Intellectual Property

Peter Menell & Suzanne Scotchmer*
University of California at Berkeley

ABSTRACT

This chapter provides a comprehensive survey of the burgeoning literature on the law and economics of intellectual property. It is organized around the two principal objectives of intellectual property law: promoting innovation and aesthetic creativity (focusing on patent and copyright protection) and protecting integrity of the commercial marketplace (trademark protection and unfair competition law). Each section sets forth the economic problem, the principal models and analytical frameworks, application of economic analysis to particular structural and doctrinal issues, interactions with other legal regimes (such as competition policy), international dimensions, and comparative analysis of intellectual property protection and other means of addressing the economic problem (such as public funding and prizes in the case of patent and copyright law and direct consumer protection statutes and public enforcement in the case of trademarks).

The digital revolution and other technological breakthroughs of the past several decades have brought intellectual property to the forefront of economic, social, and political interest. Much of the value of the leading companies in the world today resides in their portfolio of intangible assets – ranging from the better defined forms of intellectual property (such as patents and copyrights) to the least tangible of the intangibles (trade secrets (know-how) and trademarks (good will associated with a brand)). By one estimate, approximately two-thirds of the value of major industrial companies derives from intangible assets (Swiss Reinsure Company 2000). Not surprisingly, there has been a deluge of economic analyses of intellectual property law during the past decade (Menell 1998a; Landes and Posner 2003; Jaffe and Lerner 2004; Gallini and Scotchmer 2002; Merges, Menell and Lemley 2003; Scotchmer 2004b).

At the outset, it is important to clarify two important issues relating to "intellectual property." Although it draws upon certain characteristics from the law relating to real and personal "property" – most notably, the concept of exclusive rights – and many parallels can be readily identified, the differences between tangible forms of "property" and "intellectual property" are profound and numerous. To take one prominent example, whereas the traditional bundle of rights associated with real and personal property involve perpetual ownership (the classic "fee simple absolute" of real property law), two of the most prominent forms of intellectual property -- patents and copyrights -- protect rights for limited durations (although in the case of copyrights, the term is quite long). Furthermore, exclusivity in the field of

* We would like to thank Mark Lemley and the editors for comments.

"intellectual property" is far less inviolate than it is in the traditional property domains. Intellectual property law comprises a system of policy levers that legislatures tailor and courts interpret in order to promote innovation and protect the integrity of markets.

Second, the field of "intellectual property" is far from unified or monolithic. The landscape of intellectual property comprises a highly variegated array of quite distinct legal regimes: patent, copyright, trade secret, trademark, and a variety of specialized modes of protection (e.g., mask work protection). Although multiple intellectual property regimes can protect different aspects of the same work – computer software being a prime example – it is important to recognize that each mode of intellectual property protection has distinct characteristics and limitations.

For purposes of exploring the economic dimensions of the intellectual property field, it is important to distinguish between two quite distinct functions. The principal objective of intellectual property law is the promotion of new and improved works – whether technological or expressive. This purpose encompasses patent, copyright, and trade secret law, as well as several more narrow protection systems (e.g., mask works, database, design, and misappropriation). The other purpose of intellectual property law addresses a very different economic problem – ensuring the integrity of the marketplace. Trademark law and related bodies of unfair competition law respond to this concern.

I. Promoting Innovation

Economic interest in intellectual property grows primarily out of the critical importance of innovation to social welfare. Solow (1957) demonstrated that technological advancement and increased human capital of the labor force accounted for most (between 80 and 90%) of the annual productivity increase in the U.S. economy between 1909 and 1949, with increases in the capital/labor ratio accounting for the remainder. Denison (1985) extended and refined this analysis, reaching similar results for the period 1929 - 1982: 68% of productivity gain due to advances in scientific and technological knowledge, 34% due to improved worker education, 22% due to greater realization of scale economies, and 13% attributable to increased capital intensity; these factors were offset by decreases in work hours (-25%), government regulation (-4%), and other influences. It is now widely recognized that technological advancement and enhanced human capital are the principal engines of economic growth in the United States and other industrialized countries (Scherer and Ross 1990, pp. 613-14).

The role of intellectual property in contributing to innovation, however, has been more difficult to establish. As we will see, the availability of intellectual property for innovation creates incentives for investment as well as potential impediments to diffusion and cumulative innovation. The net effects are quite complex to sort out from both theoretical and empirical perspectives. As a means for surveying and synthesizing the field, we begin in Section A by clarifying the economic problem that motivates interest in intellectual property protection. Section B provides an overview of the principal modes of intellectual property protection aimed at promoting innovation and creativity. We then survey the design of intellectual property systems and discuss the principal policy levers available for tailoring such protection, focusing first

upon stand-alone innovation before turning to cumulative innovation. Part D examines administration of intellectual property regimes. We then turn to enforcement of intellectual property, its interaction with competition policy, and applied research on its role in particular industries. Part H returns to the question with which we begin – the comparative efficacy of intellectual property in promoting innovation. The final section discusses the economics of intellectual property treaties.

I.A. The Economic Problem

The principal justification for intellectual property derives from a broader economic problem: the inability of a competitive market to support an efficient level of innovation. In a competitive economy, profits will be driven to zero, not accounting for sunk costs such as research and development (R&D) or costs of authorship. From an *ex post* point of view, this is a good outcome, as it keeps prices low for consumers and avoids deadweight loss. But from an *ex ante* point of view, it produces a sub-optimal level of investment in R&D. Most firms would not invest in developing new technologies, and potential creators might not spend their time on creative works, if rivals could enter the market and dissipate the profit.

Unlike tangible goods, knowledge and creative works are public goods in the sense that their use is nonrival (Arrow 1962; Nelson 1959). One agent's use does not limit another agent's use. Indeed, in its natural state (cartooned in the digital age as "bits want to be free"), knowledge is also "nonexcludable." That is, even if someone claims to own the knowledge, it is difficult to exclude others from using it. Intellectual property law is an attempt to solve that problem by legal means; it grants exclusive use of the protected knowledge or creative work to the creator. For other forms of property, exclusion is often accomplished by physical means, such as building a fence. Intellectual property is a legal device by which the inventor can control entry and exclude users from intangible assets.

Of course, intellectual property results in deadweight loss to consumers, and that is its main defect. Two other defects are that it may inhibit the use of scientific or technological knowledge for further research, and, from an *ex ante* point of view, that there is no guarantee that the research effort will be delegated efficiently to the most efficient firms, or even to the right number of firms. Commentators have been lamenting the defects of intellectual property since the nineteenth century, in more or less the same terms as today (Machlup and Penrose 1950).

But intellectual property also has virtues, of which we mention three powerful ones. Probably the greatest virtue is that every invention funded with intellectual property creates a Pareto improvement. No one is taxed more than his willingness to pay for any unit he buys; else he would not buy it. In contrast, funding out of general revenue runs the risk of imposing greater burdens on individual taxpayers than the benefits they receive.

A second great virtue is decentralization. Probably the most important obstacle to effective public procurement is in finding the ideas for invention that are widely distributed among firms and inventors. The lure of intellectual property protection does that automatically. Decentralization is especially important if

private inventors are more likely than public sponsors to think of good ideas for innovations.

The third virtue is that intellectual property is an effective screening device. Cf. Long (2002) (emphasizing the role of patents as a signaling device). Since the private value of the invention generally reflects the social value, inventors should be willing to bear higher costs for inventions of higher value. The intellectual property mechanism encourages inventors to weed out their bad ideas.

But these are not determinative, since other incentive mechanisms may share the same virtues while at the same time reducing deadweight loss. Whereas the earlier economics literature proceeded as if intellectual property protection was the self-evident solution to the incentive problem, a more recent literature, beginning with Wright (1983) and discussed below, has tried to understand when that is true, and when other incentive mechanisms might dominate.

This shift in emphasis has led to another realization: The choice among incentive mechanisms, and even the optimal design of intellectual property laws, depends importantly on the nature of the creative process or, in economists' jargon, on the model of knowledge creation. We mention some of these up front, as our later discussion of the optimal design of intellectual property
will refer to them.

Four principal models of technological change have been proposed in the economics literature: the evolutionary model, the model of induced technical change, a production function for knowledge, and an exogenous process of idea formation, with incentives determining investments.

