

Intellectual Property as Natural Monopoly: Toward a General Theory of Partial Property Rights

John F. Duffy*

The late nineteenth and early twentieth centuries saw the rise of numerous capital-intensive industries in which principles of competitive production did not seem to function well. In these industries — particularly the railroads, electrical service and telephony — the marketplace tended inexorably, and seemingly naturally, toward monopoly rather than sustained competition. The phenomenon experienced in these industries came to be described “natural monopoly” and a variety of legal structures and institutions were created to manage and control the industries in the absence of competition. These structures formed a large part of the regulatory state that grew up during the period.

The closing of the twentieth century saw the end of many of the traditional forms of natural monopoly regulation that had arisen a century earlier. The change was brought about not only by different economic theories, but also changed conditions. Because of technological advances, some industries that had once been thought to be natural monopolies could now sustain competition.² In other areas, policy makers lost faith in governmental regulatory mechanisms that had previously been directed to solving the natural monopoly problem.³ In general, it was thought, “[t]he era of big Government is over.”⁴

But there has been another change too. At the beginning of the twenty-first century, economists and policy makers increasingly turn their attention from the “heavy” industries that dominated the economy a century ago to new industries that were characterized less by brutish strength and more by the product of minds. Alan Greenspan, among others, would look across the economic landscape and see “[t]echnological advance . . . promoting the trend toward increasing conceptualization of U.S. GDP.”⁵ With this change, the focus of policy makers shifted from the old questions about price regulation of utilities to seeming new questions about

* Professor of Law, George Washington University Law School; Visiting Professor of Law, New York University, Spring 2005.

² Electrical generation is no longer a natural monopoly, although retail distribution of electricity still is. See Bernard S. Black & Richard J. Pierce, Jr., *The Choice Between Markets and Central Planning in Regulating the U.S. Electricity Industry*, 93 Colum. L. Rev. 1341, 1430 (1993) (noting that “[e]lectric power generation is no longer a natural monopoly”); *id.* at 1362 (noting that “[r]etail distribution generally involves increasing returns to scale — marginal cost less than average cost”).

³ See, e.g., Joseph D. Kearney and Thomas W. Merrill, *The Great Transformation of Regulated Industries Law*, 1998, 98 Colum. L. Rev. 1323 (1998).

⁴ William J. Clinton, Address Before a Joint Session of the Congress on the State of the Union, 32 Weekly Comp. Pres. Doc. 90 (January 23, 1996).

⁵ Alan Greenspan, *Intellectual Property Rights*, Remarks at the Stanford Institute for Economic Policy Research Economic Summit, Stanford, California, February 27, 2004 (available at <http://www.federalreserve.gov/BoardDocs/speeches/2004/20040227/default.htm>).

fair use, compulsory licensing and the proper scope of patents and copyrights. The regulatory policies and learning of the twentieth century are invoked infrequently in analyzing the modern, high tech world of “conceptual” GDP. But this is not as it should be.

As its title suggests, the thesis of this article is that the body of regulatory law and policy built up over the last century is not so utterly divorced from issues concerning the scope and structure of intellectual rights, which have such importance to the conceptual economy of the twenty-first century. The analogy between more traditional regulated natural monopolies and intellectual property — particularly the regimes of patent and copyright — is closer than any of the previous analogies used in describing and understanding the legal and regulatory structure of intellectual property. The analogy is a good one not merely because of some abstract definitions that could be postulated to show the equivalence of the two fields, but because the modes of legal regulation in the two areas show a remarkable degree of consistency. Indeed, comparison of the two leads to insights not only about intellectual property, but also about the regulation of traditional natural monopolies.

The segregation of intellectual property regulation from more traditional regulation of natural monopolies is only one manifestation of a more general tradition of what might be termed intellectual property exceptionalism — the tradition of treating intellectual property as somehow different from physical property. This tradition is critically examined in Part I, for an appreciation of the tradition’s weaknesses is a useful first step in understanding the close relationship between intellectual property and the regulatory policy applied in the physical domain. As shown in Part I, the basis for intellectual property exceptionalism has shifted with time. In the past, commentators and courts treated intellectual property exceptionally because they believed that intellectual property rights generally, and patents in particular, were monopolies whereas rights in physical property were not. This theory was important in its time, and it fostered unnecessary and unfortunate hostility toward the enforcement of intellectual property rights. The theory has, however, been sapped of its vitality as its weaknesses and contradictions have been exposed.

The decline of that theory has not led to the abandonment of intellectual property exceptionalism. Rather, as discussed in Part I.B, another theory has arisen to justify the separation of intellectual from physical property. This new theory posits that intellectual property is special because it protects information, and information has unique attributes not shared (or not *generally* shared) with physical property. Information is said to be difficult to identify and thus to protect with rights; it is said to be difficult to sell information because of a paradox first identified by Kenneth Arrow; and it is said to be a “public good.” Yet each of these claims present severe difficulties; they are typically supported only by comparisons between a subset of information and a subset of physical property. For example, much information is a private good, and much physical property is public good. If the more general classes of all information and all physical property are considered, no categorical differences between the two can be discerned.

The proper classification of intellectual property is not an idle exercise in taxonomy. Rather, the correct classification can provide insights, analogies and history that help us to understand more thoroughly both the broader field of natural monopoly and the more specific areas of intellectual property. For example, it is often asserted nearly as a truism that intellectual property differs in some fundamental way from physical property. Landes and Posner assert that “the law ... impos[es] limitations on intellectual property that go beyond what is found in the domain of physical property.”⁶ But this instinct is not right. In both the realms of both physical and nonphysical property, the law grants highly limited, partial property rights in areas where natural monopoly problems arise.

Similarly, a comparison of intellectual property law and natural monopoly regulation provides a good rejoinder to the extravagant claim that, because intellectual property is just another species of property, deviations for full property rights (roughly, fee simple) should be suspect. Again, while not categorically different from physical property, much of intellectual property closely resembles a particular type of physical property, and the rights associated with that type of property are often qualified far more than a common law version of fee simple in Blackacre.

Parts II provides more detail for the symmetry between natural monopolies and intellectual property. This symmetry exists not only as a matter of economics, but also in the regulatory structures and polices pursued in each field. There are of course differences; the fields would not be separate if there were not. While not undermining the fundamental closeness of the two fields, these breaks in symmetry are interesting in their own right, and the contrasts can provide a better understanding of the underlying motivations for particular regulatory structures.

Finally, Part III attempts to use the symmetry between the two fields to investigate a question fundamental to both fields: Why in each field does the government curtail property rights? Different answers to that question have traditionally been emphasized in the two fields. This paper attempts to harmonize the answers.

I. The Characterization of Intellectual Property: Property, Monopoly or Public Good?

What the legal profession now generally refers to as “intellectual property” has borne, and indeed continues to bear, many different labels. Historically, rights to intellectual property have been called somewhat disparagingly “monopolies,” or where a more approving tone was desired, they have been referred to simply as “property.” These two labels, and their attendant connotations, have long been competitors to dominate the vernacular of policy makers. More recently, intellectual property has also been referred to as a “public good.” All three of these characterizations are to some degree accurate and to some degree not. But the question whether a particular label accurately applies to intellectual property is not merely a semantic dispute.

⁶ William Landes and Richard Posner, *The Economic Structure of Intellectual Property Law* 21 (Harvard 2003).

Labels are applied not only to classify the subject but also to provide analogies. The propriety of the classification should be judged not merely by semantics, but by functional considerations. It should be judged not merely on whether intellectual property fits within some possible definition of the term, but on whether the modes of regulation associated with the category correspond to the modes of regulation applied to intellectual property.

The fundamental difficulty with all of the traditional labels — monopoly, property, and public good — is not that intellectual property cannot be made to fit within some definition of those terms. Each of the terms can be construed to a broad area of disparate subjects. But if construed with such breadth, then the categorization does little good because the broad class of subjects also corresponds to many disparate forms of regulation. Alternatively, if the categories are narrowed to what might be termed their paradigmatic examples (e.g., Blackacre for property), the means of legal regulation traditionally applied to the paradigmatic case correspond little to the system of regulation applied to intellectual property.

A. Property vs. Monopoly.

Perhaps the oldest and most fierce conceptual struggle in intellectual property is the battle between those who argue that intellectual property rights, particularly patents and copyrights, are best considered as monopolies, and those who would classify intellectual property merely as another species of “property.” The struggle dates back to before the founding of the United States, and it continues into the pages of modern court opinions and law review articles.

In the late 18th century, intellectual property rights were commonly described as monopolies, but the description was not uncontroversial. Adam Smith rejected the notion that copyrights and patents could be thought of as a natural species of property; he classified them as “monopolies,” though he thought them desirable monopolies.⁷ Thomas Jefferson was also a proponent of the “monopoly” view. At the time of the framing of the U.S. Constitution, Jefferson viewed both copyrights and patents as dangerous government “monopolies” that should be strictly limited, if they were to be granted at all.⁸ Years later Jefferson would also argue

⁷ Adam Smith, *Lectures on Jurisprudence* 83 (R.L. Meek, D.D. Raphael & P.G. Stein eds., Oxford Univ. Press 1978) (1896).

⁸ Jefferson’s view is evident in his 1789 recommendation to James Madison that the then-circulating draft of the Bill of Rights should include the following provision restricting the government’s ability to grant the “monopolies” of copyright and patent:

Article 9. Monopolies may be allowed to persons for their own productions in literature, and their own inventions in the arts, for a term not exceeding __ years, but for no longer term, and no other purpose.

Letter to James Madison, August 28, 1789, in 7 *The Writings of Thomas Jefferson* 444, 451 (Andrew A. Lipscomb, ed. 1904). A year earlier, Jefferson took an even more strident position. Once again, he argued in favor of a “bill of rights” which, “[i]t seems pretty generally understood, ... should go to juries, habeas corpus, standing armies,

against the notion that “an idea, the fugitive fermentation of the an individual brain, could, of natural right, be claimed in exclusive and stable property.”⁹ He objected to the “preten[sion] ... that inventors have a natural and exclusive right to their inventions” and concluded that “[i]nventions then cannot, in nature, be a subject of property,” although “Society may give an exclusive right to the profits arising from them, as an encouragement to men to pursue ideas which may produce utility.”¹⁰

Even in this early era, however, patents and copyrights were not uniformly classified as monopolies. The United States Constitution avoided describing the subjects as monopolies, instead referring merely to “the exclusive Right” that Congress could grant in “Writings and Discoveries.”¹¹ Early U.S. statutory law also avoided using the term “monopoly” in authorizing patents and copyrights, and the Patent Act of 1793 expressly described patent rights as “property.”¹² The “property” view also had theoretical support, especially in copyright. Francis Hargrave’s famous tract *An Argument in Defense of Literary Property* answered the charge that copyrights were a form of monopoly by defining the word narrowly to mean “an *appropriation* of the right of carrying on some particular branch of trade or commerce; to which all men have *originally* a common and equal pretension.”¹³ Copyrights were different, Hargrave argued, because “before publication *only* the author has the right of multiplying and copying his works”; thus, the grant of the copyright merely extended pre-publication property rights to the time after publication.¹⁴

printing, religion and monopolies.” Letter to James Madison, July 31, 1788, in 7 *id.* at 93, 96. But in this earlier letter, Jefferson thought that “the benefit of even limited monopolies is too doubtful, to be opposed to their general suppression.” *Id.* at 98.

⁹ Letter to Isaac McPherson, August 13, 1813, in 13 *id.* at 326, 333.

¹⁰ *Id.* at 334.

¹¹ U.S. Const. art. I, §8, cl. 8.

