefits in a particular industry. This is especially useful if quantitative determination of the weight of patentee reward relative to overall private benefits of innovation cannot be made. For example, when faced with a patent-exploitation practice arising in a first-cluster industry, the courts can be less concerned about preserving innovation incentives and pay closer attention to welfare loss. The relative lack of reliance on patent protection will give the courts more room for restricting the scope of patent exploitation under antitrust law without undermining innovation.

5. Summing Up

¶110 The foregoing discussion is not meant to provide exhaustive guidance for every type of patent-antitrust case. What the practical guidelines in this Section are meant to do is to highlight the most relevant issues for the antitrust courts to consider in a patent-antitrust case. The central message is that while the courts should remain mindful of maintaining innovation incentives, achievement of that goal does not require complete deference to patent rights. Given that innovation incentives are preserved whenever innovation costs are covered, and that patentee reward is but one of the myriad appropriation mechanisms available to innovators, there is in fact significant leeway for the courts to protect and advance social welfare without undermining innovation incentives. So long as innovative incentives are unaffected, the antitrust courts should aim to maximize net social benefits, which can be accomplished by enhancing social benefits or reducing social costs, or both. The focal points are the facilitation of cumulative innovation and the minimization of consumer welfare loss.[192]

¶111 Consumer welfare has always been a central concern of antitrust. However, when facing a patent-antitrust case, the antitrust courts often seem too ready to defer to patent policy for fear of undermining innovation incentives. One of the main goals of this Article is to clarify our understanding of innovation incentives and to reveal the excessively pro-patent bias of the conventional wisdom. Two leading commentators

[190] *Id*. at 802.

[191] Mansfield, *supra* note 76, at 174.

[192] A theoretical argument can be made that, given the importance of spillovers of innovation as a source of social benefits, antitrust law should facilitate if not expand spillovers of innovation. This can be accomplished perhaps by aggressive prosecution against unilateral refusal to license. There is much room for debate as to whether this is feasible or desirable as a matter of public policy. Given the state of the current antitrust jurisprudence, which is highly hostile to imposing a duty to deal in the intellectual property context, and the prevailing views that patent rights are a kind of property rights and that imposing a duty to deal is a direct affront to private property, the imposition of a duty to license absent any independent antitrust violation, solely to facilitate spillovers, has close to no chance of acceptance in the antitrust circle. Hence, this Article will not pursue this idea.

have noted that "[l]egal scholars have failed to appreciate that unconstrained monopoly pricing is not a cost-justified means of rewarding patentees."[193] In light of this improved understanding of innovation incentives, it would seem that reluctance to tinker with patentee reward can no longer be justified in the name of preserving innovation.

¶112 The importance of cumulative innovation has not been as widely recognized within the antitrust circle as it has by patent law scholars. Given its importance in generating economic growth, antitrust should pay closer attention to it. Antitrust professes to take dynamic efficiency considerations seriously, but thus far it seems to have done so only with a certain subset of those considerations, namely the provision of incentives for the initial innovation. Other dynamic efficiency considerations have been largely overlooked. This is a grave mistake. Rectifying this mistake will require the antitrust courts to make some difficult judgments. However, there is no reason to believe that it is beyond the capability of these courts.

¶113 A further message of the foregoing discussion is the need for the courts to distinguish between different types of innovation in different industries. As has been repeatedly pointed out by patent commentators in recent years, not all innovations are the same. There is a high degree of heterogeneity among innovations that the antitrust courts need to consider. These distinctions include the nature, breadth, depth, and source of innovation; its relationship with the final product; and its relationship with other technologies. Lemley advocated an industry-specific approach to the patent-antitrust interface.[194] This Article asks the courts to go one step further and focus on the various attributes of the technology at issue. These attributes will help the courts predict the effects of its decision on the various parameters in the constrained maximization exercise.