In the evolutionary model proposed by Nelson and Winter (1982) (see also Mokyr (1990, Ch. 11)), technology evolves in an evolutionary process in which R&D investments occur whenever profit drops below a specified level. Hence, the evolutionary model is not set up to investigate incentives at all, since investment is automatic. In the model of induced technical change proposed by Hicks, technical change occurs in response to changes in factor prices: "A change in the relative prices of the factors of production is itself a spur to invention and inventions of a particular kind – directed at economizing the use of a factor which has become relatively expensive" (Hicks 1932, pp. 124-25; see also Ruttan 2001). Thus rising energy prices can be expected to spur technological advances in energy conservation (Newell et al. 1999). In the production-function model of discovery, which is the basis of almost all the literature that studies patent races, there is an exogenously given relationship which determines, as a function of research inputs or the number of researchers, either the quality of invention (de Laat 1996; Shavell and van Ypserle 2001; Che and Gale 2004) or the likelihood of success in each time period (Loury 1979; Lee and Wilde 1980; Reinganum 1982, 1985, 1989; Wright 1983; Denicolo 1996, 1997; and many others). In both the induced-technical-change model and the production-function model, the profit opportunities are common knowledge. Decentralization is not important. In contrast, the "ideas" model of O'Donoghue, Scotchmer and Thisse (1998) (see also Scotchmer (1999) and Maurer and Scotchmer (2004b)) focuses directly on the scarcity of ideas. The basis of research is "imagination," and to achieve an innovation, a researcher must both have the idea for the innovation and an incentive to invest in it.

Although it is the most widely used model, the production-function model does not lead naturally to intellectual property as superior to other incentive schemes. For example, the advantages of decentralization are more important in a model where "ideas are scarce" than where "ideas are common knowledge." A recurrent theme below is that the optimal design of incentives depends on the model of creation that one has in mind.

I.B. An Overview of the Principal IP Regimes Promoting Innovation and Creativity

Patent law affords a strong and broad form of protection for technological works so as to encourage inventors to disclose their inventions and allow them to fend off imitators that seek to copy all essential elements of an invention before the inventor can recoup the costs of invention and compensate for the risk of investment. The patent offsets this power and further encourages cumulative innovation by allowing follow-on inventors to secure rights on improvements and to enable any competitor to build upon the innovation in its entirety within a comparatively short period of time (20 years from the time of the application). Copyright law, by contrast, wholly excludes protection for ideas and functional attributes of a work but protects creators against direct or near exact copying of even a significant fragment of the whole for a longer duration of time (life of the author plus 70 years).

Trade secret law can also be seen as a means of promoting innovation, although it accomplishes this objective in a very different manner than patent. Notwithstanding the advantages of obtaining a patent – an exclusive right to practice an invention for a designated period of time – many innovators prefer to protect their innovation through secrecy. They may feel that the cost and delay of seeking a patent are too great or that they can more effectively profit from their investment through secrecy. They might also believe that the invention can best be exploited over a longer period of time than a patent would allow (Horstmann, MacDonald and Slivinski 1985). Without any special legal protection for trade secrets, however, the secretive inventor runs the risk that an employee (or a thief) will disclose the invention. Once the idea is released, it will be “free as the air” under the background norms of a free market economy. Such a predicament would lead any inventor seeking to rely upon secrecy to spend an inordinate amount of resources building high and impervious fences around their research facilities and greatly limiting the number of people with access to the proprietary information. Under trade secret law, an inventor need merely take *reasonable* steps to maintain secrecy in order to obtain strong remedies against individuals within the laboratory or commercial enterprise and those subject to contractual limitations who misappropriate proprietary information. Although trade secret law does not limit the use of ideas once they have become publicly known, it does significantly reduce the costs of protecting secrets within the confines of the research and commercial environment.

Table I provides a concise comparative summary of patent, copyright, and trade secret law – the principal modes of intellectual property protection fostering innovation.

Chart I
COMPARATIVE OVERVIEW OF THE
PRINCIPAL MODES OF INTELLECTUAL PROPERTY PROTECTION
FOR PROMOTING INNOVATION AND EXPRESSIVE CREATIVITY

	UTILITY PATENT*	COPYRIGHT	TRADE SECRET
Underlying Theory	limited monopoly to encourage production of utilitarian works in exchange for disclosure and ultimate enrichment of the public domain	limited monopoly to encourage the authorship of expressive works in exchange for disclosure and ultimate enrichment of the public domain	freedom of contract; protection against unfair means of competition
Source of Law	Patent Act (federal)	Copyright Act (federal)	state statute (e.g., Uniform Trade Secrets Act); common law
Subject Matter	processes, machines, manufactures, or compositions of matter	literary (including software), musical, choreographic, dramatic and artistic works	formula, pattern, compilation, program, device, method, technique, process
Limitations	excludes laws of nature, natural substances, business methods, printed matter (forms), mental steps	*limited by* idea/expression dichotomy (no protection for ideas, systems, methods, procedures) and useful article doctrine (only expressive elements of useful articles may be protected); no protection for facts/research	
Reqts for Protection	utility; novelty; non-obviousness; utility; adequate written description	originality (low threshold); authorship; fixation in a tangible medium	information not generally known or available; reasonable efforts to maintain secrecy; commercial value

	UTILITY PATENT*	COPYRIGHT	TRADE SECRET
Process for Obtaining Protection	Examination by the Patent Office. Limited Opposition process. Reexamination process. Maintenance fees.	now automatic, but optional registration process at the Copyright Office confers some benefits	none
Rights	exclusive rights to make, use, sell innovation as limited by contribution to art	rights of performance, display, reproduction, derivative works	protection against misappropriation -- acquisition by improper means or unauthorized disclosure
Scope of Protection	extends to literal infringement (embodiments of all materials elements of a claim) and "equivalents"(non-literal but close imitations; subject to prosecution history estoppel (no protection for reasonably foreseeable "equivalents" if claim narrowed during prosecution))	extends to substantial similarity to protected expression (even a small but significant part of an overall work)	proprietary information
Duration	20 yrs from filing (utility); extensions up to 5 yrs for drugs, medical devices and additives; 14 years (design)	life of author + 70 years; "works for hire": minimum of 95 years after publication or 120 yrs after creation	until becomes public knowledge
Disclosure	right to patent lost if inventor delays too long after disclosing before filing application; full disclosure is required as part of application; notice of patent required for damages	prior to 1978, publication without proper copyright notice (©) resulted in forfeiture; copyright notice no longer required	loss of protection (unless *sub rosa*)

	UTILITY PATENT*	COPYRIGHT	TRADE SECRET
Rights of Others	narrow experimental use exception; limited prior user right (for business methods); reverse doctrine of equivalents potentially allows radical improvements to be practiced without license of embodied patented technology	fair use; compulsory licensing for musical compositions, cable TV, et al.; independent creation	independent discovery; reverse engineering; if someone else obtains patent, then trade secret owner may be enjoined from continued use
Costs of Protection	filing, issue, and maintenance fees; litigation costs	none (protection attaches at fixation); suit requires registration; litigation costs	security expenses; personnel dissatisfaction; litigation costs
Ownership, Licensing, and Assignment	Inventors may assign inventions. Licensing encouraged by completeness of property rights, subject to antitrust constraints; non-exclusive (and possibly exclusive) licenses cannot be assigned without licensor consent.	"work made for hire" doctrine – works prepared within scope of employment and commissioned works (from designated categories and evidence by written agreement) are owned by employer *ab initio* termination of transfers – assignor has inalienable termination right between 36th and 41st years (if notice given) assignment – non-exclusive licenses (and exclusive licenses in the 9th Circuit) cannot be assigned without licensor consent	Licenses must ensure that trade secret is not disclosed; trade secret licenses cannot be assigned without authorization from licensor

	UTILITY PATENT*	COPYRIGHT	TRADE SECRET
Remedies	injunctive relief and damages (potentially treble if willful infringement); attorney fees (in exceptional cases)	injunction against further infringement; destruction of infringing articles; damages (actual or profits); statutory damages ($200-$150,000) w/i court's discretion); attorney fees (w/i court's discretion); criminal prosecution	civil suit for misappropriation; conversion, unjust enrichment, breach of contract; damages (potentially treble) and injunctive relief; criminal prosecution for theft
	* The Plant Patent Act and the Plant Varieties Protection Act separately provide exclusive rights for distinct asexually reproducing plant varieties and sexually reproduced varieties respectively; Design patent protection (affording 14 years of protection for non-functional ornamental designs) largely overlaps copyright protection.		