¹² Act of February 21, 1793, ch. 11, section 1, 1 Stat. 318. prescribes

¹³ Francis Hargrave, *An Argument in Defense of Literary Property 1774* 28-29 (Garland Publishing 1974 ed.). Hargrave’s definition of monopoly is echoed in *United States v. Dubilier Condenser Corp.*, 289 U.S. 178, 186 (1933), which stated that “[t]hough often so characterized, a patent is not, accurately speaking, a monopoly, for it is not created by the executive authority at the expense and to the prejudice of all the community except the grantee of the patent.” The Court defined monopoly to require “tak[ing] something from the people” by “the giving of an exclusive privilege for buying, selling, working or using a thing which the public freely enjoyed *prior to the grant*.” *Id.* (emphasis added).

¹⁴ *Id.* Hargrave’s view that authors held a natural or common law property right prior to publication was controversial as a matter of historical. Nonetheless, the Supreme Court endorsed the view in *Wheaton v. Peters*, 33 U.S. 591, 657 (1834) (“That an author, at common law, has a property in his manuscript, and may obtain redress against any one who deprives him of it, or by improperly obtaining a copy endeavours to realise a profit by its publication, cannot be doubted”).

The property-versus-monopoly debate continues through to modern times. Thus, for example, Harvard Professor Lloyd Weinreb confidently asserts that “the most that can be said confidently about copyright or patent is that it confers a monopoly,”¹⁵ and economists Michele Boldrin and David Levine proclaim in the prestigious pages of the *American Economic Review* that modern rights in copyright and patent “create[] a socially inefficient monopoly, and what is commonly called intellectual property might be better called “intellectual monopoly.”¹⁶ Yet there are authorities of equal weight on the other side. The Antitrust Guidelines of the Federal Trade Commission and Department of Justice refuse to presume that intellectual property rights convey any market power.¹⁷ Since substantial market power is a precondition to the legal definition of monopoly for antitrust purposes,¹⁸ the Antitrust Guideline effectively deny that intellectual property rights should be viewed as monopolies. In judicial opinions, Federal Circuit Judge Markey has written one of the most strident statements against the monopoly view of patents. Complaining that “[i]t is but an obfuscation to refer to a patent as ‘the patent monopoly’ or to describe a patent as an ‘exception to the general rule against monopolies,” Markey stresses that “[a] patent, under the statute, is property” and that “[n]owhere in any statute is a patent described as a monopoly.”¹⁹ The academic literature also continues to bristle with articles supporting the “property” characterization and disparaging the monopoly view of intellectual property.²⁰

¹⁵ Lloyd L. Weinreb, Copyright for Functional Expression 111 Harv. L. Rev. 1149, 1205 (1998).

¹⁶ Michele Boldrin and David Levine, *The Case Against Intellectual Property*, 92 Am. Econ. Rev. (Papers and Proceedings) 209 (2002). Economic literature has a long, if not consistent, tradition of referring to intellectual property as a monopoly. See Frank H. Knight, Diminishing Returns from Investment, 52 J. Pol. Econ. 26, 28 (1944) (“knowledge can yield a return to its individual possessor only if he has some degree of monopoly power enabling him to restrict its diffusion and general application. Such power may ... rest on a legal grant, such as a trademark, copyright, or patent ...”); Robert M. Hurt & Robert N. Schuchman, The Economic Rationale of Copyright, 56 Am. Econ. Rev. Papers & Proc. 421, 425-26, 432 (1966) (describing copyright as a form “monopoly”).

¹⁷ U.S. Department of Justice and the Federal Trade Commission, Antitrust Guidelines for the Licensing of Intellectual Property § 2.2 (“The Agencies will not presume that a patent, copyright, or trade secret necessarily confers market power upon its owner.”).

¹⁸ See Phillip Areeda, Monopolization, Mergers, and Markets: A Century Past and the Future, 75 Calif. L. Rev. 959, 960 (1987) (“most judges have recognized that ‘substantial’ market power is involved in the statutory concept of monopolization”).

¹⁹ Schenk, A.G. v. Nortron Corp., 713 F.2d 782, 786 n.3 (Fed. Cir. 1983). See also Howard T. Markey, Why Not the Statute?, 65 J. Pat. Off. Soc’y 331, 331-33 (1983).

²⁰ For a sample of the seemingly endless literature debating whether intellectual property is properly viewed as property or monopoly, see Roger E. Meiners & Robert J. Staaf, Patents, Copyrights, and Trademarks: Property or Monopoly?, 13 Harv. J.L. & Pub. Pol’y 911, 915-25 (1990); Frank H. Easterbrook, Intellectual Property is Still Property, 13 Harv. J. L. & Pub. Pol’y 108 (1990); Kenneth W. Dam, The Economic Underpinnings of Patent Law, 23 J. Legal Stud. 247, 253 (1994); Edmund W. Kitch, Elementary and Persistent Errors in the Economic Analysis of Intellectual Property, 53 Vand. L. Rev. 1727, 1729-38 (2000).

For all of its pitched rhetoric and passion, the debate over whether intellectual property is properly considered monopoly or property has done nothing to advance understanding of intellectual property. The problems begin with semantics, but extend much deeper. On a semantic level, one problem with the debate is that the concept of “monopoly” has more than one definition. One old and venerated definition is that a monopoly denotes merely “exclusive possession or control of something” usually granted by the state.²¹ Under this broad definition, patents and copyrights certainly qualify as monopolies. A much more narrow definition is generally used for antitrust law, where a “monopoly” is only considered to exist if a firm has substantial power to raise prices and control supply in a relevant market. Applying this definition, most patents and almost all copyrights are not monopolies, because the “market” often contains many substitutes for the subject matter covered by a particular intellectual property.²²

*United States v. E. I. Du Pont de Nemours & Co.*²³ is an excellent case to see these two definitions in operation. Early in its opinion, the Court categorically describes a patent as “the most familiar type of classic monopoly.”²⁴ Thus the opinion could be cited in support of the monopoly view of intellectual property. But, oddly enough, it could also be cited to support precisely the opposite view. The issue in the case was whether Du Pont had a monopoly for purposes of Section 2 of the Sherman Act. At the time of the litigation, Du Pont still held over 200 valid, enforceable patent rights on cellophane,²⁵ and so Du Pont had, to use the Court’s words, over 200 “classic monopol[ies].” Nevertheless, the Court held that, for the purpose of defining a monopoly under the antitrust laws, a firm must have power to affect price and supply in a “relevant market,” which requires an “appraisal of the ‘cross-elasticity’ of demand in the trade.”²⁶ The product cellophane was not a market unto itself, the Court held; rather, cellophane was just part of a larger market of all “flexible packaging materials.” Because Du Pont did not have power to control price and supply in that larger market, it could not be found to have engaged in any monopolization. In the end, Du Pont had no monopoly, even though it had over 200 classic monopolies!

²¹ Webster’s Encyclopedic Unabridged Dictionary of the English Language 927 (1989) (definition #4). See also WEBSTER’S NEW TWENTIETH CENTURY DICTIONARY (2nd ed. 1983) (defining “monopoly” to have, among other possible meanings, the following: “1. exclusive control of a commodity or service in a given market, or control that makes possible the fixing of prices and the virtual elimination of free competition. 2. an exclusive privilege of engaging in a particular business or providing a service, granted by a ruler or the state.”) [Other definitions, OED, etc.]

²² There are exceptions. For example, the Microsoft’s copyright on the Windows operation system confers monopoly power in the market for operating systems. See [cite].

²³ 351 U.S. 377 (1956).

²⁴ 351 U.S. at 392.

²⁵ *United States v. E. I. du Pont de Nemours & Co.*, 118 F. Supp. 41, 140 ¶ 470 (D. Del. 1953).

²⁶ 351 U.S. at 394-95.

Passing over the semantic dispute, we can agree that intellectual property is not necessarily a monopoly as defined under antitrust laws, but is always a monopoly if a more encompassing definition is used. This leads to another problem, however. If a broad definition is employed, then “[a]ll private property is, in a sense, a monopoly.”²⁷ Property — any property — is an exclusive right that is granted by the state, which is the broader definition of monopoly. Though this point may be frequently overlooked by those scholars and commentators who debate whether intellectual property is better classified as property or monopoly, the point is hardly new. It was familiar to Karl Llewellyn, who described property as “the monopoly of the things which are scarce.”²⁸ And the point is also well known to economists.²⁹ Indeed, one economic historian has suggested that the debate about whether patent rights constitute monopolies ultimately died down in the nineteenth century as “the judiciary came to recognize openly that the enforcement and protection of all property rights involved trade-offs between individual monopoly benefits and social welfare.”³⁰ This overlap between the two concepts explains why courts have sometimes referred to intellectual property as both “property” and “monopoly” in virtually the same breath.³¹

The property-monopoly debate is to some extent a pointless disagreement about semantics, and to some degree an incoherent dispute because both concepts of property and monopoly are broad enough to encompass intellectual property. But there is even one more flaw in the debate. Even if intellectual property is characterized as property or monopoly, the proponents of each particular characterization must then explain why the regulation of intellectual property is so different from the regulatory structure applied in the paradigmatic cases. For the “monopoly” proponents, the task is to explain why the intellectual monopolies are

²⁷ Weinreb, 111 Harv. L. Rev. at 1230, n.330.

²⁸ Karl Llewellyn, *The Bramble Bush* 10 (1960); see also *id.* (noting that “so far as the state safeguards such a monopoly and at the same time makes possible the free transfer of the things monopolized, it sets the framework within which the market, the agreement, the contract moves”). See also Kitch, 53 Vand. L. Rev. at 1729-30 (noting that the term “monopoly” can be used “in the sense of an exclusive right” but that then the term includes “all property rights” because those rights always “enable[] the owner to exclude others from the use of the subject matter of the right”).

²⁹ See, e.g., Senate Subcomm. On Patents, Trademarks, and Copyrights of the Senate Comm, on the Judiciary, 85th Cong., *An Economic Review of the Patent System* 53 (Comm. Print 1958) (prepared by Fritz Machlup) (“There is the idea that ‘property’ and ‘monopoly’ are one and the same thing from the economic point of view, and that the ‘owner’ of an invention has a monopoly of its use just as the owner of a house has a “monopoly” of the use of the house.”).

³⁰ B. Zorina Khan, *Property Rights and Patent Litigation in Early Nineteenth-Century America*, 55 J. Econ. Hist. 58, 92 (1995).

³¹ See, e.g., *Fox Film Corp. v. Doyal*, 286 U.S. 123, 127 (1932) (“The owner of the copyright, if he pleases, may refrain from vending or licensing and content himself with simply exercising the right to exclude others from using his *property*. The sole interest of the United States and the primary object in conferring the *monopoly* lie in the general benefits derived by the public from the labors of authors.”) (emphasis added; citation omitted).

fostered when the general government policy is one of hostility to monopolies. For the property proponents, it is to account for why intellectual property law is so dramatically different from the paradigmatic regulation of fee simple in Blackacre. After all, two of the major branches of intellectual property — patent and copyright — are *constitutionally* required to have limited terms. What accounts for such a fundamental difference from fee simple?

Of course, both sides of the debate can recite good policy reasons to explain why the regulation of intellectual property departs from the paradigmatic regulatory approaches to property and monopoly. But these policy reasons only highlight that both general labels are largely worthless. Neither provides good intuitions or analogies for analyzing the basic problems of intellectual property.

B. The Public Good Formulation.

In recent years, commentators have begun to describe intellectual property in another way. Intellectual property covers information and, so the argument goes, information is different from physical goods because “information is a public good.”³² Here is how one standard text explains the point:

Information has the characteristics of what economists call a “public good” — it may be consumed by many people without depletion, and it is difficult to identify those who will not pay and prevent them from using the information. Once the idea of the intermittent windshield wiper is disclosed, others can imitate its design relatively easily. Once a book is published, others can copy it at low cost. It is difficult to exclude nonpurchasers. Ideas and information can also be used by many without depleting the enjoyment of others. Unlike an ice cream cone, a good story or the concept of intermittent windshield wipers can be enjoyed by many without diminishing the enjoyment of these creations by others.³³

This argument has numerous problems. Let us begin with the claim that information is somehow different from physical property because information, generally speaking, has public good qualities. To analyze this claim, we need to know what is a public good. Perhaps the most common definition in the literature is that a public good is characterized by two essential

³² See, e.g., Yochai Benkler, *Constitutional Bounds of Database Protection: The Role of Judicial Review in the Creation and Definition of Private Rights in Information*, 15 Berkeley Tech. L.J. 535, 543 (2000); Gideon Parchomovsky and Peter Siegelman, *Towards an Integrated Theory of Intellectual Property*, 88 Va. L. Rev. 1455, 1467 (2002) (describing “information [as] the quintessential public good”); Robert A. Heverly, *The Information Semicommons*, 18 Berkeley Tech. L.J. 1127, 1157 n127 (2003) (“No one seriously disputes the public-good nature of information.”).