¶114 Lastly, it is important to clarify that this Article does not advocate drastic restrictions of patent rights. While some of the theoretical discussions in Section **Error! Reference source not found.** suggest that patentee reward can be dramatically reduced without undermining innovation incentives (assuming that private benefits of innovation under the current patent system significantly exceed innovation costs for most innovations), which may lead the reader to surmise that this author advocates aggressive curtailment of patent rights by antitrust, the extent of intervention proposed in this Article is much more modest. Substantial improvement of net social benefits can be achieved with minor restrictions of patent rights. As Ayres and Klemperer observe, "[t]he last bit of monopoly pricing produces large amounts of deadweight loss for a relatively small amount of patentee profit."[195] Thus, small restrictions of patent rights and monopoly pricing will not significantly undermine innovation incentives.

¶115 The theoretical discussion in Section **Error! Reference source not found.** admittedly lends support to more substantial modifications of patentee reward. From a social efficiency perspective, such modifications may be desirable. That, however, would be a usurpation of the fundamental tradeoff between static and dynamic efficiencies struck by Congress in designing the current patent system. As crudely as that

[193] Ayres & Klemperer, *supra* note 88, at 987.

[194] *Industry-Specific Policy*, *supra* note 13.

[195] Ayres & Klemperer, *supra* note 88, at 987. This is so because of the offsetting gains in profit as a result of a decrease in price (and the resultant increase in quantity sold). The resultant loss in patentee profit should be small. Meanwhile, the improvement in social welfare, in the form of elimination of deadweight loss, can be substantial. Therefore, inducing infringement to cause a small drop in price would be socially efficient.

tradeoff was struck, any major changes to it should be made by Congress, and not the courts. Calls for reform of the patent system and rollback of the increasingly expansive patent rights abound.[196] There is ample evidence that the patent system as it currently operates is broken.[197] Yet, these defects of the patent system do not justify massive interference on the part of antitrust law. Patent law's problems remain its own to solve. Drastic alterations to patent rights should not be made by the antitrust courts.

IV. AN APPLICATION OF THE PROPOSED APPROACH: UNILATERAL REFUSAL TO LICENSE

¶116 To demonstrate the application of the proposed approach, this Section will focus on how it applies to the analysis of a unilateral refusal to license a patent.[198] Unilateral refusal to license a patent is one of the most controversial areas of the patent-antitrust interface. To those staunch defenders of intellectual property rights, imposing a duty to license on a monopolist patentee constitutes a direct affront to private property. In some ways, compulsory licensing has become a taboo in the U.S. antitrust circle.[199] Mention of compulsory licensing seems to be greeted with skepticism, if not derision.

¶117 The most common defense against imposing a duty to license is that it will undermine innovation incentives. The conventional wisdom is that if an innovator is forced to share its creation with competitors, it will be less likely to invest in innovations in the future. Other innovators will be similarly deterred. By refusing to order compulsory licensing, the courts will effectively force rivals to develop their own

---

[196] *See, e.g.*, JAFFE & LERNER, *supra* note 6, at 178–205 (making detailed and varied proposals for patent reform); BESSEN & MEURER, *supra* note 67, at 235–53 (suggesting ways to improve notice function of patents); BOLDRIN & LEVINE, *supra* note 20, at 23–67 (advocating abolition of patent system altogether).

[197] JAFFE & LERNER, *supra* note 6, at 25–77; BESSEN & MEURER, *supra* note 67, at 46–94.

[198] Hovenkamp distinguishes four types of compulsory licensing scenario: (1) concerted refusal to license; (2) conditional refusal to license (such as tying and exclusive dealing); (3) compulsory licensing imposed as a remedy for an independent antitrust violation; and, (4) unilateral refusal to license as an antitrust violation itself. This Article will focus on the last two instances. The competitive harm of a concerted refusal to license stems mainly from the concerted aspect of the conduct rather than the denial of access to a patented technology. It does not raise as immediate concerns about innovation incentives as the other instances of compulsory licensing. Conditional refusal to license, it is argued, is better analyzed under the rubric of the alternative characterizations. This is because the doctrine and the economic understanding about conduct such as tying and exclusive dealing seems to be more sophisticated than that concerning unilateral refusal to deal. Analysis under that rubric is likely to be better informed. Moreover, treating conditional refusal to license as instances of tying or exclusive dealing carries the added benefit of avoiding the stigma that is usually attached to compulsory licensing within antitrust law. It will probably help to allow the courts to approach the issue in a more objective manner. However, given the dearth of refusal to license cases that belong to these last two categories, cases falling under the first two may be discussed where relevant.