I.C. Policy Levers

From an economic perspective, the modes of intellectual property protection as well as the system as whole can be seen as an interrelated set of policy levers. The ones that have been stressed in economics journals are length and breadth, and increasingly, the threshold for protection. More recent literature, especially in law journals, has examined a wider range of rules and institutions affecting the incentive effects of intellectual property regimes.

The policy levers operate differently in different creative environments. They also operate differently in the contexts of stand-alone innovations and innovations that lay a foundation for future innovations – referred to as "cumulative innovation." In order to distinguish the economic effects, we begin with models of the stand-alone environment and then move onto the more important and complex domain of cumulative innovation.

1.C. 1. Stand-Alone Innovation

Much of the early economic modeling of the role of intellectual property in promoting innovation posed the following question: what system of incentives or rewards would best promote the attainment of a particular invention. Such models provide the basis for analyzing legal protection for a distinct and relatively narrow class of inventions which do not ultimately generate follow-on innovation. Examples from this class include the safety razor, the ballpoint pen, and pharmaceutical innovations for which the scientific mechanism is poorly understood (Nelson and Winter 1982; von Hippel 1988, p.53). Even where inventors depend on prior knowledge, which is almost inevitable, the lag may be such that prior rights have expired, so that the incentive system treats inventions as stand-alone. Examples are the bicycle and the early development of the light bulb (Dyson 2001).

Models focusing on stand-alone innovation can also be helpful in analyzing legal protection for expressive creativity. Although such works often draw upon prior art for inspiration or common reference points for the work's audience, most authors, musicians, and artists have not traditionally built so extensively upon the work of prior creators as to require express permission.[1] This proposition obviously turns on the underlying right structures – copyrights tending to be relatively narrow in comparison to patents – but it also reflects a fundamental difference between the fields of technological and expressive innovation: "Science and technology are centripetal, conducing toward a single optimal result. One water pump can be better than another water pump, and the role of patent and trade secret law is to direct investment toward such improvements. Literature and the arts are centrifugal, aiming at a wide variety of audiences with different tastes. We cannot say that one novel treating the theme, say, of man's continuing struggle with nature is in any ultimate sense 'better' than another novel--or musical composition or painting--on the same subject. The aim of copyright is to direct investment toward abundant rather than efficient expression" (Goldstein

[1] The ease of incorporating prior musical, graphic, and audiovisual works into "new" digital works has fostered greater cumulative creativity in the content fields (Lessig 2004).

1986).

The critical inquiry in seeking to promote stand-alone innovation is how much profit an inventor or creator should receive and how it should be structured. The focus for stand-alone innovation, therefore, is upon ex ante incentives. As we will emphasize below (also emphasized by Gallini and Scotchmer (2002)), all of the results in this area depend sensitively on what is assumed about licensing. Collaboration and the exchange of technological knowledge across firm boundaries encounter substantial transaction costs. Arora et al. (2001) find evidence that changes in the technology of technical change – most notably the growing use of digital information technologies – facilitate greater partitioning of innovation tasks activities across traditional firm boundaries. They foresee markets for technology – licensing and specialized technology transfer and innovation service firms – playing a more significant role in the production of innovation. When we turn to cumulative innovation, ex post incentives enter the analysis. The principal categories of policy levers affecting incentives to invent are the threshold for protection, duration, breadth, rights of others (and defenses), remedies, and channeling doctrines (for determining priority where intellectual property regimes overlap).

I.C.1.a. Threshold for Protection. As noted above, intellectual property protection results in deadweight loss to consumers. Therefore, it should only be available for significant innovation – works that are new and would not be readily forthcoming without legal encouragement. Works already in the public domain should not be protectable and the threshold for protection should be sufficiently high (or the rights sufficiently narrow) to prevent easily achieved ("obvious") advances from being insulated from free market competition.

Intellectual property regimes erect several types of threshold doctrines limiting protection: (i) subject matter rules – categorical limitations on protection; (ii) substantive requirements – minimum criteria for protection; and (iii) formal requirements – administrative and technical rules that must be complied with in order to obtain and maintain protection. Patent law applies broadly to all classes of innovation (i.e., few subject matter limitations), but applies relatively rigorous standards (utility, novelty, non-obviousness (or inventive step), and adequate disclosure) through a formal examination system. By contrast, copyright applies a very low threshold for protection – a work need only be fixed in a tangible medium of expression and reflect a modicum of originality – and does not require examination (registration is optional). As we will see, such a low threshold is counterbalanced by a relatively narrow scope of protection. Trade secret law requires merely that information derive economic value from not being generally known or readily ascertainable (by proper means) by others and be the subject of efforts that are reasonable under the circumstances to maintain its secrecy.

Patent law's threshold requirements have received the most economic scrutiny. Due to the relatively uniform nature of patent protection, some have argued that certain classes of innovation (such as computer software and business methods) that may not require such lengthy protection should be subject to a sui generis form of protection (Menell 1987 (software); Samuelson et al. 1994) or excluded from intellectual property protection altogether (Thomas 1999 (business methods); Dreyfuss 2000). The basic

contours of patent law were established during an age of mechanical innovation and were designed with this model (and the guild system that predominated) in mind (Merges, Menell, and Lemley 2003, pp. 105-11). Mechanical innovation continued to comprise the bulk of patent applications well into the 20th century. During the past half century, however, various newer fields – such as chemistry, software (and business methods), and biotechnology -- have increasingly come into the patent system (Allison and Lemley 2002), calling into question the premises on which patent law was built. If specialized protection systems are not developed to address new and distinctive fields of innovation (as was partially done in the case of semiconductor chip designs -- see: Semiconductor Chip Protection Act of 1984), the challenge remains of reshaping the relatively uniform patent system to accommodate the growing heterogeneity of inventive activity (Cf. Burk and Lemley 2002, 2003).

Patent law's novelty requirement – what it means to be "first" – turns on the location of the "finish line" in the race to invent. Most patent systems in the world apply a first-to-file standard; the United States determines the winner on the basis of who was the first to invent. In principle, the first-to-invent system rewards the first inventor to discover new knowledge, even if they lack the specialized patent filing resources of others. Thus, many small inventors defend the first to invent system as a means of leveling the playing field relative to large companies which may have more resources available and personnel in place to file applications more expeditiously. The first-to-file system significantly reduces the administrative costs of operating a patent system – priority depends solely on the time and date stamped on an application. Evidentiary disputes over the subtle nuances of who was first to grasp an invention can be quite costly to resolve (Macedo 1990). Empirical studies cast doubt on the notion that small inventors tend to do better under a first to invent system, likely reflecting the high costs of resolving priority disputes (Mossinghoff 2002; Lemley and Chien 2003).

The first-to-invent system also has incentive effects as to the choice between trade secrecy and disclosure (Scotchmer and Green 1990). Inventors may be inclined to delay their applications in order to effectively extend the expiration date of a patent (20 years from the date of filing). In order to counteract this effect and promote prompt filing, U.S. patent law adds an additional layer of legal complexity (and hence uncertainty and cost): requiring that an inventor file an application within one year after the invention is disclosed (either through patenting or publication in the anywhere in the world or in public use or on sale in the United States). This reduces the delay in disclosure of new knowledge, but does not eliminate it. The first-to-file system promotes earlier disclosure of technological advances. Grushcow (2004) finds that the growing interest in patenting by academic institutions since 1980 has delayed the publication of research, potentially increasing the risk of wasteful duplication of research.