³³ Robert Merges, et al., *Intellectual Property in the New Technological Age* 13 (Aspen 1997).

properties: It is “nonrival” and “nonexcludable.”³⁴ These two properties are easily explained. A “nonrival” good is one that, once produced, can be consumed by all without any person’s consumption impairing any other’s consumption. In effect, additional “units” of the good can be produced (and consumed) with zero marginal cost. A “nonexcludable” good is one that, once produced, is available to all because it is not possible to exclude anyone from the good. The two conditions have been neatly summarized by Nobel laureate Joseph Stiglitz, who defines “pure public goods” as ones for which it is neither desirable nor feasible to ration their use.³⁵ The classic example of such a good is national defense. Once a country provides for its defense, all within the nation can enjoy the good without diminishing each other’s enjoyment, and individuals could not feasibly be excluded from protection of the defense forces.³⁶

Defining a public good as both nonrival and nonexcludable traces back at least to Paul Samuelson’s seminal article on *The Pure Theory of Public Expenditure*.³⁷ Samuelson defined

³⁴ See, e.g., See Robert Cooter & Thomas Ulen, *Law and Economics* 46 (4th ed. 2004) (defining a “public good” as “a commodity with two very closely related characteristics: 1. *nonrivalrous consumption* . . . , and 2. *nonexcludability*); Yochai Benkler, *Constitutional Bounds of Database Protection: The Role of Judicial Review in the Creation and Definition of Private Rights in Information*, 15 *Berkeley Tech. L.J.* 535, 543 (2000) (“Information is a public good. It is nonexcludable and nonrival.”). Julie Cohen, *Copyright and the Perfect Curve*, 53 *Vand. L. Rev.* 1799, 1807 n.21 (2000) (“A public good is a good that can be consumed without depletion (non-rivalrous consumption) and that can be withheld from nonpaying beneficiaries only at prohibitive cost (non-excludability).”); see also id. at 1807 (claiming that information products “have strong public good characteristics”). C. Edwin Baker, *Giving the Audience What It Wants*, 58 *Ohio St. L.J.* 311, 316 n.14 (1997) (“Economic definitions of “public good” usually emphasize two aspects: “nonrivalrous use,” . . . and “nonexcludability.””); Daphna Lewinsohn-Zamir, *Consumer Preferences, Citizen Preferences, and the Provision of Public Goods*, 108 *Yale L.J.* 377, 377 (1998). (defining a public good based on nonrivalrous consumption and nonexcludability).

³⁵ Joseph E Stiglitz, *The Economics of the Public Sector* 99 (1986).

³⁶ Nearly all commentators invoke national defense as the quintessential public good. See id. at 102 (“National defense is one of the few *pure* public goods”); Cooter & Ulen, *supra* note ___, at 46-47 (also invoking national defense to illustrate the concept of a public good); [more cites]. Of course, even national defense does not always fit the definition in its purest form. See Fred S. Hoffman, *The Economic Analysis of Defense: Choice without Markets*, 49 *Amer. Econ. Rev. (Papers & Proc.)* 368, 369 (1959) (noting that, although “[n]ational may well be the public good par excellence, [this] does not mean that even it is a pure public good”). For example, a nation may in some cases have to ration defense forces between two areas of the country. For example, a nation could choose to protect one city with more anti-aircraft guns than another city. In such cases, consumption of defense by some citizens (e.g., those residing in City A) may preclude consumption by others (e.g., those living in City B), and of course, the rationing of defense is then feasible at least on the level of choosing the localities to protect. See id. (noting that “[a]n antiaircraft missile battalion on Long Island contributes more to the security of a New Yorker than does one in Los Angeles”); see also John C. Head, *Public Goods and Public Welfare* 171 (Duke 1974) (noting that exclusion is never completely impossible). Moreover, in an era when national defense is more concerned with terrorists than blitzkriegs, many aspects of defense (i.e., security) may very well be rival and excludable, and thus can be efficiently provided by private security firms.

³⁷ Paul Samuelson, *The Pure Theory of Public Expenditure*, __ *Rev. of Econ. & Stat.* 387 (1954). Samuelson’s article is generally acknowledge to be the “central contribution to the modern discussion” on public good. John C. Head, *Public Goods and Public Welfare* 68 (Duke 1974).

“collective consumption good” as a good “which all enjoy in common in the sense that each individual’s consumption of such a good leads to no subtraction from any other individual’s consumption of that good.”³⁸ Samuelson emphasized that, under this definition, a collective consumption good “differs from a private consumption good in that each man’s consumption of it . . . is related to the total [consumption] by a condition of *equality* rather than of summation.”³⁹ Because Samuelson’s definition requires that *all* consumers consume the good in *equal* quantities (though not necessarily in equal degrees of satisfaction), his definition implies that the collective consumption good is nonexcludable.⁴⁰ As other commentators have noted, a direct implication of Samuelson’s definition is all enjoy in common in the sense that each individual’s consumption of such a good leads to no subtraction from any other individual’s consumption of that good

Asserting that intellectual property constitutes a public good under this common, two-feature definition poses an obvious problem: Intellectual property law is all about granting rights to exclude. If intellectual property — or, perhaps more accurately, the underlying innovations, writings and other informational products — really were nonexcludable, then the law in this field would be trying to do the impossible. Other commentators have noted, though usually in passing and without further thought, the problem of classifying information as a public good despite its excludability. Thus, as Stanford economist Paul Romer explains, “[e]ven though the information from discoveries is nonrival . . . , economically important discoveries usually do not meet the other criterion for a public good; they typically are partially excludable, or excludable for at least some period of time.”⁴¹

Nevertheless, many legal and economic commentators continue to characterize the subjects protected by intellectual property as a public good even though it fails the classic two-part definition. Three justifications have been made for doing so. First, some commentators have asserted that the subjects protected by intellectual property *would be nonexcludable* if the government did not grant and enforce exclusive rights. Thus, Yale’s Yochai Benkler writes:

Information is a public good. It is nonexcludable and nonrival. *Its nonexcludability makes it unsusceptible to appropriation, except by grant of exclusive right coupled with the threat of state enforcement; consequently, this*

³⁸ Rev. of Econ. & Stat. at 387.

³⁹ Id. at 388.

⁴⁰ See John C. Head, Public Goods and Public Welfare 171 (Duke 1974) (“it should be clear that impossibility of exclusion is a direct implication of [Samuelson’s] formulation of the equal consumption condition for public goods”); see also id. at 80 (noting that one distinct characteristic of the public good is that “once a unit of service is made available to one individual, a service unit of similar quality not only *can* but *must* be made available to all other individuals”).

⁴¹ Paul M. Romer, *The Origins of Endogenous Growth*, 8 J. of Econ. Perspec. 3, 13 (1994) (also noting that “[b]ecause people and firms have some control over the information produced by most discoveries, it cannot be treated as a public good”).

characteristic is the potential justification for recognizing and enforcing such rights.⁴²

In short, information is nonexcludable only in the sense that it is *potentially* nonexcludable: It would not be susceptible to appropriation if the government were unwilling to grant and enforce legal rights. But if “nonexcludable” has this definition, then all property is nonexcludable. Consider land.⁴³ If the government does not grant or enforce legal rights in real property, the land becomes a commons, and no one person could exclude another. Nor is it an answer to note that, even without legal rights, possessors of land could resort to fences, locks and weapons — i.e., self-help — to protect the land they possess. If no property rights exist in the land, interlopers have just as much of a “right” to break the fences and locks and otherwise use force to gain access to the land. Moreover, in the absence of property rights, a person possessing information might have a greater practical ability to exclude than a person possessing land. Valuable information can be exploited as a secret even without legal rights. In such circumstances, those not in possession of the information may be able to obtain the information easily because they do not know where to look, and the possessor can have a practical ability to exclude.

Indeed, far from presenting a problem of nonexcludability, information often presents what might be termed a problem of nonincludability — that even with active governmental action to force disclosure, the possessor remains capable of excluding others. Consider, to take an extreme example, one piece of information that is surely known to some people: the physical location of Osama bin Laden. Though no government in the world is enforcing any exclusive rights in that information and several powerful governments are taking action to force disclosure of the information, Mr. bin Laden has been successful in excluding others from obtaining it. It is not an exaggeration to say that the U.S. government might find it easier to possess any single square mile of land in most countries in the world than to possess that single bit of information. This problem of nonincludability cannot be dismissed as exceptional. Even if we limit consideration to wholly domestic law situations, the entire investigative branch of law enforcement can be viewed as dedicated to fighting the nonincludability problem associated with information.

The nonincludability problem has obvious implications for at least one branch of intellectual property: One justification for trade secret protection is that legal protection of secrets can decrease the expenditures that otherwise would be made to keep the material secret. In other words, policymakers accept that, with or without legal protection, the information will be kept

⁴² Yochai Benkler, *Constitutional Bounds of Database Protection: The Role of Judicial Review in the Creation and Definition of Private Rights in Information*, 15 Berkeley Tech. L.J. 535, 543 (2000).

⁴³ In a footnote, Yochai Benkler suggests that excludability should be based on “the available technology for exclusion, and the institutional (legal) framework that permits or facilitates such technically feasible exclusion.” Id. n. 23 (emphasis added). Yet, if this approach were followed, information would have to be viewed as excludable in legal systems, like ours, that grant and enforce exclusive rights in information.

secret. Legal policy is then animated not by the question whether the information should be private or public, but instead by the question whether, given that the information will be kept private, privacy can be most efficiently obtained with some form of legal protection.

Because all property could qualify as “nonexcludable” if that concept were defined by reference to the situation where the government chooses not to grant or enforce legal rights, the economics literature defines “nonexcludable” to mean situations where exclusion is so costly as to be not feasible.⁴⁴ In other words, the relevant question is whether the government could, in practice, grant and enforce exclusive rights. If it can, then the property does not qualify as nonexcludable. Under that standard definition, both land and information are excludable; there is no justification for distinguishing them.

The second justification for classifying intellectual property as a public good is more sound than the first. This approach classifies intellectual property as a public good by defining the concept of public good solely by nonrivalry. Richard Posner and William Landes take this approach by defining “‘public good’ in the economist’s sense that consumption of it by one person does not reduce its consumption by another.”⁴⁵ Other commentators in both the legal and economic literature follow this definition though, as previously noted, the practice is not uniform and the foundational work in the field of public goods such as Samuelson’s suggest a different definition.⁴⁶

Even if we accept this definition of public good, we still must ask whether information is particularly likely to be nonrival, or is more likely to be nonrival than physical property. Here we

⁴⁴ Daphna Lewinsohn-Zamir, *Consumer Preferences, Citizen Preferences, and the Provision of Public Goods*, 108 *Yale L.J.* 377, 377 (1998) (“Nonexcludability means that it is impossible or impractical to prevent those who do not pay for the good from enjoying its benefits.”); R. A. Musgrave, *Provision for Social Goods*, in J. Margolis & H. Guitton, *Public Economics* 124, 128 (1969) (defining non-excludability in terms of cost, with the polar case being where “excludability is impossible”).