[199] For a general view within the literature on unilateral refusal to license and compulsory licensing and the related doctrine of essential facilities, see Howard A. Shelanski, *Unilateral Refusal to Deal in Intellectual and Other Property*, 76 ANTITRUST L.J. 369 (2009) (arguing that IP and other property should be subject to similar treatment as far as refusal to deal is concerned); Brett M. Frischman & Spencer Weber Waller, *Revitalizing Essential Facilities*, 75 ANTITRUST L.J. 1 (2008) (taking a more pro-enforcement view of refusal to deal); Dennis W. Carlton, *A General Analysis of Exclusionary Conduct and Refusal to Deal—Why Aspen and Kodak Are Misguided*, 68 ANTITRUST L.J. 659 (2001) (critiquing Supreme Court decisions in *Aspen* and *Kodak* and describing limited circumstances under which refusal to deal could cause consumer harm); Glen O. Robinson, *On Refusal to Deal with Rivals*, 87 CORNELL L. REV. 1177 (2002); Robert Pitofsky et al., *The Essential Facilities Doctrine Under U.S. Antitrust Law*, 70 ANTITRUST L.J. 443 (2002); Spencer Weber Waller, *Areeda, Epithets, and Essential Facilities*, 2008 WIS. L. REV. 359 (2008); Philip Areeda, *Essential Facilities: An Epithet in Need of Limiting Principles*, 58 ANTITRUST L.J. 841 (1989).

technologies, which leaves society better off. Therefore, compulsory licensing undermines the innovation incentives of the patentee, the putative licensees, and potential future inventors. This innovation incentive argument presumes that any reduction in patentee reward will reduce innovation incentives, and that the putative licensees possess the technological capacity and commercial interest to innovate.

¶118 As has been repeatedly asserted in this Article, the former presumption is not true. So long as the innovator is compensated for its innovation costs, including the opportunity cost of innovation, innovation incentives will be preserved. Prominent commentators and empirical evidence have affirmed this view. Hovenkamp has advocated a nuanced and balanced stance on compulsory licensing:

> One corollary of the principle that an IP right is simply property is that no special deference is due to the IP laws when courts fashion remedies for proven antitrust violations. For example, ordering compulsory licensing for a proven antitrust violation is no different than fining a firm or ordering divestiture of a plant. While we do not want to deter innovation, we do want to deter antitrust violations either.[200]

In fact, it seems that patent lawyers and economists are less apprehensive about compulsory licensing than antitrust lawyers themselves.[201] Levin and his co-authors concluded that compulsory licensing does not undermine innovation incentives in any significant manner.[202] Scherer found that "the substantial amount of evidence now available suggests that compulsory patent licensing, judiciously confined to cases in which patent-based monopoly power has been abused . . . would have little or no adverse impact on the rate of technological progress . . . ."[203] He further referred to specific conversations with executives of Xerox, which had been subject to compulsory licensing order in the 1970s, that refuted the popular belief that compulsory licensing had adversely affected the firm's R&D.[204] Scherer goes so far as to conclude that "a massive antitrust attack on business firms' use of patents to monopolize markets or enhance profit returns appears to have had negligible adverse consequences for the vigor of innovative activity in the United States."[205]

¶119 While this Article in no way advocates a massive attack on patent rights, this author believes that imposing a duty to deal in limited instances will enhance social welfare. Preservation of innovation incentives cannot be treated as a trump card in every refusal to license case. It must be scrutinized with care to verify that those incentives will truly be undermined if compulsory licensing is ordered. In particular, the applicability of the innovation incentive argument depends on market conditions. There are at least two circumstances under which the argument is unlikely to be valid: (1) aftermarkets; and, (2) patent holdup.

---

[200] *Age of IP Expansion*, *supra* note 16, at 238.

[201] Burk & Lemley, *supra* note 20, at 1667; BAUMOL, *supra* note 6, at 215–16 ("if the price is right, it is never most profitable for the owner of such technology to prevent others, even its most direct competitors, from using it.").