From an economic standpoint, patent law's non-obviousness standard plays the most important role in determining which innovations qualify for protection (and hence what type of innovation patent law encourages). Patent law specifies that a claimed invention must go beyond readily predictable or conventional solutions to technical, engineering, or business problems. Articulating an objective and determinative standard for non-obviousness, however, has proven elusive. In the 1940s, U.S. courts interpreted the law to require a "flash of creative genius" test. (Cuno Engineering Corp. v. Automatic

Devices Corp., 314 U.S. 84, 91 (1941)). Such a demanding formulation generated a backlash within the patent community, leading Congress to frame the standard in the following manner: a patent may not be obtained "if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to at the time the invention was made to a person having ordinary skill in the part to which said subject matter pertains." (17 U.S.C. §103) What raises the non-obviousness hurdle above the novelty standard is that the patent examiner may consider multiple references simultaneously where there is a suggestion, teaching, or motivation to combine elements across these references. The examiner must also consider circumstantial evidence of non-obviousness (so-called "secondary considerations") – long-felt but unsolved need, commercial success of the claimed invention, failed efforts by others, copyright by others, praise for the invention, unexpected results, and disbelief of experts – but only to the extent that such factors are connected to the inventive aspects of the claim.

In its actual formulation and application, the non-obviousness rule falls short of implementing the economic gatekeeping principle. Whereas an economist would consider paramount among relevant considerations the level of research and development expense in making an invention (Merges 1992), the U.S. Patent Act states that patentability "shall not be negatived by the manner in which the invention was made," implying that inventions requiring minimal effort (and hence likely to obtain even without protection) may nonetheless qualify for protection. In addition, research expense and effort are not listed among the traditional secondary considerations, although several court decisions on non-obviousness take note of such factors (Merges 1992 (noting that the threshold for patentability should be lowered with regard to high-cost research); Oddi 1989, p. 1127 (recommending that courts expressly consider "qualitative and quantitative investment in research and development" among the secondary factors)). The legal standard for non-obviousness does consider the level of uncertainty involved in research. The fact that a research project is "obvious to try" does not render a resulting discovery "obvious" unless there was little or no ex ante uncertainty about the outcome – i.e., those skilled in the art could readily predict the outcome of the experiment. In practice, the test depends on the number of parameters and the extent to which the relevant prior art guides the experimentation process.

The role of commercial success in the non-obviousness determination has produced conflicting economic analyses and prescriptions. Drawing upon historical and empirical research on the innovation process, Merges (1988) finds commercial success to be a poor proxy for technical advance. What succeeds in the market tends to reflect product strategy and marketing more than technical advances over the prior art. Hence, Merges (1988) argues for downplaying this factor and scrutinizing the connection between market success and the technical advance. By contrast, Kitch (1977, p.283) sees market success as consistent with the prospect theory. By using subsequent economic success as a factor favoring patentability, the patent law increases "the security of the investment process necessary to maximize the value of the patent." Both analyses support the idea that the consideration of market success in assessing non-obviousness promotes commercialization (Merges and Duffy 2002, pp. 727-36), although it is not clear that a patent system is needed to achieve this end. Where adequate incentives exist to invent, free market forces should be adequate to promote commercialization. (But cf. Kieff (2001b) (articulating a commercialization theory of patent law)).

Several observers of the patent system perceive that the Federal Circuit has significantly lowered the non-obviousness hurdle over the past two decades (Desmond 1993; Barton 2003; Lunney 2000-01). Based on both statistical and doctrinal analysis, Lunney (2004) concludes that the Federal Circuit has effectively "eviscerated" the non-obviousness requirement, routinely affirming decisions upholding patents and frequently overturning decisions by holding patents invalid on non-obviousness grounds. See also Dunner (1995). Since 1982, the year that the Federal Circuit was created, the rate at which patents have been held invalid has plummeted (Lunney 2004). Allison and Lemley (1998) find that non-obviousness remains the most frequent ground for invalidating patents at the trial and appellate levels (42% of invalidity judgments), but often fails when raised (63.7%). As we discuss in the section on administration of intellectual property (I.D), however, empirical studies have not discerned a significant decline in patent quality.

Some commentators believe that it is now far too easy to obtain a patent, particularly with regard to business method and DNA sequence patent (Barton 2003; Hall 2003). They recommend raising the non-obviousness hurdle, although articulating a standard that appropriately balances the concerns of over- and underprotection has proven elusive. In fact, Scotchmer (2005b, chapter 3) argues that the low bar to patentability is a misplaced worry, and shifts the discussion from the standard for patentability to the breadth of the right. Unnecessary patents are not harmful to competition if the patents are narrow, so that similar products compete in the market.

The non-obviousness standard may have some perverse collateral effects on the nature and timing of disclosure of new knowledge. By preemptively publishing work in progress, a firm that is ahead may induce a shake-out among rivals by raising the level of prior art to render a rival's subsequent invention obvious (Scotchmer and Green 1990); a laggard in a patent race may be able to reduce the likelihood that a leader will be able to obtain a patent by raising the level of the prior art sufficiently to defeat patentability by a leader (Lichtman, Baker, and Kraus 2000; Parchomovsky 2000; cf. Bar-Gill and Parchomovsky 2003). This possibility might also lead competitors to collude or collaborate to maximize patent opportunities. Under this theoretical account, a higher standard of non-obviousness increases the viability of a preemptive patenting strategy. The likelihood that such a strategy would be pursued by rivals has been questioned on doctrinal and practical grounds (Merges 2004b, p.195-96); Eisenberg 2000; cf. Hicks 1995).

Much of the economics literature on trade secrets addresses the optimal level of expenditures to maintain secrecy, i.e., what constitutes "reasonable efforts" under the circumstances. Kitch (1980) argues that all such "fencing costs" are inefficient and would require only such expenses as are necessary to provide evidence of the existence of a trade secret, i.e., a notice or marking function. Friedman, Landes, and Posner (1991) make the related point that trade secret protection should be available when it is cheaper than the physical precautions that would be necessary to protect a particular piece of information.

I.C.1.b. Duration. Nordhaus (1969) offered the first formal model of the optimal duration of intellectual property protection. Nordhaus asked why the life of the intellectual property right should be

limited, since a longer right leads to more innovation, and more innovation creates social benefit. He argued that there is a counterveiling cost. The longer right might increase innovation, but it also increases deadweight loss on all the inframarginal innovations that would occur even with shorter protection – i.e., innovations that would be forthcoming even in the absence of the longer right. The optimal duration of a patent or copyright should balance the incentive effect against the deadweight loss in order to maximize social welfare. Many economists believe that copyright duration (life of the author plus 70 years) is much longer than justified to provide an appropriate ex ante incentive for creation of new works. See Akerlof et al. (2003); but cf. Landes and Posner (2003, p. 218) (noting that the deadweight loss from copyright protection is relatively small due to the narrow scope of copyright protection).

To see the Nordhaus argument, suppose that there is a universe of "ideas" available for investment. Let an idea be a pair (s,c) where s measures the value of the resulting innovation and c is its cost. An idea with higher s can be interpreted as leading to a larger market; a higher s means that the demand curve is shifted out. Let $?(s,T)$ be the profit available to a rightholder with an intellectual property right of length T and an idea of quality s. $?$ is increasing in both T and s. Let $W(s,T)$ be the corresponding social welfare associated with investment in the idea. The welfare $W(s,T)$ is the sum of consumers' surplus for the infinite life of the innovation, sold at the competitive price, minus the deadweight loss during the period of protection. Thus, W is increasing in s and decreasing in T. Finally, suppose that for each R&D cost c, the distribution of "ideas" is given by a distribution function F with density f, where $F(s|c)$ is the fraction of ideas with cost c that have value less than s.

Then the social value of investment in ideas with cost c is $W^{\wedge}(T,c)$ defined below, where $s(T)$ is the minimum value that will elicit investment ($?(s(T),T) = c$). That is,

$$\hat{W}(T,c) = \int_{s(T)}^{\infty} [W(s,T) - c] f(s|c) ds$$

Notice that $\sigma\mathbb{N}(T) < 0$. A marginal increase in T will increase investment in amount $-W(\sigma(T),T)\sigma\mathbb{N}(T)$. However, even though investment goes up with T, total social welfare $W^{\wedge}(T,c)$ may go down. The change in social welfare is

$$\frac{\P}{\P T}\hat{W}(T,c) = \int_{s(T)}^{\infty} \frac{\P}{\P T} W(s,T) f(s|c) ds - W(s(T),T) s'(T)$$

The last term represents the welfare due to new innovations called forth by longer protection, but the first term, which is negative, represents the loss in consumers' surplus on all the inframarginal innovations that would have been achieved even with shorter duration. As T becomes large and $\sigma(T)$ becomes small, it is reasonable to think that the first term becomes large relative to the last term. Increasing the duration T beyond that point will not be in the social interest.