⁴⁵ William M. Landes and Richard A. Posner, *The Economic Structure of Intellectual Property Law* 14 (Harvard 2003). Landes and Posner argue, however, that even under this definition, intellect property is not always a public good because it is not always purely nonrival. They view at least some copyrighted works as being similar to a congested highway, an example where the consumption of addition consumers could decrease the enjoyment of the good by other. See *id.* at 222-228 (arguing that copyrighted works may suffer from “congestion externalities”); see also William M. Landes and Richard A. Posner, *Indefinitely Renewable Copyright*, 70 *U. Chi. L. Rev.* 471, 484-488 (2003).

⁴⁶ See, e.g., William H. Oakland, *Theory of Public Goods*, in 2 *Handbook of Public Economics* 486-487 (A.J. Auerbach & M. Feldstein eds., 1987) (defining “public goods” as “involv[ing] no ‘rivalry’ among consumers” and allowing for the “possibility of exclusion”); See, e.g., Glynn S. Lunney, *Fair Use and Market Failure: Sony Revisited*, 82 *B.U. L. Rev.* 975, 993 n. 84 (2002) (citing Oakland for the proposition that “[a]lthough legal commentators are prone to defining ‘public’ goods as goods that are characterized by: (1) nonrivalrous consumption, and (2) an inability to exclude non-payers, in truth, only the first is required.”); Richard A. Posner, *Economic Analysis of Law* § 3.3 at 41(6th ed. 2003) (“A public good is a good that can be consumed without reducing any other person’s consumption of it.”)

must consider the entire realm of information, not just songs, stories and inventions, but also financial and medical data, corporate insider information, embarrassing facts, private information, industrial secrets and the like. When this whole range of possible forms of information is considered, it is not at all clear that information in itself is necessarily more nonrival than physical property. A whole branch of law — the law of privacy — is premised on a view that a person’s consumption and enjoyment of the information could depend on possessing the information to exclusion of others. Similarly, insider trading cases show another area of law where the consumption of the relevant information reduces its consumption by others. Trade secret law too involves information — like customer lists — the value of which may be destroyed if shared. In these situations, which are common, information is not a nonrivalrous good.

While it is not accurate to assert that information is nonrivalrous, it might be accurate to say that copyrights and patents apply to a subset of information that usually is nonrivalrous. The difference here is significant. Merely because intellectual property rights cover information does not necessarily mean that the rights cover nonrivalrous goods. The point can best be illustrated with a case.

Salinger v. Random House, Inc., involved the extent to which a biographer could make “fair use” of J.D. Salinger’s unpublished letters. Salinger had never assigned away the copyrights to these letters, so it was undisputed in the case that he owned the copyrights. No one could publish the letters without his permission. It was also undisputed in this case that Salinger had consistently refused to allow the letters to be published, even though the rights to publish them were estimated to be worth about a half million dollars.⁴⁷ The physical copies of the letters were owned initially by their recipients, but many had been donated to various university libraries. A biographer gained access to them only after signing agreements promising “not to copy, reproduce, circulate or publish [the letters] without the permission ... of the owner of the literary property rights.”⁴⁸ Nonetheless, the biographer copied and paraphrased parts of the letters without Salinger’s permission. Salinger sued and won an injunction barring publication of the biography.

If this case is analyzed as if the material in Salinger’s letters is a nonrivalrous public good, then the case seems wrongly decided. Landes and Posner provide the best analysis from this prospective. For nonrivalrous but excludable public goods, the justification for enforcing exclusive private rights cannot rest on a rationale that the exclusive rights serve to allocate the good between competing consumers because, if the good is truly nonrivalrous, then one person’s consumption will not adversely affect another’s. Thus, the standard assumption is that exclusive rights are provided solely as a means for providing the incentives to create the good in the first place. With this premise, the optimal copyright protection becomes a tradeoff between providing

⁴⁷ 811 F.2d 90, 99 (2d Cir).

⁴⁸ 650 F. Supp. 413, 416-17 (S.D.N.Y. 1986).

good incentives for authors to write letters (and to be candid in them!) and providing access to these letters, once they are produced, to biographers and the interested public. Landes and Posner devote an entire chapter to their analysis of this issue and develop a formal model to illustrate the relevant calculus. They conclude that weak copyright protection is likely to be optimal because “weak copyright protection [is not] likely to cause writers of unpublished materials to trim their sails much.”⁴⁹ Or, in other words, “copyright is unlikely to influence [the writer’s] behavior significantly.”⁵⁰ The focus here is on the writer’s incentives to produce the letters in the first place, which is the correct focus if the material is a nonrivalrous public good.

The facts of the case, however, are wholly irreconcilable with the view that the content of Salinger’s letters was a nonrivalrous public good. Salinger himself could have allowed the letters to be published and reaped a half million dollars. Why did he not grant permission? The facts do not suggest that he was holding out for a better offer. Rather, he refused publication because keeping the content of the letters private was worth more than a half million dollars to him. His enjoyment of the content would have been destroyed by others’ consumption.⁵¹ Thus, the content of the letters cannot be viewed as a nonrivalrous public good. With this insight, the court’s decision to protect the content of Salinger’s letters — and to restrict the normal level of fair use so that Salinger had something much more closely approximately full property rights in the contents — makes much more sense. The copyright in this case was serving a function that property often does for rivalrous goods: It was allocating rather absolute rights to the party who received greater enjoyment from the good.

Salinger is a famous but not an isolated case. The whole law of privacy can be viewed as dealing with the problem of information that is rivalrous in consumption. Even in the area of industrial innovation, where tastes for privacy might be appropriately ignored, not all information is a public good, and recognition of this is necessary to explain some major features of patent and trade secret law. Consider, for example, the obviousness doctrine of patent law. The patent law in this and other countries awards patent rights only if the applicant had made a nonobviousness advance, or equivalently an inventive step, over the prior art. One argument for not providing patent protection is that such minor, obvious advances represent nothing really new, so there is nothing to protect. But this explanation is difficult to accept because the law is also quite clear

⁴⁹ Landes and Posner, at 139.

⁵⁰ *Id.*

⁵¹ Landes and Posner assert their “doubt[] that the expected harm of publication to the author of unpublished materials is likely to be great.” *Id.* at 141. The facts of the case, however, suggest that for Salinger, the harm exceeded the amount that any biographer was willing to pay. Landes and Posner also suggest that the harm to the author of the unpublished letters should be excluded because “[s]ince the biographer doesn’t capture the consumer surplus, too few biographies are created.” *Id.* Excluding the harm to the author of the unpublished letters “counter[s] the incentive to produce too few biographies.” *Id.* This is curious reasoning. In every case, the writer of a biography or book cannot capture the full social surplus of the consumers. Subsidizing one class of books (those that rely on unpublished letters) would seem to distort the market for books. Moreover, the subsidy comes solely at the expense of one individual — the author of the unpublished letters.

that a trade secret need *not* meet patent law's nonobviousness threshold for protection.⁵² Thus, another field of law recognizes that obvious developments *may* constitute protectable intellectual property.

A better explanation is that nonobviousness doctrine reflects a recognition that not all information is public good. The public good formulation assumes that, once created, information can be reused with little or no cost. This formulation, however, ignores the costs of *transmitting and reapplying* the information. Even in an era of inexpensive communications, the costs of transmitting the information might be large where the information takes time to find and to grasp.⁵³ If these costs are considered, much of the information manufactured in firms might be private goods in the sense that the social costs of creating the information might be lower than the costs of transmitting the information. A common example of this phenomenon is well-known to all lawyers. Law firms often instruct young lawyers to beware of "reinventing the wheel." The warning is supposed to remind lawyers that they should not duplicate research already done by others. Yet young lawyers instinctively know that, while it may be a bad idea to reinvent the wheel, it's often a good idea to recreate prior research. Even if some other lawyer in the firm has previously researched the same point, finding that research takes time and, more importantly, applying the research to even a slightly different situation requires understanding the research. Very often, the costs of finding and understanding someone else's prior research may be higher than the costs of recreating the research and the concomitant understanding generated in the process of research.

This perspective sheds significant light on the nonobviousness requirement. Information failing the nonobviousness test might very well be a private good; it is best produced competitively by multiple firms who constantly create and consume the information. If that information were patented, no social gain would result. Firms would not find it worthwhile to search the patent library to find the information even after the patent expired. They would instead continue to recreate the information on their own. Indeed, this analysis suggests a good test to see whether the nonobviousness doctrine has been set at an appropriate level. If infringement actions are frequently brought against firms that unknowingly reinvented the patented information, or if patented information is not treated as a useful category that should be searched when an engineering challenge is confronted, then courts should worry that the nonobviousness standard has been lowered to the point where information not having public good aspects is being patented. Such information would be better produced competitively, without granting exclusive rights to a single firm.

⁵² See Robert G. Bone, A New Look at Trade Secret Law: Doctrine in Search of Justification, 86 Cal. L. Rev. 241, 248 (1998) ("Unlike patent law, which only protects inventions that are 'nonobvious,' trade secret law protects all inventions that confer a competitive advantage, even ones that are not especially new.").

⁵³ See George J. Stigler, *An Introduction to Privacy in Economics and Politics*, 9 J. Leg. Stud. 623, 640-41 (1980) (noting that "dissemination of such knowledge is often enormously more expensive than its production" and providing as an example the mathematical quadratic equation, for which society "invest[s] perhaps a day of every high school student's life in instructing them in its use").

Rather than asserting that information is a public good and that the structure of intellectual property law is tailored to accommodate this basic nature of information, it is much more accurate to say that the major branches of intellectual property law (copyright and patent) are generally structured to cover that subset of information having some public good qualities, particularly nonrivalrous consumption. This difference here is important. If all information is posited to be a public good — or at least a nonrivalrous good — then cases such as *Salinger*, doctrines such as nonobviousness and indeed whole areas of law like trade secret and privacy law are difficult to explain.

Yet all of this is only part of the objection to the public good characterization of intellectual property. Another, and perhaps more troubling, problem is similar to the problem with the property and monopoly theories of intellectual property. As with “property” and “monopoly,” the concept of a “public good” is malleable enough that, under some definitions, large areas of intellectual property fall within the concept.⁵⁴ Once again, however, the regulation applied to intellectual property is quite different than that applied to paradigmatic public goods. For example, national defense is frequently mentioned as a prototypical public good. Indeed, even writers who define a public good simply by nonrivalrousness still refer to national defense as a prototypical case. Yet national defense is underwritten entirely by the government, and no attempt is made to exclude citizens from enjoying its benefits.

Many commentators recognize that public goods exist on a spectrum between pure public goods (which are nonrival and nonexcludable, like national defense) and pure private goods (rival and excludable, like Blackacre). Under this approach, intellectual property is said to have “some public good aspects” or to be a “mixed” public good because it is nonrival even though it is usually excludable. This approach admits that a broad class of goods might constitute public goods or mixed public goods, and that broad class of goods may be regulated in radically different ways. But if the forms of regulation applied to intellectual property do not bear any consistent resemblance to the heterogeneous legal structures associated with the category of (broadly defined) public goods, what then is the significance of describing intellectual property as a “public good”?

⁵⁴ The malleability of the public good concept has generated a rather large number of slightly different definitions in the literature. For example, Professor Sunstein has defined a public good as one that *is, in fact*, equally available to all. See, e.g., Cass Sunstein, *Television and the Public Interest*, 88 Calif. L. Rev. 499, 534 (2000) (defining a public good to mean “something that if provided to one is also provided to all or many”). This definition requires that the power to exclude others is not exercised, even if it could be. It also does not necessarily require nonrivalry; thus, food aid distributed to all in an area of famine would qualify as a public good under this definition. Curiously, Professor Sunstein states in a footnote that even if exclusion is practiced and fees are charged for access to information, still information is still not converted into a private good. *Id.* n.159. A similar approach is taken by Robert J. Barro and Xavier Sala-I-Martin, who give as example a public highway that is open to everyone but that suffers congestion so that the enjoyment of the road by some prevents (or decreases) the enjoyment of it by others. Robert J. Barro and Xavier Sala-I-Martin, *Public Finance in Models of Economic Growth*, 59 Rev. Econ. Stud. 645, 650 (1992) (considering the case of a public highway that experiences congestion and noting that, “[i]n this case, the public good is rival but not excludable”).