[202] Levin et al., *supra* note 46, at 804.

[203] F.M. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 456–57 (1980).

[204] *Antitrust, Efficiency, and Progress*, *supra* note 25, at 1016.

[205] *Antitrust, Efficiency, and Progress*, *supra* note 25, at 1018.

A. Aftermarkets

¶120 As suggested earlier, the conventional argument against a duty to license presumes that reduced patentee reward will undermine innovation incentives and that the spurned rivals would have developed their own technologies. The extent to which the latter is true depends on a variety of factors. Of particular interest to us is the consideration about the level on which the rivals compete with the innovator. If the rivals do not compete in the primary technology market and only compete in the aftermarkets in the provision of derivative products or services, denying them access to the technology will be unlikely to encourage innovation in the primary market. They will simply exit the aftermarkets altogether. This is aptly illustrated by the Federal Circuit *CSU v. Xerox* case,[206] and the Ninth Circuit *Image Technical Services v. Kodak* case,[207] which contains practically identical facts as *CSU*. In those two cases, the request to deal was not made by Kodak's and Xerox's rivals in the primary photocopying machine market, but by independent service organizations ("ISOs") that provide repair services for photocopying machines. It is difficult to see how a rejection of a duty to deal would have spurred these ISOs to develop their own photocopying machine technology to compete with Kodak and Xerox. Compulsory licensing need not undermine the innovation incentives of the putative licensees.

¶121 The patentee will no doubt focus on the first presumption and argue that profits generated from the aftermarkets are necessary for it to recoup R&D investment. Unlike the case for the second presumption about rivals' innovation incentives, it does not seem possible to come up with a categorical rule that will help the courts to distinguish valid and invalid claims about the first presumption. The courts must instead scrutinize such claims closely and not accept them at face value. The patentee should be required to produce concrete evidence to substantiate the claim, including evidence that shows that profits from the primary market is insufficient to cover the R&D costs.

¶122 For instance, this claim about aftermarket profits would have been unlikely to hold in *Image Technical Services*, where Kodak had initially welcomed ISOs to provide maintenance and repair services. It was only after the maintenance and repair market had become lucrative and Kodak had lost an important maintenance contract with the state of California that Kodak altered its prior policy of supplying parts to ISOs.[208] If Kodak had considered profits from the aftermarket essential to its recoupment of R&D investment from the start, it would have been unlikely to have adopted an open policy with the ISOs initially. The Ninth Circuit eventually dismissed Kodak's business justifications for the refusal to deal on the ground that they were pre-textual, focusing on the subjective state of mind of Kodak's employees.[209] This approach has been criticized as being inconsistent with the modern antitrust focus on objective effects rather than subjective intent.[210] By focusing on Kodak's employees' subjective state of mind, the Ninth Circuit was in fact attempting to verify Kodak's innovation incentives argument. This focus is

[206] 203 F.3d 1322 (Fed. Cir. 2000).
[207] Image Technical Serv., Inc. v. Eastman Kodak Co., 125 F.3d 1195 (9th Cir.1997).
[208] 125 F.3d at 1214.
[209] *Id.* at 1218–20.
[210] U.S. DEP'T OF JUSTICE & FEDERAL TRADE COMMISSION, ___SERIAL NUMBER ____, ANTITRUST ENFORCEMENT AND INTELLECTUAL PROPERTY RIGHTS: PROMOTING INNOVATION AND COMPETITION 5 (2007).

clearly correct. The Ninth Circuit's mistake was its reliance on the wrong type of evidence. To determine the veracity of the innovation incentive argument, the Ninth Circuit should not have relied on what Kodak's employees thought, but whether profits from the aftermarket were objectively necessary for Kodak to recoup its investment.