Of course the best length T must be established by adding up the marginal effects for all c. Depending on the distributions of (s,c) in different product classes, the one-size-fits-all nature of the patent

system may provide excessive protection in some product classes, and deficient incentives in others.

Races for the intellectual property right introduce another inquiry as to how profitable the intellectual property right should be, regardless of how the profit is achieved.[2] Unlike the Nordhaus argument, the inquiry leads to an argument for limited duration that applies even if the profit is given as a prize out of general revenue and involves no deadweight loss. The argument concerns the optimal amount of R&D effort. A more profitable right will encourage more entry into the race (the extensive margin) or more collective effort as each participant accelerates its effort (the intensive margin).

The potential benefits of inciting more effort by offering more profit depend on the creative environment – the nature of the R&D process. Nordhaus implicitly addressed a creative environment where "ideas are scarce" so that duplication of costs is not the focus. Suppose, however, that more than one potential innovator can serve the same market niche. Then there is a second reason to limit duration. Not only will there be excessive deadweight loss on inframarginal innovations, but the disparity between profit and cost will also lead to duplication of R&D cost as firms vie for the very profitable rights.

Thus, part of the inquiry into the optimal strength (profitability) of an intellectual property right concerns the extent to which additional effort is duplicative. This issue takes us back to the question, What is the right model of the creative environment? If "ideas are scarce," then races are not an issue. But if all investment opportunities are commonly known, then races may or may not be efficient, depending on the "production function for knowledge." If successes and failures in the R&D process are perfectly correlated, then a race is duplicative. If successes and failures are independent, then a race increases the probability of at least one success, or in another interpretation, accelerates progress (Loury 1979; Lee and Wilde 1980; Reinganum 1982, 1985, 1989). Further, if the creative environment is one in which different firms have different unobservable ideas for how to address a given need, then entrants to a race need not be the most efficient firms or those with the best ideas (Scotchmer 2005b, chapter 2).

The number of entrants in a race may be too large or too small, as compared to the efficient number, depending on the size of the private reward. Suppose, for example, that two firms have different ideas about how to fill a market niche with value s. Suppose that each firm's cost is c, and that each has probability 1/2 of succeeding. Suppose that the value of the property right will be $\Pi(s,T)$, and that the firms' prospects for success are independent. If both firms succeed, each will receive the property right with probability 1/2.

---

[2] Innovation races are more suited to patents and patentable subject matter than to copyrights and creative works. Such races can only occur if several rivals are vying for a right that only one of them will receive. Rights to creative works are generally narrow enough in scope that several authors can obtain protection for works that have some similarity (and hence can compete). Thus, an author may fear a reduced market due to competition from another author, but does not generally fear that he or she will be wholly excluded from the market through a rival completing their work first.

Then a second firm will enter the patent race if $\Pi(s,T)(3/8)>c$, since its probability of receiving the patent is 3/8. On the other hand, entry by the second firm is only efficient if $W(s,T)(3/4)-2c> W(s,T)/2-c$ or $W(T,s)/4>c$. Thus, if $\Pi(s,T)(3/8)>\underline{c}>W(s,T)/4$, there will be excess entry to the patent race – the second firm will enter even though that is not efficient – and if $\Pi(s,T)(3/8)<c<W(s,T)/4$ there will be too little entry.

Entry into a race may provide a private value to the entrant that is greater than the social value of the entry, and always provides a private value that is greater than the increment to private value of both firms. The latter is because of the "business-stealing effect." The second entrant's expected profit is $\Pi(s,T)(3/8)-c$, while the increment to joint profit is only $\Pi(s,T)(1/4)-c$. The second entrant's chance of winning the race and getting the patent comes partly at the first entrant's expense. It is this externality that may lead to excessive entry into a race. It also implies that, if the reward were as large as the social value, there would be too much entry. In fact, the only thing that is clear in this environment, without imposing additional structure, is that the optimal reward is smaller than the social value of the innovation. But this is not a very useful design principle, because rewards given as intellectual property will have that attribute almost inevitably.

Landes and Posner (2003, pp. 222-28) suggest that some works may be diminished by a congestion externality. They illustrate their point by reference to the Disney Corporation's self-imposed restraint on commercialization: "To avoid overkill, Disney manages its character portfolio with care. It has hundreds of characters on its books, many of them just waiting to be called out of retirement . . . . Disney practices good husbandry of its characters and extends the life of its brands by not overexposing them . . . They avoid debasing the currency" (Britt 1990). Landes and Posner (2003) assert that this concern justifies perpetual protection for some works. To balance the costs of protection, they advocate a system of indefinitely renewable copyright protection, with the renewal fee acting as policy lever for diverting works not subject to congestion externalities into the public domain. They note that a similar over-saturation can arise with regard to some rights of publicity (use of persona in advertising) and trademarks.

I.C.1.c. Breadth. The breadth or scope of an intellectual property right has critical bearing on its economic value, and hence its incentive effect. A broader right preempts more substitutes than a narrow right.

The scope of a patent is determined by the language of the claims (which define the boundaries of literal infringement) and the extent to which such boundaries will be stretched to cover similar, but not quite literal, embodiments. Under the "doctrine of equivalents," courts will find infringement where the accused device "'performs substantially the same function in substantially the same way to obtain the same result.'" Graver Tank & Mfg. Co. v. Linde Air Products Co., 339 U.S. 605, 608 (1950) (quoting Sanitary Refrigerator Co. v. Winters, 280 U.S. 30, 42 (1929)). *See also* Warner-Jenkinson Co. v. Hilton Davis Chemical Co., 520 U.S. 17 (1997). The scope of copyright is determined by the substantial similarity test in conjunction with copyright's limiting doctrines (e.g., originality, scenes a faire, non-protectability of ideas and facts, fair use) – does the defendant's work embody *substantial similarity* to *protected elements*

of the plaintiff's work? In practice, a copyright is quite narrow with regard to newly created works. It is unlikely that different authors operating from the same ideas will produce substantially similar novels. Breadth issues arise, however, with regard to works built on copyrighted works, such as sequels and film adaptations. We deal with these issues in the context of cumulative innovation. Breadth does not generally arise in the context of trade secrets.

These legal tests do not map directly onto the economic concepts of breadth that economists have developed. Economic models of breadth have been developed for two market contexts: where an innovation is threatened by horizontal competition, and where an innovation might be supplanted by an improved innovation. We take up the latter question in the analysis of cumulative innovation (section I.C.2). For horizontal competition, breadth has been modeled in two ways: in "product space," defining how "similar" a product must be to infringe a patent, and in "technology space," defining how costly it is to find a noninfringing substitute for the protected market.

In the first notion of breadth, introduced by Klemperer (1990) using a spatial model, the size of the market for the patented product depends on the closeness of noninfringing substitutes. A broader patent covers more of the product space, meaning that more substitutes infringe. The right to keep a substitute out of the market is profitable for the patent holder in two ways: by shifting the demand for the protected good outward (where the intellectual property owner excludes the substitute from the market) or by allowing the intellectual property owner to charge higher prices for both the patented good and the infringing substitute.

The second notion of breadth for horizontal substitutes, due to Gallini (1992), is that it determines the cost of entering the market. In this conception, the goods are exact substitutes, and breadth implicitly refers to the technology of production (process innovation) rather than closeness of substitutes in the market. Entry by a second firm does not cause demand curves to shift, but instead causes the firms to compete in a given market. A narrower patent (lower cost of entry) will lead to more entry and lower prices. Entry stops when the cost of entry can no longer be covered by competing in the market.

In both conceptions of breadth, a narrower patent leads to lower per-period profit. Thus, breadth might be conceived as a policy lever that governs profit, as described above, in a one-size-fits-all system where the duration of protection cannot be tailored to the cost of innovation. In the patent system, such tailoring is not generally done in any systematic way by the Patent Office. Examiners focus solely on ensuring that the application meets the threshold criteria and that the claims are clear. They do not adjust the "breadth" of claims. The courts exercise a modest degree of tailoring. In applying the doctrine of equivalents, courts accord "pioneering" inventions greater scope than more modest inventions. Such a rule increases the reward for major breakthroughs. The copyright system does not systematically vary the scope of protection with the cost or importance of the work.