As with the terms property and monopoly, the term “public good” biases the discussion in a particular fashion. For even though the term public good can encompass a broad range of goods that have various degrees of nonrivalry or nonexcludability, the economic response typically associated with public goods involves some form of governmental underwriting. Thus, for example, Robert Cooter and Thomas Ulen write:

How can public policy correct the market failure in the provision of public goods? There are two general correctives. First, the government may undertake to *subsidize* the private provision of the public good Second, the government may undertake to provide the public good itself and to pay the costs of providing the service through the revenues raised by compulsory taxation. This is, in fact, how national defense is supplied.⁵⁵

Similarly, on the very first page of his book on *The Demand and Supply of Public Goods*, Nobel prize winner James Buchanan introduces the distinction between private goods and public goods in this way:

People are observed to demand and to supply certain goods and services through market institutions. They are observed to demand and to supply other goods and services through political institutions. The first are called private goods; the second are called public goods.⁵⁶

Other such statements are ubiquitous.⁵⁷ The general assumption is that the public good will be supplied, or at least underwritten, by the government.⁵⁸ Now, of course, the writers of these passages are sophisticated enough to know that if public goods is broadly defined, then of course many such public goods are privately produced without government subsidies.⁵⁹ Statements such as the above are usually careful enough to speak in “general” terms. But these statements nicely

⁵⁵ Cooter and Ulen, *supra* note ____, at 47.

⁵⁶ James M. Buchanan, *The Demand and Supply of Public Goods* 1 (1968).

⁵⁷ See, e.g., Russell D. Roberts, *Financing Public Goods*, 95 *J. Pol. Econ.* 420, 420 (1987) (“The two ways to finance public goods are direct taxation and subsidies to private spending.”); see also R. A. Musgrave, *Provision for Social Goods*, in J. Margolis & H. Guitton, *Public Economics* 124, 128 (1969) (stating that, if even it is possible to exclude people from access to a good, “exclusion . . . should be applied only with regard to goods, the benefits of which are rival, but not with regard to goods whose benefits are non-rival”).

⁵⁸ Later in his book, James Buchanan expressly notes that “organizational-institutional implications [have been] read into the theory of public goods” with the “natural” tendency to believe that “public good should be public.” Buchanan, *supra* note ____, at 171.

⁵⁹ See, e.g., Buchanan, *supra* note ____, at 184-186 (noting that even for “purely public” good — here defined merely as a good which always can be supplied to additional users at zero marginal cost — private provision of the good through positive prices may be better than government provision of the good).

indicate the focus of the public goods literature, which is typically centered on goods that are provided by, or at least financed by, the government.

Thus, as with property and monopoly, the public good label seems to portend an agenda: As one scholar has written, “many commentators urge us to remember that because information is a public good, Congress should keep the legal restrictions on information — the intellectual property laws — as narrow as possible.”⁶⁰ But just as with the “property” and “monopoly” views, the public good description of intellectual property is either erroneous or is so general as to have little value.

II. Intellectual Property as Natural Monopoly

The major species of intellectual property law — patent and copyright law — are best considered as form of natural monopoly regulation. The legal, historical and economic parallels between intellectual property and natural monopoly law are strong, and a comparison of the two fields is valuable because it can provide new insights into the operation of both areas. Drawing this parallel does not require any assertion that information inherently possesses some special economic attribute that differentiates it from all physical property and thereby dictates the contours of intellectual property law. To the contrary, the point here is that the gap between physical property and intellectual property is much less than it may first appear. This does not mean that intellectual property law looks much like the traditional property law as applied to Blackacre. Intellectual property law regulates a species of information that has natural monopoly characteristics. Where physical property has such characteristics, owners generally do not possess fee simple rights unencumbered by regulatory limitations. The same holds true for informational property that has natural monopoly characteristics.

A. The Basic Symmetry

A “natural monopoly” is typically said to be present in any market where the average cost of providing a good continues to decline throughout the relevant range of demand. Such “declining average cost” conditions occur, for example, where the fixed costs averaged over each unit of production are relatively high compared to marginal cost.⁶¹ Production and distribution of electricity has frequently been considered a case of declining average cost because an efficient electrical plant and distribution system require large capital expenditures but, once such a plant and distribution system are built, the marginal cost of producing and delivering electricity is relatively small.

⁶⁰ I. Trotter Hardy, *Not So Different: Tangible, Intangible, Digital, and Analog Works and Their Comparison for Copyright Purposes* 26 Dayton L. Rev. 211, 245 (2001).

⁶¹ See, e.g., RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* § 12.1, at 317 (3rd ed. 1986). Declining average costs will always occur where the per unit fixed cost is greater than marginal cost. It is also always present whenever the marginal cost is constant or decreases throughout the range of production.

There are two important implications of declining average costs. First, the classic condition of perfectly competitive markets — a market price at marginal cost — is impossible to achieve as a long run equilibrium without some sort of governmental subsidy. The reason for this result is that, where average costs are declining, the marginal cost of production is always *below* average cost. Thus, if the market price of the good were driven to marginal cost, producers would be unable to recover their fixed costs and they would not enter the industry in the first place. A second corollary of declining average costs is that concentrating all production in a single firm is more efficient than having multiple firms undertake production.⁶² The intuition supporting this result for high-fixed-cost industries is that, by concentrating all production in a single firm, that firm is able to realize an economy of scale by spreading the fixed costs over more units of production. Thus, having a single firm avoids wasteful duplication of the fixed costs of production.

The class of property known as intellectual property generally — particular the core of traditional intellectual property, patent and copyright law — covers a special case of goods with declining average cost. The fixed costs associated with producing intellectual property are the costs associated with the intellectual or creative effort. They are the costs of writing a book, filming a movie or researching and developing an innovation. Once the intellectual property has been created, the marginal cost of using it an additional time is very low; in fact, in most cases, it is essentially zero. The zero marginal cost of copyrighted works is easy to see in the modern world where photocopiers, computers and VCRs are ubiquitous. The information contained in a patent also generally has a zero marginal cost of reusing it even though the marginal cost of producing the good embodying that information may be substantial. Thus, if a patent covers a new type of car, the production costs of making each car might be substantial, but the social costs of using the patented information in the making of the car are zero or very close to zero.

Recognizing the zero marginal cost of intellectual property is consistent with the point made earlier that intellectual property generally covers a species of information that is nonrivalrous. The concept of nonrivalry is really identical to zero marginal cost, for we could just as easily imagine that, each time a person uses intellectual property, the intellectual property is consumed but another copy is instantaneously produced at zero cost.⁶³ Indeed, precisely because nonrivalry and zero marginal cost are identical concepts, many traditional natural

⁶² This point follows directly from the assumption that average costs decline throughout the entire range of demand. The total cost of satisfying the demand (i.e., the sum of each firm's average cost times the quantity produced by the firm) will be minimized by having all the production undertaken by one large firm, which will have lower costs than smaller firms.

⁶³ The marginal cost of the intellectual property should not be confused with the marginal cost of a product incorporating the intellectual property. Thus, the marginal cost of the intellectual property incorporated in, for example, an innovative new mousetrap is not the cost of producing an additional mousetrap according to the invention, but merely the cost of supplying the information so that the new mousetrap can be produced. If the production of the actual mousetraps is not subject to declining average costs, the supplier of intellectual property — i.e., the patent holder — can license several different competing firms to construct the physical devices and charge those firms a royalty that will have the effect of raising price and constraining output.

monopolies have been viewed as a particular type of public goods. As Paul Samuelson noted, “[w]here marginal cost pricing will not cover full costs, we encounter a version of the public-good problem.”⁶⁴ Samuelson used an electric grid to provide an example:

Suppose two (or more) families have summer homes out on a peninsula. The cost of bringing electricity to them has separable-cost and joint-cost aspect. Thus, each can properly be charged the separable marginal cost of the electricity that he draws from outside the system. But the common cost of the line is a public good.⁶⁵

It is important to recognize, though, that natural monopolies are merely “*a version* of the public-goods problem.” The difficulty with analogizing intellectual property to public goods generally is not inaccuracy. Rather, the category of public goods is so large that it provides little useful insight into the structure of, and policies underlying, intellectual property. For insight, we must look to the specific category of public goods that most resembles intellectual property.

Because intellectual property embodied in patents and copyright has zero marginal cost of use, it is subject to the phenomenon of declining average costs just like other natural monopolies. It is also subject to the two effects discussed above, the unsustainability of marginal cost pricing and the desirability of a single producer. Indeed, these effects are familiar. First, a market price at marginal cost is not a long run equilibrium solution without some sort of governmental subsidy. Marginal cost pricing for intellectual property could be achieved by, for example, refusing to grant exclusive legal rights but, as in the case of other natural monopolies, that policy would mean that producers would not be able to recover their fixed costs and thus would not produce the property in the first place. The second effect also holds: By concentrating production of an intellectual property good in a single firm, that firm is able to realize an economy of scale by spreading the fixed costs over more units of production. Thus, having only one producer avoids wasteful duplication of fixed costs.

B. Government Policies for the Declining Average Cost Phenomenon.

As Milton Friedman famously noted in 1962, the phenomenon of declining average cost presents the government with “a choice among three evils: [1] private unregulated monopoly, [2] private monopoly regulated by the state, and [3] government operation.”⁶⁶ Once intellectual property is viewed as a natural monopoly, we can see that Friedman’s choices should be stated in

⁶⁴ P.A. Samuelson, *Pure Theory of Public Expenditure and Taxation*, in *Public Economics* 98, 118 (J. Margolis & H. Guitton, eds., 1969).

⁶⁵ *Id.* at 115-116. See also C. Edwin Baker, *Giving the Audience What It Wants*, 58 *Ohio St. L.J.* 311, 317 (1997) (“Typically, utilities or other “natural” monopolies exhibit this public good quality in their infrastructure, which multiple consumers can use with no or very small extra expense.”).

⁶⁶ MILTON M. FRIEDMAN, *CAPITALISM AND FREEDOM* 128 (1962).

slightly more general terms as (1) unlimited private rights, which will lead to private monopoly; (2) partial private property rights, which include governmental limitations ranging from price regulation, fair use restrictions, and temporal limitations on rights; and (3) public ownership, either with or without public subsidies.

These categories are, of course, not wholly discrete. For example, categories (2) obviously blends into (1) as the regulatory restrictions on property rights are decreased. Also, categories (2) and (3) are sometimes combined; the government sometimes subsidizes the creation of the natural monopoly property but still allows private parties to take ownership of some partial interest. For example, in the nineteenth century, the government provided massive subsidies, often in the form of land, as a means of encouraging the construction of railroads.⁶⁷ Once the railroads were built, however, they were private, not public, properties (though the properties were soon subjected to price regulation). A modern equivalent is the Bayh-Dole Act, which allows federally funded researchers to obtain and keep the patents derived from their research.⁶⁸ But of course, patents obtained from federally sponsored research, like all patents, are limited in time and scope. Predictably, policies such as the railroad subsidies and Bayh-Dole lead to controversy over whether the government is giving away benefits to private interests with no corresponding benefit to the public.⁶⁹

Both for intellectual property and for traditional natural monopolies, governments have usually adopted some variant of choice number (2), granting some sort of partial private property right. At first glance, the forms of the partial property rights conferred might seem wholly different in intellectual property as compared to traditional natural monopolies. Property rights in traditional natural monopolies — natural monopolies in the physical world — are generally curtailed by administrative price regulation. In intellectual property, price regulation seems to be the exception rather than the rule; instead, property rights are curtailed by finite property term and qualitative limitations on the scope of the right. But these apparent difference are less dramatic than they seem. Let us examine the limitations on property rights that occur in both fields.