B. Patent Holdup

¶123 Patent holdup presents another situation in which imposing a duty to license may be appropriate. Patent holdups are detrimental because they are likely to retard cumulative innovation. They are also likely to provide a windfall to the patentee that is above and beyond the profit necessary for the recoupment of innovation costs. The previous two Sections explain the importance of cumulative innovation as a major source of social benefits. Among the three models of cumulative innovation, patent holdup may be a particularly serious problem in the trunk-branch and the anticommons models.[211] As discussed earlier, both situations present great potential for social welfare loss due to patent holdups. Moreover, in the trunk-branch model, the Ramsey intuition means that antitrust can restrict patent exploitation at the margin without causing much effect on patentee reward. Cumulative innovation hence can be unlocked without substantial loss of patentee rewards. Patent holdup is likely to provide a windfall to patentees.[212] In the event of a holdup, the negotiation usually takes place *ex post* after the innovation has already been made and the R&D costs are sunk. The follow-on innovator is in a very weak bargaining position to secure a surplus from the negotiation that will allow it to recover those costs. It may need to accept any licensing arrangement that allows it to cover its variable costs. In fact, Lemley and Shapiro argue that *ex ante* negotiation would not significantly improve the follow-on innovator's bargaining position.[213] Cumulative innovation will be deterred over the long haul.

¶124 Imposing a duty to license on opportunistic patentees may solve this problem. If these patentees know that the courts may step in and mandate licensing at a reasonable royalty rate,[214] they will be induced to enter into negotiations with follow-on innovators in good faith.[215] The threat of compulsory licensing may become a default background legal rule against which negotiations between initial and follow-on innovators take place. The instances in which the courts need to intervene could be few.

---

[211] The use of compulsory licensing to resolve holdup problems in anticommons situations has been endorsed by Burk and Lemley. Burk & Lemley, *supra* note 29, at 1667. And the use of compulsory licensing in the trunk-branch scenario is vividly illustrated in the Hybritech monoclonal antibodies case. SCOTCHMER, *supra* note 12, at 154. The pioneering research on monoclonal antibodies was performed by scientists George Koehler and Cesar Millstein on public funding. Hybritech, the patentee seeking to enjoin others from utilizing the technology, only developed diagnostic kits. The innovation costs were likely to be insubstantial compared to the initial research, while the harm to social welfare would be great due to the importance of the technology.

[212] Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991, 2008 (2007).

[213] *Id*. at 2003–04.

[214] The royalty of course needs to allow the initial innovator to recover its innovation costs.

[215] The need for antitrust intervention in patent holdup cases has arguably receded after the Supreme Court's decision in *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388 (2006), in which the Court unanimously overturned the longstanding presumption in patent law that a patentee is entitled to a permanent injunction following a finding of infringement. This case has arguably reduced the possibility of holdup behavior.

C. Caveats

¶125 A number of caveats about the use of compulsory licensing are in order. First, compulsory licensing should be rarely granted when the patentee itself practices the patent. In such a case, there is a greater likelihood of reciprocity of interests, which should encourage the parties to negotiate in good faith. There is thus a lower likelihood of patent holdups. When one party to the negotiation has no need for the counterparty's technology, as in the case of patent trolls, opportunistic behavior and holdup are more likely. Also, the fact that the patentee practices and commercializes the patent itself means that it is more likely to compete with the putative compulsory licensee. As will be elaborated below, compulsory licensing should be rarely granted when the patentee and the licensee are competitors in the technology market. Antitrust should encourage competition at the technology development level so that society can benefit from competing technological offers.

¶126 Second, the courts should be cautious about imposing compulsory licensing when the putative licensee competes directly with the patentee in the primary technology market. In relation to this point, it is worthwhile to evaluate Professor Daniel Crane's objection to imposing a duty to deal with respect to intellectual property. Crane distinguishes the denial of injunctive relief under patent law from antitrust imposition of a duty to deal, rejecting the latter while endorsing the former. The crux of his objection is that the latter requires the monopolist patentee positively to assist rivals in the acquisition and deployment of the technology.[216] This objection raises some fundamental questions about the relationship between compulsory licensing and innovation incentives.