Even within the one-size-fits-all system, there is a policy question as to whether, on average, rights designed to give a pre-specified reward should be structured as long and narrow or short and broad. The

inquiry into how market rewards should be structured has led in several papers to a ratio test: a policy reform is desirable if it increases the ratio of profit to deadweight loss. The ratio test was devised by Kaplow (1984) in the patent/antitrust context, and also used by Ayres and Klemperer (1999) in the enforcement context. It reappears in the cited discussions of patent breadth. The basic notion is that deadweight loss is the consumer cost of raising money through proprietary pricing. If the ratio of profit to deadweight loss is higher, the money being raised through proprietary pricing is raised more efficiently.

In a broad class of demand curves including linear ones, any price reduction from the monopoly price will increase the profit-to-deadweight-loss ratio, but will also reduce profit, thus necessitating a compensation such as longer protection. This can be seen in Figure 1, where the monopoly price is ½ and the lower price 1/3 is the duopoly price. At the price 1/3, the ratio of profit to deadweight loss is the ratio of the cross-lined areas (of size 2 ×C) to the triangle D. At the monopoly price 1/2, the ratio of profit to deadweight loss is the ratio of the outlined box that represents monopoly profit to the triangle (B+M+D). One can see by inspection that the ratio of profit to deadweight loss is smaller at the monopoly price ½ than at the lower price 1/3. In fact, with the linear demand curve, this argument generalizes for any reduction in price: the lower the price, the higher the ratio of profit to deadweight loss. This is the argument given by Tandon (1982), arguing for compulsory licenses to lower prices, and Gilbert and Shapiro (1990), arguing for narrow patents, which they interpret as lower prices, although they do not say how price reductions in a given market might flow from narrower scope.

How, though, do narrow patents lower the price in a given market? In Gallini's conception, breadth determines the cost that an imitator must pay to enter a proprietary market. Entry is only tempting if the market will be protected long enough so that the entrant, in competition with the patent holder, can still cover the cost of entry. If entry occurs, competition between the entrant the rightholder will lower the price.

Figure 1

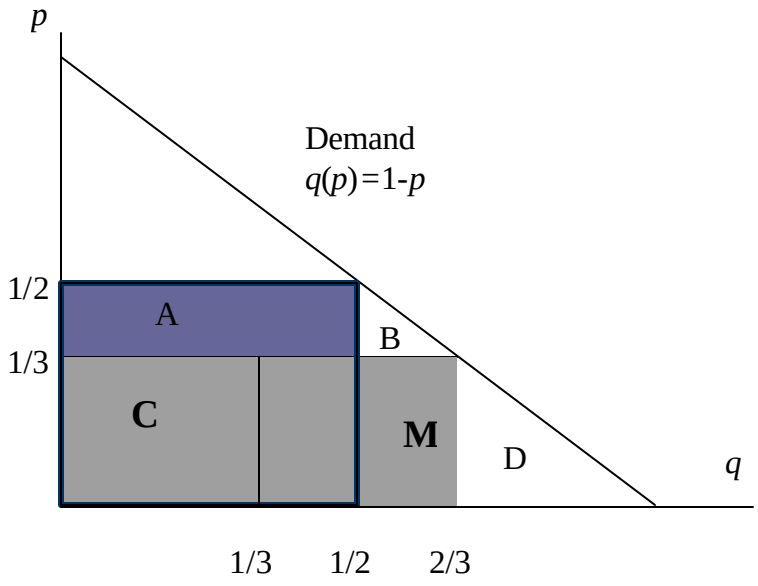

Figure 1 can be used to compare a relatively short period of protection where entry by an imitator is not tempting, with a longer period of protection, where entry is tempting even though the imitator must pay a cost. With the shorter period of protection, say T^M, consumers will pay the monopoly price ½, but with the longer period of protection, say $T^D > T^M$, they will pay the duopoly price 1/3. Suppose that T^M and T^D are chosen so that the patent holder makes the same discounted profit in both regimes, and the cost of entry is such that exactly one imitator will enter if the patent lasts for length T^D. Then by the above argument, consumers would be better off in the duopoly regime, despite the longer period of protection, because of the lower price.

However that argument does not account for the fact that the imitator must pay real resource costs to enter the market. Gallini argues that the duplication of costs is severe enough to overturn the above argument. Given that the price can only be reduced by costly entry, it is better for society as a whole – including consumers, the patent holder and the imitator – to have a short period of monopoly pricing than a longer period that attracts entry.

However, we have already stressed that the best design of intellectual property rights depends importantly on what one assumes about licensing. In this case, licensing again overturns the conclusion. Maurer and Scotchmer (2002) argue that the patent holder will anticipate entry, and offer a license instead of tolerating unlicensed entry. In this way the patent holder can increase his own profit without reducing the profit of the entrant, and at the same time can eliminate the wasteful duplication. The narrow patent thus has the effect of lowering price without imposing the social cost of duplication, and the above welfare analysis is restored. The better policy is a narrow patent for a relatively long time.

I.C1.d. Rights of Others (and Defenses). The rights afforded others in protected works directly affect the profits from intellectual property. Many of these rules – such as blocking rights (patent and copyright) and exceptions for experimental use (patent), fair use (copyright), and reverse engineering (copyright and trade secret) – find their economic justification in the cumulativeness of innovation, and therefore we take them up in section I.C.2. Doctrines relating to independent invention, prior user rights, and "first sale" (or exhaustion of rights) relate to stand-alone invention, as do proposals about extending user rights to limited sharing of copyrighted works.

Rights Arising from Independent Invention. A right of independent invention means that, provided the independent inventor was actually an "inventor" (and, in particular, did not learn the invention from any other party, such as a prior inventor), he or she is free to practice the invention. Both copyright law and trade secret law immunize independent inventors from liability, but patent law does not. In the case of trade secrets, it would be impossible for an independent inventor to know what had previously been invented. In the case of copyrights, which protect expression, any re-expression escapes liability (broadly speaking). In the case of patent law, the right is defined with respect to claims, and (broadly speaking) not with respect to how a potential infringer achieved the potentially infringing innovation. These principles have

doctrinal nuances, some of which are mentioned below.

Scholars have made three types of economic arguments about independent invention. First, in the context of trade secrecy, the absence of an independent-invention defense would stifle innovation because inventors would be uncertain as to whether they could practice the new knowledge they create.

Second, a right of independent invention can reduce the duplication of R&D costs in patent races (La Manna et al. 1989; Blair and Cotter 2002; Maurer and Scotchmer 2002; Leibovitz 2002; Ottoz and Cugno 2004). If the value of an exclusive right in the market is $100, and the R&D cost is $20, five firms may enter a race. But if all five firms have rights ex post, competition will reduce the private value of the right below $100, and fewer than five firms will enter. The right of independent invention reduces the duplication of costs, and at the same time affords lower prices to users, all without undermining the incentive to invent.

Landes & Posner (2003, pp. 361-62) make a similar argument for trade secrets. They compare the American rule, under which the owner of a trade secret loses his right to the invention if someone else patents it (W.L. Gore & Associates v. Garlock, Inc., 721 F.2d 1540, 1550 (Fed. Cir. 1983)), to the prior-user-right that prevails in some other nations. The prior-user-right divides the entitlement, enabling multiple independent inventors to share its value through an effective oligopoly structure. As in the foregoing argument, duplicative entry will only occur to the extent that all firms cover their costs.

Third, giving rights to independent inventors can induce patent holders to license ex post on terms that reduce market price, in order to discourage ex post entry through independent invention (Maurer and Scotchmer 2002). Suppose that a single patent holder is in the market. Then, whether or not the patent holder licenses, a right of independent invention will reduce the price in the market below the proprietary price. Without licensing, the price will fall due to entry by independent inventors. Instead, the patent holder can license at a fee equal to the cost of independent invention. Then independent inventors are indifferent to paying the license fee or paying the costs of independent invention, but the patent holder prefers licensing. The price reduction in the market (determined by the terms of license and number of licensees) must be large enough to deter further entry.