⁶⁷ Carter Goodrich, *GOVERNMENT PROMOTION OF AMERICAN CANALS AND RAILROADS: 1800-1890* (1960). See also Salvatore Massa, *Surface Freight Transportation: Accounting for Subsidies in a "Free Market,"* 4 *N.Y.U. J. Legis. & Pub. Pol'y* 285, 292-93 (2000/01); Abraham Bell and Gideon Parchomovsky, *Givings*, 111 *Yale L.J.* 547, 576-77 (2001).

⁶⁸ Act of Dec. 12, 1980, Pub. L. No. 96-517, 94 Stat. 3015-3028 (codified as amended at 35 U.S.C. 200-211, 301-307 (1994)). See generally Rebecca S. Eisenberg, *Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research*, 82 *Va. L. Rev.* 1663 (1996).

⁶⁹ *Peter S. Arno and Michael H. Davis, Why Don't We Enforce Existing Drug Price Controls? The Unrecognized and Unenforced Reasonable Pricing Requirements Imposed upon Patents Deriving in Whole or in Part from Federally Funded Research*, 75 *Tul. L. Rev.* 631 (2001) (criticizing the government for failing to assert its rights under the Bayh-Dole Act to limit the prices of patented drugs developed with federal subsidies).

1. *Limited Term of Right, and a Problem of Waste.*

The limited term of property rights is a major feature of intellectual property law. The property rights in both patents and copyrights last only for a term of years; indeed, the Constitution itself imposes the requirement of “limited times” for these branches of intellectual property. This requirement constitutes one of the most obvious ways in which intellectual property deviates from the property interests typically conferred in the paradigmatic slice of physical property, fee simple in Blackacre. In property law, the historical alternative to perpetual rights has been the usufruct, which grants a package of land use rights “that terminate when the usufruct’s owner dies or ceases the use.”⁷⁰ But as Professor Ellickson has noted, “the inherent efficiencies of perpetual private land rights have led to their spontaneous appearance on every continent,” and usufructs have generally been limited to primitive, “preliterate” societies.⁷¹ It is therefore surprising to find limited term grants of property on the very forefront of high technology.

Nevertheless, the prevalence of limited terms for intellectual property does not indicate some fundamental difference between the physical and the intellectual. In fact, where physical property exhibits natural monopoly characteristics, the limited property term has been one method for regulating property rights. A good example is the system of exclusive franchises awarded in the eighteenth and nineteenth centuries to encourage the production of bridges. These franchises were limited in time and frequently included the condition that, at the end of the franchise period, the physical property became public property.⁷² Nor is this mode of regulation merely a historical phenomenon. The exclusive franchise granted for the construction of the Channel Tunnel also terminates after a period of years, at which time the tunnel will become public property no longer subject to the leasehold of the franchisee.⁷³ In fact, limited term franchise agreements have recently enjoyed a renaissance because of a variety of factors, including “[t]ight budgetary constraints, a growing faith in the virtues of privatisation, the extent

⁷⁰ Robert C. Ellickson, *Property in Land*, 102 Yale L.J. 1315, 1364 (1993).

⁷¹ *Id.* at 1367, 1368 & 1370-71.

⁷² See, for example, *Charles River Bridge v Warren Bridge*, 36 US 420, 423 (1837) (quoting the 1785 charter granted to the Charles River Bridge Company, which provided for a forty-year term and required that “at the end of the said term, the said bridge shall revert to, and be the property of, the commonwealth”).

⁷³ Title to the tunnel actually vested in the government as it was being built, but the holder of the franchise or concession, Eurotunnel, also received a leasehold in the property for the duration of the concession. See Channel Tunnel Act 1987, Ch. 53, s. 7 (providing that “[t]he land comprising the seaward section of the tunnel system shall, as it becomes occupied by or on behalf of the Concessionaires working from England, vest in the Secretary of State” but that the interest in the tunnel covered by the “Concession lease ... shall vest in the Concessionaires under the lease as that land becomes so occupied, as if granted by the Secretary of State immediately on the vesting in him of that land”). The original franchise period was 55 years, but it was extended to 65 and then 99 years.

and urgency of the need for new infrastructure, ... and the rapid evolution of the techniques and skills deployed in structures of this kind.”⁷⁴

Still, although limited term franchises are not unknown in traditional natural monopoly regulation, they are less common in that field than in intellectual property law. The reason for the difference was discerned quite clearly in an early twentieth century treatise on municipal franchises, which noted “two objections” to having franchise property “revert to the city at the expiration of the grant without compensation”:

One is that the company will inevitably let its plant and property run down toward the close of the franchise period, so that the public will suffer from bad service and the city will come into the possession of an out-of-date equipment, needing immediate and complete rehabilitation. This condition runs counter to reason and good public policy with reference to public utilities. In the second place, the plan of paying for the entire plant out of earnings within a comparatively short period many necessitate higher rates than are consistent with public welfare and justice to the present generation. This objection is not as strong as the other one While it may be an injustice to require this generation to make a present to the next one of a fully-equipped gas plant or street railway system, nevertheless it is a much greater injustice for this generation to saddle upon the next one a debt for a public utility plant that has been used up or has gone out of date or is so inadequate that it has to be rebuilt.⁷⁵

Thus, the problem of waste provides a compelling reason to avoid having a limited property term that ends in public ownership of the property. Where waste is not a problem — where the costs of maintaining the property are low — limited franchises are a more common regulatory mechanism. Bridges provide one example where maintenance costs may not be unduly high, and limited franchises for bridges were a common regulatory method in the eighteenth and early nineteenth century. Intellectual property is, of course, another type of property that generally does not involve high maintenance costs, and this fact explains the prevalence of strictly limited franchises in patent and copyright.

A theory of waste also explains the exceptions to the general reliance on limited terms for patents and copyrights. For example, patent law includes the doctrine of so-called “lost art,”

⁷⁴ Christopher Clement-Davies, *Public/private partnerships on Central and Eastern Europe: Structuring the concession agreement*, 1 Bus. L. Int'l 18, 18 (Sept. 1999). See also Eduardo Engel, Ronald Fischer and Alexander Galetovic, *A New Approach to Private Roads*, 25 Regulation 18 (Fall 2002) (noting that recent road project in California and elsewhere have relied upon a franchising system under which the franchise owner builds the road, collects tolls during the franchise term, and then transfers the road to the government “usually 20 to 30 years” after the start of the franchise). [Review: Eduardo Engel, Ronald Fischer, and Alexander Galetovic, *Least-present-value-of-revenue Auctions and Highway Franchising*, 109 J. Pol. Econ. (no. 5) (2001).]

⁷⁵ 1 Delos F. Wilcox, *Municipal Franchises* 63 (McGraw-Hill, New York 1910).

which allows an inventor to patent an old invention if knowledge of a prior invention has become so “completely lost as if it had never been discovered.”⁷⁶ This doctrine recognizes that, if the maintenance costs on a piece of prior art were so high that it failed to survive in the public domain, then no good objection can be raised to granting new property rights over the subject even though this will add a new period during which the information will be out of the public domain. The forgotten invention is like an old poorly maintained bridge that has been washed out in a storm. In such a case, a new bridge will have to be built, and granting a limited franchise is one plausible regulatory tool to provide the proper incentives. The waste problem also explains why the Patent Office keeps a library of all issued patents. In effect, the government subsidizes (though at trivial cost) the maintenance of the knowledge previously obtained through patents; this practice eliminates any problem of waste as to issued patents.⁷⁷

Copyright law offers another example of waste in the public domain. In enacting the Sonny Bono Copyright Term Extension Act,⁷⁸ Congress believed that the imminent termination of copyright protection was causing film studios to allow some old films to waste away; the term extension was said to be a remedy for this problem because it would “provide[] copyright owners generally with the incentive to restore older works and further disseminate them to the public.”⁷⁹ In sustaining the constitutionality of the extension, the Supreme Court relied in part on this justification, holding that Congress “rationally credited projections that longer terms would encourage copyright holders to invest in the restoration and public distribution of their works.”⁸⁰ Although many commentators have been justifiably skeptical that this justification was an important reason for the copyright extension, still the point here is merely that the economic

⁷⁶ *Gayler v. Wilder*, 51 U.S. (10 How.) 477, 495 (1850). The Supreme Court held that a prior art safe built by a man named Connor, which had been lost and whose details of construction had been forgotten by all concerned, did not anticipate the patented invention at issue in the case:

[I]f the Connor safe had passed away from the memory of Connor himself, and of those who had seen it, and the safe itself had disappeared, the knowledge of the improvement was as completely lost as if it had never been discovered. The public could derive no benefit from it until it was discovered by another inventor. And if [the patentee] made his discovery by his own efforts, without any knowledge of Connor’s, he invented an improvement that was then new, and at that time unknown

Id. at 495.

⁷⁷ Historically, the problem of waste was not necessarily eliminated by the Patent Office’s library. In 1836, a large fire destroyed the Patent Office. Many of the thousands of patents issued between 1790 and 1836 were lost forever, thus creating the theoretical possibility that the lost art doctrine could have lead to the grant of two patents. See Merges & Duffy, *Patent Law and Policy* 22 n.2 (2002) (discussing the incompleteness of the Patent Office’s records due to the 1836 fire); see also <http://www.uspto.gov/web/offices/com/speeches/01-60.htm> (noting that only about 28% of the 10,000 patent issued prior to 1836 were recovered).

⁷⁸ 112 Stat. 2827 (1998).

⁷⁹ H. R. Rep. No. 105-452, p. 4 (1998).

⁸⁰ *Eldred v. Ashcroft*, 537 U.S. 186, 207 (2003).

theory of waste is a sound justification for moving away from limited property right terms, and it is a symmetry between traditional natural monopoly regulation and intellectual property law.

Finally, patent law gives one last example where limited terms may be detrimental. As Professor Eric Kades has noted, patented antibiotics are a wasting asset.⁸¹ While the information in the patent is never lost, its value does decline because the commercial exploitation of the patent leads to resistant strains of microorganisms. Early in the life of a patented antibiotic, the patentee will have an incentive to restrict sales of the antibiotic or take other precautions to minimize the growth of resistant strains. But not so in the end of the patent life. The growth of resistant strains might be viewed as similar to a corrosion on a physical property, and towards the end of the patent term, the patentee will have little incentive to maintain the property against this corrosion. The remedy that Professor Kades suggests for this problem is the precisely same remedy that was and is used for wasting properties in physical natural monopolies — infinite term franchises.⁸²

2. Limited Scope of Right: Price Regulation and Rights Regulation.

Another common technique in regulating a natural monopoly is to allow infinite property terms but to impose limitations on the rights during the term. In traditional natural monopolies, the limitation is usually manifested as price regulation by an administrative body. In intellectual property, the limitations tend to come in the form of qualitative restrictions such as the fair use doctrine, experimental use rights, and other restrictions on what is generally called the scope of rights. For both price regulation and rights regulation, the overarching goal is the same — to decrease the static deadweight loss associated with the property while still allowing the property owner a fair possibility to recover the fixed costs of creating the property.

The broad equivalence of these two techniques can best be seen in a 1999 article by Ian Ayres and Paul Klemperer.⁸³ Ayres and Klemperer propose decreasing the scope of patent rights by introducing more uncertainty into patent enforcement. Patents would be enforced only in probabilistic damage actions at the very end of the patent term. In a certain percentage of randomly selected suits (for example, 25%) the court would dismiss the infringement charge. In the other remaining suits (the other 75%), the court would calculate and award the damages that the infringer caused the patentee. The introduction of uncertainty would create an incentive to infringe, and thus during the life of the patent, the patentee would endure some competition from infringers, which would push the price below the monopoly level. Price would still be held above marginal cost because the infringers would have to store up sufficient profits so that their

⁸¹ See Eric A. Kades, *Plagues, Policy, & Patents: Addressing Overuse of Antibiotics*, ___ Nw. U. L. Rev. ___ (forthcoming 2005) (available at http://papers.ssrn.com/so13/papers.cfm?abstract_id=387241).