¶127 This author agrees with Crane that a patentee should not be required to provide positive assistance to its rivals if doing so will impair innovation incentives, both on the part of the patentee and the putative licensees (the incentive effect on future potential inventors presumably is a derivative of the effect on the patentee). The purpose of antitrust law is to encourage competition in technology and innovation. Having multiple firms pursue alternative routes to technological development is likely to be beneficial to society in the long run. However, requiring a patentee to provide positive assistance would only reduce innovation incentives if the putative licensee competes with the patentee in the technology market; *i.e.*, they offer competing technologies or products in the primary market. It will have negligible impact on overall innovation incentives in the case of aftermarket competitors.[217] ISOs and other aftermarket operators are ill-equipped to contribute technological alternatives in the primary market. In fact, in most cases, the aftermarket operators do not license the technology in order to produce a competing product. They obtain a license so that they can sell or use products manufactured by the patentee. There is no concern about positive assistance to rivals that will undermine their innovation incentives.

¶128 The same argument applies to follow-on innovators in the trunk-branch case as well as the anticommons case. To the extent that these follow-on innovators do not compete with the initial innovator in the development of basic scientific research or

[216] Crane, *supra* note 1, at 97.

[217] In the case where the putative licensee competes with the patentee in both the primary technology market and the aftermarket, the overriding need to preserve innovation incentives in the primary market means that compulsory licensing should be disfavored.

research tools, compulsory licensing does not undermine innovation incentives on the part of these follow-on innovators. The incentives of the initial innovator of course still need to be considered, but those will be preserved so long as private benefits of innovation cover innovation costs.

¶129 Third, a unilateral refusal to license claim should only be upheld if imposing a duty to license would substantially improve competition or promote innovation. Compulsory licensing may bring two main benefits: (1) greater competition in a market where competition has been hitherto precluded by patent protection, and (2) facilitation of innovation that builds upon the patented technology. If neither benefit will materialize as a result of compulsory licensing, there would be few good reasons to alter the patentee's reward, regardless of its impact on innovation incentives. If the market at issue is already highly competitive, consumers will have little to gain from the competition introduced by compulsory licensing. While some rivals may prefer to gain access to the patented technology to increase their competitive advantage, compulsory licensing in such a case would only result in wealth transfer between competitors with little attendant gains in consumer welfare. This is clearly an insufficient justification for antitrust intervention. Similarly, if innovation will not be promoted by compulsory licensing, such as by overcoming patent holdup, compulsory licensing should not be imposed.

D. Case Law

¶130 The proposed framework to unilateral refusal to license cases is largely consistent with the outcome, if not the analytical approach, of the existing case law.[218] Two leading cases, the Ninth Circuit *Image Technical Services v. Kodak* case and the *In re Independent Service Organizations Antitrust Litigation* case, both involved ISOs of photocopying machines that did not produce photocopying machines or engage in research in photocopying technology. The ISOs did not compete with the photocopying machine manufacturers in the primary market. The two courts reached divergent outcomes with essentially the same facts. The Ninth Circuit upheld the ISOs' unilateral refusal to license claim, while the Federal Circuit did not. It is suggested that despite its misplaced focus on the subjective intent of Kodak's employees, the Ninth Circuit's result is more consistent with an approach to the patent-antitrust interface that strikes a balance between preserving competition and safeguarding innovation incentives. With its expansive holding that absent a fraudulently obtained patent, sham enforcement action or illegal tying, a patentee is free to refuse to license its intellectual property, the Federal Circuit displayed too much deference to the patentee. Apart from these enumerated circumstances, there are situations in which compulsory licensing may be imposed to preserve or promote competition and innovation without impairing innovation incentives. There was no evidence in that case that Xerox relied on the profits from the maintenance market to recoup its R&D costs. Therefore, giving the ISOs access to Xerox's parts was unlikely to have reduced the manufacturer's innovation incentives. Meanwhile, there was evidence in both cases that requiring Kodak and Xerox to deal with rivals would

[218] *See, e.g.*, Image Technical Serv. v. Eastman Kodak Co., 125 F.3d 1195 (9th Cir. 1997); *In re* Indep. Serv. Orgs. Antitrust Litig., 203 F.3d 1322 (Fed Cir. 2000), Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147 (1st Cir. 1994); SCM Corp. v. Xerox Corp., 645 F.2d 1195 (2d Cir. 1981); Intergraph Corp. v. Intel Corp., 195 F.3d 1346 (Fed Cir. 1999).

have substantially improved competition in the maintenance market: Kodak held at least 80% of the maintenance market. And there was evidence that the unavailability of spare parts had made it increasingly difficult for the ISOs to compete with Xerox. There was also evidence in Kodak that many consumers actually preferred services provided by the ISOs. Gains in consumer welfare in both cases from compulsory licensing would have been substantial. With little impact on the patentee's innovation incentives and substantial gains in consumer welfare, compulsory licensing should have been imposed in both cases.