The market price with licensing will thus depend on the cost of independent invention. If the cost is high enough, the right of independent invention can benefit users without undermining the incentive to invent. In fact, in plausible models, the cost of independent invention only needs to be greater than half the cost of the original innovator (Maurer and Scotchmer 2002; Ottoz and Cugno, 2004). Nevertheless, Blair and Cotter (2002) rightly point out that the economic consequences depend critically on the relative costs of first inventors and imitators, which will differ across technologies. Giving a right of independent invention can have harmful consequences if imitation or independent invention is too cheap.

Lichtman (1997) made a similar argument in the context of unpatented inventions, advocating on grounds of cost that independent inventors be allowed to copy but not clone them. Armond (2003)

proposed that independent discovery be available as a defense to a preliminary injunction motion.

Although independent inventors are not generally exempted from liability under U.S. patent law, the law does, in fact, recognize user rights in two circumstances: (1) prior secret use of business methods – as a narrow statutory exception with regard business method patents (17 U.S.C. §273); and (2) shop rights – under state law governing employment agreements and the employment relationship, an employer obtains a royalty-free, non-exclusive, non-transferable license to use an employee's invention where the employee makes a patented invention using the employer's facilities. In most research environments today, employers require employees involved in research-related activities to assign their inventions to the employer, although some state laws limit such agreements to inventions developed within the scope of employment or developed using the employer's facilities. (E.g., CAL. LAB. CODE § 2870) Even where no express agreement has been signed by an employee, patents invented by the employee may nonetheless be deemed to have been assigned where an employee has specifically been employed to invent in the field in which the invention was made. In these circumstances, a court may imply an assignment clause into the employment contract.

Rights after Sale. Under what is commonly referred to as the "first sale" or "exhaustion" doctrine, the intellectual property owner "exhausts" the legal monopoly in a product by selling it to the public, thereby enabling the purchaser to use the work and resell it without infringing. Such a default right structure reduces transaction costs for subsequent transactions. Similarly, purchasers of patented products are deemed to have an implied license to make repairs, although this license does not extend to "reconstruction" of the patented product. Intellectual property owners can, subject to anti-competitive restrictions, circumvent the first sale doctrine by imposing licensing restrictions upon the conveyance of a product.

Rights to Share Copyrighted Works. Even though the purpose of copyright law is to prevent copying, a controversial idea that keeps resurfacing is that copying or sharing is less harmful to creators than meets the eye, at least where the sharing of each legitimate copy is limited. Where it is unlimited, such as in peer-to-peer networks or when users make copies of copies, sharing poses a greater threat to appropriability.

The argument is that the proprietor will price in a way that anticipates sharing. Sharing allows the proprietor to charge a higher price, since demand is determined by the willingness to pay several parties. Limitations on sharing may arise because copies of copies degrade (Liebowitz 1982, 1985; Liebowitz and Margolis 1982 (arguing in the era of analog copies)) or because it is less costly to facilitate sharing than to produce a copy for every user, as in a video rental market (Varian 2000), or because the probability of detection increases with the size of the sharing group.

The earlier set of papers in this vein relied on the fact that copying is costly. Novos and Waldman (1984) and Johnson (1985) argued that proprietors may reduce price to avoid copying, but that the cost of copying will nevertheless preserve the proprietor's market. The market price will be lower than without copying, reducing the deadweight loss of excluding users, but the per-period reward to creative works will

also be reduced, especially when there is heterogeneity in tastes as well as in copying costs. The welfare effects are different according to whether the cost of copying is per-copy or per-user, as when it requires the purchase of a copying device. Scholars have also argued that copying can have an affirmative benefit for rightholders because it builds network effects (Conner and Rumelt 1991; Shy and Thisse, 1999).

A second set of papers focus on the fact that prices can be tailored to the groups that form. Liebowitz (1985) emphasized price discrimination according to whether the purchaser will make the copy available to many users, as libraries do. See also Ordover and Willig (1978). Bakos, Brynjolfsson, and Lichtman (1999) argued that, depending on the groups, sharing might actually be more profitable than selling to individual users. This is true if, first, the willingnesses to pay within groups are negatively correlated, or, second, if there is variance in the sizes of groups. Thus, whether sharing enhances profit depends on what governs group formation. However, Scotchmer (2005a) argued that sharing groups will not be formed exogenously or even randomly, and if they form in a way that is efficient for the group members conditional on the proprietors' prices, then group formation has no effect at all on profit opportunities. Sharing is neither profit-reducing nor profit-enhancing.

Given that copying can have salutary effects as well as deleterious effects, these arguments have led authors to consider an additional set of policy levers specific to the copying context, such as taxes and subsidies on prices of legitimate copies, and taxes and subsidies on copying devices, as well as the optimal mix of enforcement activities and other incentives. See Besen and Kirby (1989); Chen and P'ng (2003); Netanel (2003); Fisher (2004).

I.C.1.e. Remedies. As in other bodies of law, the remedial opportunities in intellectual property law are injunctions and damages. There are two branches of thought about the relative efficacy of these rules, one branch focusing on whether remedies will lead to efficient use of the property ex post, and the other branch focusing on the ex ante effects.

The first set of arguments (Calabresi and Melamed 1972; Polinsky 1980; Kaplow and Shavell 1996) for the general framework, and Blair and Cotter (1998) for the intellectual property context considered whether property rules (injunctions) are more or less likely than judicially imposed liability to encourage bargaining to an efficient outcome ex post. For example, property rules (injunctions) may be preferred when transaction costs of exchange are low and the costs of valuing violations of rights by courts are high. For the intellectual property context, Ayres and Klemperer (1999) add the consideration that "soft" remedies, which do not actually restore the proprietary price, can be socially beneficial because they increase consumers' surplus without impinging much on profit, at least for small price reductions.

The second set of arguments are not concerned with what would happen in the out-of-equilibrium event of infringement, but focus instead on how potential remedies affect equilibrium profits and the ex ante incentives for R&D (Schankerman and Scotchmer 2001; Anton and Yao 2004). In these arguments, remedies are only important because they do or do not deter infringement, and because they determine the terms of an ex ante license. The terms of license that will be accepted by a potential licensee/infringer

depend on the consequences for infringement, and this threat has an affect on the ex ante division of profit. Schankerman and Scotchmer argue that if infringement leads to profit-eroding competition between the infringer and rightholder, a wide range of remedies will deter infringement, at least for stand-alone innovations, and are therefore equivalent from an ex ante point of view. However, this is not necessarily true for research tools and other potentially licensed intellectual property where infringement does not dissipate profit (see below).

Merges and Duffy (2002) and Blair and Cotter (1998) argue, again from the ex post perspective, that patent and copyright law are better suited to a property-rule paradigm than a liability-rule paradigm. Since intellectual property rights are relatively well defined, disputants or potential disputants should have little trouble resolving their differences by negotiating licenses against the backdrop of an injunction. In contrast, if the setting of damages (an ex post "compulsory license") is left to a generalist judicial institution under a liability rule, the court may have difficulty placing a value on the intellectual property or on the injuries caused by infringement. Further, judicially imposed licenses can undermine the prospect function of patent law (Kitch 1977). Merges (1996) argues that for complex transactions involving many players, a property rule will facilitate the creation of private exchange institutions, such as patent pools, that can evolve in response to changing circumstances and draw upon industry and institutional expertise.

Although infringed rightholders generally have a right to enjoin unauthorized use, injunctions are backed up by compensatory damages for past violations. These may include enhanced (punitive) damages for patent infringement, statutory damages for copyright infringement, and attorney fees and costs in "exceptional cases." In several areas such as the covers of musical compositions, juke boxes, cable television broadcasts, and webcasting, copyright law provides for compulsory licensing. These regimes arguably economize on transaction costs, although commentators are divided on the economic effects of such compulsory licenses. Cf. Merges (1996) and Netanel (2003)

Consistent with traditional economic analysis of damages (harm internalization), patent and copyright law award intellectual property owners the greater of lost profits or a reasonable royalty for the defendant's unauthorized use of the protected works (Blair and Cotter 1998). Calculating these measures, however, is quite complex, involving numerous subtle determinations of how markets would have evolved had infringement not taken place (Blair and Cotter 2001).