⁸² *Id.*

⁸³ Ian Ayres and Paul Klemperer, *Limiting Patentees' Market Power Without Reducing Innovation Incentives: the Perverse Benefits of Uncertainty and Non-injunctive Remedies*, 97 Mich. L. Rev. 985 (1999).

expected profits from the business (sales profits minus the expected damages) would not be negative. Yet holding price slightly below the monopoly price is likely to increase social welfare because, as Ayres and Klemperer note, the “last bit of monopoly [pricing] is so disproportionately damaging” — it increases the monopolists profits very little while inflating deadweight loss the most.⁸⁴ As Ayres and Klemperer emphasize, they are not trying to diminish the patentee’s overall rents or decrease the incentives for innovation. To offset the reduction in monopoly price, they propose “to hold patentees’ profits constant by appropriately lengthening the patent life.”⁸⁵

At first blush, the Ayres and Klemperer article seems to uncover a new and previously unrecognized benefit of uncertainty in intellectual property rights. Their ultimate conclusion is “not that judges should actually flip weighted coins in adjudication,” but rather that “understanding the perverse benefits of uncertainty and delay suggests that policymakers may not want to spend inordinate amounts of money to eliminate certain types of uncertainty and delay from the current system.”⁸⁶ The tenor of the article is that uncertainty in patent enforcement produces unusual benefits. But in fact Ayres and Klemperer have really uncovered nothing new.

The key to understanding the Ayres and Klemperer system — and to seeing that this system is merely price regulation by another name — is to recognize that all market participants are setting their prices based on the expected damages to be awarded by courts. If firms expect the courts to award high damages, the market price will be higher because infringers will have to store up more to pay their expected damage bills. If firms expect courts to award lower damages, the market price will be lower. Uncertainty in enforcement is irrelevant. Indeed, the Ayres and Klemperer article makes clear that precisely the same effect could be achieved if courts awarded with certainty “partial damages” — a percentage (e.g., 75%) of the full monopoly damages — rather than awarding full monopoly damages in a random percentage of the cases. A “partial damages” regime is easily seen as compulsory licensing system; the only difference is that the royalty is established *ex post* rather than *ex ante*. “Infringers” (compulsory licensees) are allowed to make the patented product provided that they pay “damages” (a government-calculated *ex post* royalty) to the patentee. Thus, the Ayres and Klemperer proposal has nothing to do with uncertainty; it is just a form of compulsory licensing, which is a type of price regulation.

All of this can be summarized very simply: The price of a patented technology is directly related to the scope of expected remedies afforded by infringement courts. If injunctions and full monopoly damages are provided (broad scope), the price for the patented technology will be at the full monopoly level (limited, of course, by the market substitutes for the invention). If courts afford less than full damages (more narrow scope), then the price for the invention will be below

⁸⁴ *Id.* at 990.

⁸⁵ *Id.* at 991.

⁸⁶ *Id.* at 995.

the monopoly level. Limitations on scope of right can thus function equivalently to price regulation.

Yet even if the Ayres and Klemperer system can be unmasked as merely compulsory licensing by another name, we are still left with the question why a patent system does not include a system of compulsory licensing such that any party can use a patented technology provided that he or she pay to the patentee partial damages. Or, to rephrase the question, why does the patent system specifically, and intellectual property law more generally, generally not rely on price regulation? The answer to this question cannot be that IP creators will be under-rewarded. As noted by Ayres and Klemperer as well as many others, price regulation or some other diminution in scope can be offset by an increase in the duration of the right. If, for example, patent life were extended while patent scope narrowed (via limited damages or price regulation), patentees could recoup the same amount of net rents. Indeed, as Ayres and Klemperer suggest, trading longer lives for a narrower scope seems to be socially beneficial because the longer but narrower right will impose less deadweight loss on society.⁸⁷

Optimal monopoly regulation would seem to require narrow but infinite property rights. Price regulation is one method to achieve a narrow right and, in fact, price regulation is not unknown in intellectual property. Compulsory licenses of copyright include royalties fixed by government action, and patented pharmaceuticals are subject to direct administrative price regulation in many developed nations (such as Canada). Still, price regulation in intellectual property is more the exception than it is in traditional natural monopolies. Conversely, although traditional natural monopolies are often burdened with more qualitative restrictions on their rights — including restrictions on rate discrimination, universal service duties, and obligations to provide free service in particular instances — price regulation remains a central feature of regulatory policy regulation in traditional natural monopolies.

The break in symmetry between the two fields is, once again, not due to any fundamental difference between the physical and the intellectual. It is instead attributable to several more subtle effects. First, pieces of intellectual property often have many reasonable substitutes. In the case of copyrighted works, this point is familiar, but even for patents, substitutes in terms of older technology or alternative technologies (possibly covered under different patents) compete in the marketplace. This competition may not be sufficient to drive price down to marginal cost, but it may be sufficient so that, if the patentee raises prices above a certain level, demand for the patented good begins to drop very rapidly. In these circumstances, competition will prevent the patentee from raising price to the point where patentee profits come with a high cost in deadweight loss. Since price regulation is not designed to push price to marginal cost but only to

⁸⁷ The basic reason is for this effect is that, if the patentee's reward is increased by lengthening the term, every incremental increase in the reward raises deadweight loss by a fixed amount. By contrast, if the patentee's reward is increased by broadening the patent's scope (i.e., by allowing the price to rise closer to the monopoly level), then each additional incremental increase in the reward raises deadweight loss by an *increasing* amount

lower price below the full monopoly price, competition may serve as the check that otherwise would be imposed by price regulation.

A further reason is that, even for those pieces of intellectual property not subject to much competition from substitutes, the traditional mechanisms of price regulation will not work well because of one feature commonly seen in intellectual property — a high level of risk. Traditional price regulation typically requires administrative determinations of the capital investments expended in creating the property and the market rate of return for such investments. This approach may be adequate for investments that are not so risky — like investments in the physical plant of a local telephone or electrical distribution system.⁸⁸ In those cases, the capital investments prudently made in building the system can be measured, and the market rate of return, adjusted for risk, can usually be estimated within a moderate range.⁸⁹ For intellectual property, both the capital investments made and the rate of market return for such investments will be much more uncertain. For example, developing a single drug may cost tens of millions of dollars if investments in that one drug are considered, but the costs can run over a billion dollars if investments in all failed projects are considered.⁹⁰ The level of risk in such investments will very much depend on the field of technology, and estimates of riskiness will have a much larger range.

Riskiness of investment — and the difficulty in measuring it — provides a good reason that *traditional price regulatory* mechanisms are not seen much in intellectual property. But why does intellectual property law not use nontraditional mechanisms similar to the “partial damages” system put forward by Ayres and Klemperer? The answer is quite simple: The courts are unlikely to have enough information to determine the damages necessary for the system to function. In calculating damage awards under a partial damages system, the courts would need not only information about the number of units sold by the patentee and infringers and the prices charged by each; they will need to know the *shape* of the demand curve for the product and the

⁸⁸ A. Lawrence Kolbe and William B. Tye, *The Duquesne Opinion: How Much "Hope" is There for Investors in Regulated Firms?*, 8 *Yale J. on Reg.* 113, 120 (1991) (noting that traditional “utilities are virtually always public monopolies dealing in an essential service, and so relatively immune to the usual market risks”).

⁸⁹ See *id.* at 153 (noting that traditional price regulation does not work well where the rate of return must be adjusted for risks and “the probability and expected size of the loss are difficult to determine”). See also Stephen Breyer, *Regulation and Its Reform* 43-44 (1982) (noting the difficulties of measuring risks which “vary widely among industries over time”); *id.* at 47 (noting that “betas” — a measure of a common stock’s riskiness which is necessary for setting a regulated rate of return under the capital asset pricing model — “are notoriously difficult to measure”).

⁹⁰ Peter Landers, *Cost of Developing a New Drug Increases to About \$1.7 Billion*, *Wall Street Journal* (12/8/2003). For risky investments, an administrative agency could count all investments (successful and failed) and then apply a relatively modest rate of return. The difficulty with this course is Averch-Johnson effect — the incentive of price regulated firms to make excessively heavy capital investments. To combat this effect, agencies often allow only prudently made investments, but it might be exceedingly difficult to distinguish between prudent and imprudent failures to innovate.

marginal costs of production.⁹¹ The courts would need this information because they will have to calculate the “monopoly price” for the innovation — i.e., the price that the patentee could have charged *if* the patent right had been fully enforced. That price would be — to put it mildly — difficult to determine because *no one will have ever paid that price!* If some government actor (the courts or another entity) could determine that price, then the government could impose a direct price cap: The government could simply require that prices for patented goods not exceed a certain percentage (e.g., 75%) of the monopoly price.

In sum, price regulation is best viewed as a species of scope-of-rights regulation. The overarching goal of scope-of-rights regulation is to constrain the monopoly that occurs at any given time. The broad similarity between intellectual property and natural monopoly regulation is that both areas have scope of rights limitations in one form or another. Still, the two fields rely predominately on different forms of scope regulation. This break in symmetry is due to subtle differences between intellectual property and the properties typically regulated with natural monopoly regulation.

3. Competition for Rights.

A third method of regulating natural monopoly is to foster competition for an exclusive franchise with the bidders competing to minimize their own monopoly profits and maximize the benefits received by the public. This form of competition is generally known as a Demsetzian auction; the name memorializes the inventor of the system, Harold Demsetz. His insight was to see that, even if it is efficient to have only one firm serve the entire market and thus competition is not possible once that firm begins serving the market, competition is still possible *before* any one firm is selected as the monopolist. Demsetz proposed that the government should hold a particular type of auction for the exclusive franchise to serve the market; “bids” in this auction would be made in terms of the price and quality of the service that would be offered to consumers if bidder were selected as the exclusive franchisee. Thus, one bidder might offer to provide electric service for \$.10 per kilowatt-hour; another, \$.07 per kilowatt hour. The bidders would therefore be competing against each other to diminish their monopoly rents. The government would recognize as the winner the firm that had offered the best terms to the public. The competition to gain the exclusive right would be harnessed to reduce private rents while increasing social surplus.⁹²

⁹¹ Ayres and Klemperer argue that, by increasing uncertainty in the patent system, “*lawmakers* can improve on the current regime without knowing the exact shape of the demand and cost curves.” *Id.* at 1008-09 (emphasis added). This is a bit confusing, however. While *lawmakers* would not have to know the shape of the demand and cost curves, *courts* making the damage awards would. The necessity for courts to know this information can be seen earlier in the article, for when Ayres and Klemperer calculate the damages necessary to induce a certain level of competition, they need to make assumptions about the shape of the demand curve and the marginal costs of production in order to complete the calculation. See *id.* at 995-97.

⁹² See Harold Demsetz, *Why Regulate Utilities?*, 11 *J L & Econ* 55 (1968). If price discrimination is not possible and elasticity of demand is positive, deadweight loss will not be eliminated because the bidders will offer service at their average, not marginal, costs.

This sort of competition has been tried for physical natural monopolies, though the results have been mixed.⁹³ The main difficulty is that, once the franchisee has been selected, governments have had trouble enforcing the franchisee's "bid" — i.e., the terms of the regulatory contract that the franchisee accepted as a condition of getting the franchise. Franchisees may skimp on quality in ways not contemplated by the original regulatory contract or lobby the government to make the terms more favorable to the franchise.⁹⁴ Because of these drawbacks, a Demsetzian auction has remained a theoretically interesting but rarely attempted regulatory device.