¶131 In *SCM Corp. v. Xerox Corp.*, the Second Circuit refused to require Xerox to license its patent on plain paper copier to a competitor to allow the latter to compete with Xerox in the copier market.[219] The court's language was very much in favor of patentees and has been interpreted as creating a *per se* legality rule for unilateral refusal to license a patent.[220] While a *per se* legality approach (if the Second Circuit did in fact intend the case to be interpreted as such) would clearly be too permissive, this author believes that the court reached the correct outcome in that case. The reason is simply that the plaintiff was a direct competitor of Xerox's in the photocopying machine market. Imposing compulsory licensing would have directly affected Xerox's patentee reward. It would have also reduced SCM's incentive to engage in its own research, which it presumably would have been prepared to do as it had planned to compete in the photocopying machine market.

¶132 Another leading unilateral refusal to license case is *Data General Corp. v. Grumman Systems Support Corp.*,[221] where the First Circuit rejected Grumman's claim that Data General illegally maintained its monopoly in the market for servicing Data General's computers by refusing to license its diagnostic software. Similar to the ISOs in *Image Technical Services* and *In re Independent Service Organizations Antitrust Litigation*, Grumman provided repair services for Data General's computers. This may suggest that the First Circuit should have upheld the unilateral refusal to license claim and ordered Data General to license its diagnostic software. The crucial difference between this case and the two photocopying machine cases is that in those cases, the spare parts were simply part of the photocopying machines that did not require independent effort to invent. The diagnostic software in *Data General* was itself a separate invention that required investments on the part of the computer maker. Therefore, at issue are two inventions, the computer and the diagnostic software. Allowing Grumman to gain access to the software may not have had much impact on the R&D for the computer, but it would most probably have undermined Data General's research effort in the software. Therefore, the First Circuit's decision was consistent with the framework proposed in this Section.

¶133 Lastly, in *Intergraph v. Intel*, the Federal Circuit rejected a unilateral refusal to deal claim by Intergraph, the owner of a technology known as "Clipper," which Intel incorporated into its microprocessor chips. Intel and Intergraph did not compete with each other in any relevant market. Intergraph competed in the market for graphics workstations, which according to the court was already highly competitive. Intel had

[219] *SCM Corp.*, 645 F.2d at 1195.

[220] Michael A. Carrier, *Refusals to License Intellectual Property After* Trinko, 55 DEPAUL L. REV. 1191 (2006).

[221] 36 F.3d at 1147.

been providing Intergraph technical assistance and privileged access to proprietary information, pre-release products, and allocation of new products. Intergraph's claim arose when Intel terminated their technical assistance and access to information services following repeated failure to settle a patent infringement suit between the two parties. While the technical assistance and access to information no doubt gave Intergraph an important commercial advantage in the graphics workstation market, there was no evidence that Intergraph needed the information and assistance for its own R&D. The resulting benefit in innovation from compulsory licensing would have been minimal. Nor would consumers stand to gain much from it; the market in which Intergraph competed was already highly competitive. The Federal Circuit thus reached the correct decision by rejecting Intergraph's claim.

¶134 The proposed approach to unilateral refusal to license is also consistent with the admonitions by Justice Scalia, writing for the majority in *Trinko,*[222] about the harm of judicial imposition of a duty to deal. One of the gravest concerns expressed by Justice Scalia about imposing a duty to deal was that it would discourage firms from making future investments in physical facilities.[223] The proposed approach specifically takes that concern into account by requiring that the innovator be fully compensated for its innovation costs. The incentive to invest would not be impaired, as feared by Justice Scalia.