Demsetz's basic insight can also be seen at work in intellectual property law, particularly in patents. Through many legal doctrines, the patent system encourages prospective inventors to race to the Patent Office. Most importantly, the system permits (indeed, encourages) patentees to patent very early in the development of the patented technology, before the patentee has spent much to develop or perfect the invention. The race often means that the early years of patent exclusivity are largely worthless because the patented device has not even been marketed yet. Inventors would be better off waiting and getting their patents later, after their development is done and the patented technology is just ready to come onto the market. But they can't wait because of the threat of competition. Thus, inventors compete to file early and, in so doing, they compete to have their patents expire soonest. A patent race can be seen as a competition to place the technology into the public domain soonest. Though the competition is based on time rather than on price, still the patent race produces the same effect as the Demsetzian auction: The winner is the party who will maximize the public benefit by shouldering the term most favorable to the consumer (i.e., the lowest price or earliest expiration date).

Again we see here a broad similarity between natural monopoly and intellectual property regulation. And again the chief difference is a subtle one of emphasis and prevalence. Though Demsetzian competitions are rare in traditional natural monopolies, patent racing and the encouragement of that behavior is built into the very fabric of the patent system. Identifying the reason for the difference is challenging. Competitions to innovate are supported and policed by antitrust enforcement authorities, which preclude attempts to monopolize or curtail the competition that exists in what is termed "innovation markets" (where this sort of competition occurs).⁹⁵ Yet, the antitrust authorities maintain of innovation competition cannot explain the

⁹³ See Kip Viscusi, et al., *Economics of Regulation and Antitrust* 395-431 (2000) (discussing the virtues and vices of franchise bidding as a regulatory device).

⁹⁴ See *id.* at 407-08 (noting the incentives for franchisees to engage in opportunistic behavior). See Oliver E. Williamson, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting* 359 (Free Press 1985) (detailing the extensive revisions to the franchise terms approved by the Oakland City Council after the award of the franchise).

⁹⁵ See U.S. Department of Justice and Federal Trade Commission, *Antitrust Guidelines for the Licensing of Intellectual Property* § 3.2.3, 5.1 at 11, 24 (Apr 6, 1995), online at <http://www.usdoj.gov/atr/public/guidelines/ipguide.htm> (defining the concept of an "innovation market" to include all "research and development directed to particular new or improved goods or processes" and noting that horizontal restraints among competing firms may be

prevalence of competition for gaining innovation rights but not for securing Demsetzian franchise rights. Attempts to monopolize or to collude in a Demsetzian auction is also likely to trigger antitrust prosecution.

A more substantial reason for the difference is that, after the competition ends in a patent race, the government has a much easier regulatory bargain to enforce. The patentee can charge any price; only the expiration date of the patent needs to be enforced. Even here we can see that the difference is not so great. Just as franchise winners of Demsetzian auctions try to lobby government to modify the terms of the bargain, so too do intellectual property owners have an incentive to lobby to be relieved of their expiration dates.⁹⁶

III. The Basis for Regulation: A Theory of Partial Property Rights.

The overarching similarity between natural monopoly regulation and intellectual property regulation is that both fields allow private property but sharply curtail the ordinary rights associated with property. The policy reasons for allowing private property rights in both fields is the same: The rights encourage the development of the property in the first place. If too few rights are conferred, the investment necessary to create the property (be it an electrical distribution system or a new drug) will not be made. The policy reason for curtailing rights is much more elusive. This section examines three justifications for curtailing rights; only the third is found satisfactory.

1. *Eliminating Deviations from Marginal Cost.* One justification for imposing special restrictions on natural monopolies and intellectual properties is that the rightholder will price access to the property above marginal costs. Supra-marginal-cost pricing has long been viewed as problematic because classical economic theory suggests efficient prices for goods must remain at the marginal cost for the good. Thus, so the justification goes, government intervention and perhaps even subsidies are necessary to drive price to marginal cost.

This justification for regulation is deeply flawed. First, supra-marginal-cost prices are ubiquitous in our society. Consider even the very simple example given in our earlier discussion of public goods: A commentator contrasted information with an ice cream cone. The earlier discussion pointed out that the two were not always different because information is often a private good, like an ice cream cone. Here we can make the different point. Like intellectual property, ice cream cones are often priced above marginal cost. Consider an ice cream store that sells cones for \$2.00. Is this the marginal cost of the cone? Certainly not if we consider only the short run marginal costs. The store must not only buy ice cream and cones and pay for someone

per se illegal where those agreements allocate markets or reduce output).

⁹⁶ See, e.g., Uruguay Round Agreements Act, 108 Stat. 4809, 4973-4990 (1994) (retroactively extending patent terms for many patents remaining in force on June 8, 1995); see *Eldred v. Ashcroft*, 537 US 186 (2003) (sustaining the constitutional validity of the retroactive copyright extension contained in the Copyright Term Extension Act, Pub. L. No. 105-298, 112 Stat. 2827 (1998), amending 17 USC § § 302, 304).

to put the two together. It must also cover its fixed costs of rent and overhead. These costs could and do drive prices far above the marginal cost of making another cone.

The ice cream cone is but one example. Others abound. The marginal cost of a transcontinental flight is trivial, but tickets cost hundreds of dollars. Even heavily regulated, government-operated entities such as passenger railroads do not charge marginal costs for carriage. Moreover, even if the government wanted to eliminate deviations from marginal cost pricing, it could not do so. The government would have to subsidize firms' fixed costs in order for them to reach marginal cost pricing. But then to pay the subsidies, the government would have to raise revenue through taxation, which also causes prices to deviate from marginal costs.

For these and other reasons, governments have not pursued marginal cost pricing perfection in regulating natural monopolies.⁹⁷ Despite recent suggestions that they should do so, governments have also not pursued marginal cost pricing policies for intellectual property.⁹⁸

2. *Deadweight Loss Minimization.* A better argument for curtailing property rights is that, for monopolies, the government should regulate to push the monopolist's price closer to marginal cost. This justification is well summarized by Justice Breyer, who writes that, for natural monopolies, the government regulator will impose a price below the monopoly level "to induce the natural monopolist to expand its output to a socially preferred level — where buyers do not inefficiently substitute consumption of socially more costly goods for consumption of the monopolized good."⁹⁹ The goal of regulation is then not necessarily to achieve marginal cost pricing, but at least to curb the extremes of monopoly pricing. Curbing those extremes increases social welfare because, as Ayres and Klemperer note, the "last bit of monopoly [pricing] is so disproportionately damaging" — it increases the monopolist's profits very little while inflating deadweight loss the most.¹⁰⁰

A weakness of this justification is that it assumes the whole of the monopoly rents are not necessary to pay the regulated firm's total costs. If regulating a monopolist will lead not to a decline in price of the natural monopoly product, but to the unavailability of the product, then the regulation plainly does not increase social welfare. This effect has a prominent position in intellectual property theory, which has long noted that curtailing rights to intellectual property in the hopes of diminishing deadweight loss is a mirage if the decrease in rights results in the

⁹⁷ See Ronald Coase, *The Marginal Cost Controversy*, 13 *Economica* 169 (1946).

⁹⁸ See John F. Duffy, *The Marginal Cost Controversy in Intellectual Property*, 71 *U. Chi. L. Rev.* 37 (2004).

⁹⁹ Breyer, *Regulation and Its Reform*, cited in note ____, at 16.

¹⁰⁰ *Id.* at 990.

property not being produced in the first place.¹⁰¹ Indeed, the whole field of intellectual property is based on the insight that it is not necessarily a good thing to have government regulation that moves the price of a good (information) closer to marginal cost. The government regulation necessary to achieve that goal — no formal intellectual property rights, so that creators would be left with whatever advantages they can obtain from being the first to have the information — would be quite easy to obtain, but society still chooses the more expensive option of enforcing intellectual property rights.

3. *Dissipation of Social Rents and the Discredited (?) Excessive Competition Theory.* Minimizing deadweight loss can be pursued as a policy only to the extent that the natural monopolist's total revenue exceeds the total cost of producing the product. It is that difference that provides the most compelling reason for curtailing property rights. The difference between the monopolist's total revenues and costs is a rent; it is by definition an above-market return on the capital. In a free market, firms will compete to capture that rent. Since no firm has any inherent advantage, the firms could very easily compete away — dissipate — the rent. Indeed, efficient capital market theory suggests that, as a whole, firms would compete away the rent because no firm should be able to earn supra-market rates of return.

This theory of limiting rights to curtail competition is well known in intellectual property theory. In the field of patents, for example, Yoram Barzel has demonstrated that innovation presents a classic common resource or “common pool” problem: Because the right to innovate is a common right (it is not under exclusive control of any one firm), competition among firms will lead to inefficient races-to-invent that can dissipate any social rent associated with an invention.¹⁰² A theory of “excessive competition” has also been a part of regulated industries theory, but in recent years it has been denigrated as “an empty pejorative phrase.”¹⁰³ Of course, at times, the theory of excessive competition was overstated: Fans of government regulation sometimes used the phrase (or its close substitutes “ruinous competition” or “destructive competition”) to posit that such competition would “inevitably lead[] to economic collapse” and

¹⁰¹ See, among the many articles noting this effect, John F. Duffy, *A Minimum Optimal Patent Term* (2002) (available at www.ssrn.com); R. Polk Wagner, Information Wants to Be Free: Intellectual Property and the Mythologies of Control, 103 Colum. L. Rev. 995, 1017-24 (2003) (noting that increasing intellectual property rights can increase the amount of information produced and even perhaps the total amount in the public domain).

¹⁰² Yoram Barzel, *Optimal Timing of Innovations*, 50 REV. ECON. & STAT. 348, 348-349 (1968). For a similar point made in the copyright literature, see Michael Abramowicz, [add cite to the excessive copyright paper].

¹⁰³ Justice Stephen G. Breyer, *Justice Brandeis as Legal Seer*, 42 Brandeis L.J. 71, 715 (2004) (describing “the consensus of modern regulatory economists, who think that ‘destructive competition’ was generally an empty pejorative phrase used by established firms in regulated industries like trucking, maritime shipping, or airlines, to stop the competition that new entrants might provide”).

calamity in the industry.¹⁰⁴ That view has “long since been discarded by most economists.”¹⁰⁵ Excessive competition will not lead to the collapse of the industry; entrants can still expect to receive normal returns on their investments. But there will be too much investment.

The solution to this problem is the same for both intellectual property law and traditional natural monopolies: Rights are curtailed so that the monopoly is worth less. The monopolist receives a partial property right; consumers or society in general can then capture the rent that otherwise would have been destroyed. Ideally, rights should be curtailed just to the point where the returns on investment are just sufficient to attract capital to efficient projects. In traditional natural monopoly, this curtailment is accomplished through price regulation that tries directly to achieve the correct return on all socially desirable investments (i.e., “prudent” investments, to use the regulatory lingo) in the monopoly. In intellectual property, the curtailment is accomplished indirectly through limits on the duration and scope of property rights. But in both fields, the goal is to preserve against competition social rents.

IV. Conclusion

Intellectual property has been called many things. To its champions, it has been just “property.” To its skeptics, it is “monopoly” or a “public good.” Yet none of these labels has been helpful in generating a comprehensive and useful framework for analyzing intellectual property. They are at best general categories which cover not only intellectual property but also many other things. At worst, they are political slogans designed to bias discussion of the field in one particular way.

This paper has tried to compare intellectual property to its closest cousin from the physical world, natural monopoly regulation. The resemblances are strong, but even the differences are interesting — and even enlightening. The comparison between the two should yield new insights into both fields. Indeed, this paper suggests that the two fields reveal an overarching basis for regulation: Partial property rights are conferred precisely because more complete rights would lead to excessive competition and destruction of the social rent implicit in the monopoly. That basis has been recognized in both fields, though previously “discarded” in one. It is hoped that this insight is but the first that closer analysis of the two fields will yield.

¹⁰⁴ Daniel R. Fischel, Andrew M. Rosenfield, and Robert S. Stillman, *the Regulation of Banks and Bank Holding Companies*, 73 Va. L. Rev. 301, 303 (1987).

¹⁰⁵ *Id.*