¶135 Furthermore, Justice Scalia believes that "[e]nforced sharing also requires antitrust courts to act as central planners, identifying the proper price, quantity, and other terms of dealing[.]"[224] That concern may have been valid in the context of *Trinko*, which involved a heavily regulated industry where the identification of the parameters of sharing is more akin to the kind of rate-setting activities that are usually undertaken by the regulatory agencies. It has little relevance for the unilateral refusal to license cases. The federal courts regularly engage in royalty setting in patent damages determination. The kind of "rate-setting" activity required in compulsory licensing is not unfamiliar to the courts and should not raise any special concerns. Moreover, at least in the case of patent holdup, once the threat of compulsory licensing becomes a real one, judicial intervention will only be necessary if negotiation breaks down, which should be rare once judicial imposition of a duty to deal becomes a distinct possibility.

¶136 Justice Scalia further argues that enforced sharing encourages collusion.[225] While that is no doubt a valid concern, policing against collusion is the bread-and-butter of antitrust agencies. There is no reason to believe that the agencies will be uniquely incapable of detecting collusion among technology firms. Joint ventures may facilitate collusion, and that has not prevented the agencies from approving them in most instances.

---

[222] Verizon Comm. v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398, 124 S.Ct. 872 (2004). Whether the courts will take the same approach to both unilateral refusal to deal regarding a physical property and unilateral refusal to license an intellectual property is still an unsettled question. Shelanski, *supra* note 199, at 373–78. This author believes that physical property should be accorded no less deference than intellectual property as far as unilateral refusal to deal is concerned. Given that the Supreme Court has not ruled on the appropriate standard for unilateral refusal to license intellectual property, and the lower courts have adopted a divergent range of approaches, the ensuing discussion will proceed under the assumption that the framework set out in *Trinko* applies to unilateral refusal to license patent rights.

[223] *Trinko*, 540 U.S. at 407–08.

[224] *Id*. at 408.

[225] *Id*.

The likelihood that collusion may materialize in compulsory licensing situations is arguably lower than that within joint ventures, given the involuntary nature of the former.

V. CONCLUSION

¶137 This Article calls for the restoration of a full and more probing consideration of innovation incentives in the patent-antitrust interface. A full consideration should not solely focus on the incentives of the original innovator, but those of the follow-on innovators as well. This is important because cumulative innovation creates substantial social benefits. The technological development process does not end after the initial innovation. Subsequent innovations will build on it and create more value for consumers and society. A probing consideration should not accept any innovation-based defenses of patent exploitation practices at face value, as some antitrust courts seem to have done. Instead, it should ask searching questions about the extent to which a patentee depends on patentee reward for innovation incentives. There should be a return to full recognition and application of the fundamental premise of the patent system that innovators will continue to produce innovations as long as their innovation costs are fully compensated. While most courts are fully aware of this premise, they seem to overlook it as soon as they turn their attention to a concrete case and become singularly focused on preservation of patentee reward. Once this premise is again fully recognized, the courts will realize that there is considerable room to advance consumer welfare without impairing innovation incentives. Antitrust need not show the same degree of deference to patent policy as has often been advocated.

¶138 This, however, does not mean that antitrust should interfere with the patent system on a substantial scale. To the extent that the specific policy issues presented by the patent-antitrust interface are already resolved internally within the patent system, antitrust should wisely abstain from intervention. There are signs that the patent system is moving in the right direction and beginning to curtail expansive patent rights. But where antitrust can play a useful role, the courts should not shy away from it simply because dynamic efficiency arguments are by their nature more difficult to evaluate and rebut. It is the job of the antitrust courts to subject every pro-competitive justification presented to them to close scrutiny. No argument should be allowed to become a trump card; not even one based on something as important to economic growth as innovation. The antitrust courts have underestimated their ability to assess dynamic efficiency arguments. Once the courts recognize the diversity of innovations and acquire a fuller understanding of the various attributes of innovation, the courts will be in a better position to assess these arguments. This Article has only illustrated the application of this new awareness in the context of the unilateral refusal to license a patent. The implication of the proposed approach in other areas of the patent-antitrust interface remains to be explored. The hope is that, in time, a more balanced approach to the patent-antitrust interface with full and fair consideration of dynamic efficiency arguments will emerge.

Northwestern Journal of Technology and Intellectual Property

Volume 11, Number 5 (April 2013)