It follows from general economic principles that enhanced damages should be awarded where improper behavior is costly to detect and where full compensatory damages are costly to prove (Polinsky and Shavell 1997; Cooter 1982; Cotter 2004). Excessive damages (i.e., where expected damages exceed actual damages) could lead to overdeterrence in the sense that parties may exercise caution in order to avoid a risk of liability. Several recent studies indicate that courts may well be overdeterring patent infringement based on the high rate of enhanced damage awards (Federal Trade Commission 2003; Cotter 2004; Moore 2000; (finding that requests for enhanced damages were fully litigated in nearly half (45%) of cases and awarded in 64% of cases between 1983 and 1999, thereby resulting in awards in 29% of all cases)).

Due to the very different nature of trade secret protection, the remedies available for unauthorized use and dissemination of a trade secret are more limited. Where the secret has not been disclosed to the public, courts will generally enjoin further use of the secret by a misappropriating entity. But where the secret has been disclosed, the trade secret owner will be limited to damage remedies or limited injunctive relief against the entity that misappropriated the trade secret (such as a "head start" injunction which excludes the misappropriator from the market for a designated period). Disclosure to the public destroys the secret and therefore it would be inappropriate (and infeasible) to enjoin use of the information by others. Nonetheless, Sidak (2004) argues an injunction against a misappopriating entity should be perpetual in order to encourage efficient post-litigation bargaining over the value of continued use. Such a rule would also avoid the expense and difficulty (error costs) of having courts adjudicate the option value of a trade secret.

I.C.1.f. Channeling Doctrines. The various modes of intellectual property provide overlapping protection. For example, patent, copyright, and trade secret law all cover computer software. As noted earlier, however, copyright doctrines exclude ideas, processes, and methods of operation from copyright protection. Thus, software developers cannot gain copyright's long duration of protection for the functional aspects of computer software. Such inventions must comply with patent law's formal examination requirements and surpass patent law's higher thresholds in order to obtain legal protection. In this way, intellectual property law prevents inventors from obtaining protection for functional features through the "backdoor" of the copyright system.

The relationship between patent and trade secret law is somewhat more complicated. Both regimes cover technological innovation. Where an inventor obtains a patent before a subsequent researcher invents the same technology, the patent trumps the subsequent inventor, regardless of whether the subsequent inventor seeks to protect the invention as a trade secret. (Such secrecy may well conceal infringement, particularly in the case of process inventions, but that does not suggest that the trade secret would have any validity vis á vis the patent).

A somewhat more complicated issue arises where the first inventor chooses to protect the technology as a trade secret. If a subsequent inventor independently discovers the same invention and obtains a patent, two issues overlap issues arise: (1) does the trade secret invalidate the patent (on novelty grounds); and (2) if the patent is valid, can the trade secret owner continue to practice the invention – in essence, does the first inventor enjoy a prior user right. As suggested earlier, the trade secret will not invalidate the patent because it does not fall within the body of prior art that may be considered in judging novelty. Therefore, assuming the second inventor meets the other requirements of patentability, she will obtain a valid patent. As regards the rights of a trade secret owner, U.S. patent law holds that the trade secret owner infringes upon the patent by continuing to practice the invention. Some nations recognize a prior user right in this circumstance, which places the technology under duopoly rather than monopoly control. The profits available to the patentee are reduced accordingly. It can be argued, however, that such a structure of rights might partially improve the screening function of patent law – inventions that have been independently developed may not have needed as much of an ex ante incentive in the first place. To

the extent that ex ante incentives are more than sufficient to generate the innovation, duopoly improves social welfare by reducing the deadweight loss from exclusive exploitation.

I.C.2. Cumulative Innovation

In the context of stand-alone inventions or creations, intellectual property rewards reflect the social value of the contribution, since the profit is determined by demand. That is one of the main virtues of intellectual property as an incentive system. However, when innovation is cumulative, the most important social benefit of an innovation may be the boost given to later innovators, and this may make the benefits harder to appropriate (Scotchmer 1991). Moreover, the innovation may enable rivals to enter with improved products. In that case, social success may mean private failure: the boost given to the rivals may cause the innovation's own demise (Scotchmer 1991, 1996; Green and Scotchmer 1995; Chang 1995; O'Donoghue, Scotchmer and Thisse 1998; O'Donoghue 1998; Besen and Maskin 2002; Hunt 1999, 2004). Merges and Nelson (1990) give a rather opposite perspective on the cumulative problem. Instead of worrying that later improvers pose a threat to earlier innovators, they worry that earlier innovators (earlier patents) pose a threat to later improvers.

The intellectual property system must resolve these contradictions. In general, the problem of appropriating benefits has two facets: the overall level of profit, and how it is divided among the sequential innovators. As we will see, the roles of the policy levers are closely intertwined in the cumulative context, and the best design of the system will depend on the transaction costs of licensing.

Many scholars have emphasized the importance of cumulativeness in the process of knowledge creation, especially economic historians. As expressed by David (1985, p. 20), "Technologies . . . undergo . . . a gradual, evolutionary development which is intimately bound up with the course of their diffusion." Secondary inventions – including essential design improvements, refinements, and adaptations to a variety of uses – are often as crucial to the generation of social benefits as the initial discovery. See, e.g., Nelson and Winter (1982); Taylor and Silbertson (1973); Mak and Walton (1972); Rosenberg (1972). Cumulative technologies tend to involve multiple components, serve as building blocks for further incremental innovation, and often spur wide-ranging applications. Automobiles, aircraft, electric light systems, semiconductors, and computers fall within this category. Some chemical technologies are hybrids of discrete and cumulative models. New chemical compounds are typically discrete in terms of the product market that they serve, but can suggest promising new lines of research (e.g., penicillin, Teflon) (Nelson and Winter 1982). The biotechnology field reflects several cumulative features. The development of research tools provides the means for decoding genomic information. Research decoding genomes provides the input for downstream biomedical research.

Cumulativeness also extends to expressive creativity. All authors and artists draw, to some extent, on prior works. Sequels, translations, and screenplays build directly upon prior works. Parodies and satires comment on or employ other works. Most musical compositions reflect rhythm and other elements of established genres. The hip-hop and rap musical genres embody prior recordings through the use of

digital sampling.

Computer software, which is written work intended to serve utilitarian purposes, falls between the technological and expressive realms. Cumulativeness plays a particularly important role here, whether in operating systems, technical interfaces, peripheral devices or application programs (Menell 1987, 1989; Lemley and O'Brien 1997).

I.C.2.a. A Preliminary Model: The virtues of licensing

One of the lessons that emerges powerfully below is that the best design of the intellectual property system depends on how fluid the market for licenses is. Before turning to a more detailed analysis of design issues and how licensing affects them, we illustrate the importance of licensing by modifying the "reduced form" model of Landes and Posner (2003), where a variable z is taken as the "strength" of a right. For example, the strength of the right may be affected by breadth or exemptions such as fair use.

Let $q(\cdot)$ be the demand function for a protected innovation, where $q(p)$ is decreasing with p. Referring back to our discussion of copying, and how the threat of copying affects the market price, let $y(p,z)$ be the supply of illicit copies. Then the net demand faced by the proprietor is $q(p)- y(p,z)$. Assuming for convenience that the marginal cost of copies is zero, the proprietor maximizes $p[q(p)- y(p,z)]$, and sets a profit-maximizing price $p^*(z)$ that depends on the strength of protection through the threat of copying.

Now consider how the profit-maximizing price and the proprietor's profit depend on the strength of protection, z. Assume that the supply $y(p,z)$ of illicit copies increases with p and decreases with z. Because the supply of imitations $y(p^*(z),z)$ depends on the proprietor's price as well as the level of protection, there is a potential indeterminacy in the model. An increase in protection could conceivably lead to a decrease in price and an increase in illicit copying. However, under reasonable assumptions we can assume that the profit-maximizing price increases with the level of protection, even though the increase in price has a feedback effect of increasing imitation or copying.

Nevertheless, the profitability of creations and hence their supply will not necessarily increase with the level of protection z. This is because the cost of creation can also depend on z, e.g., by creating a burden to innovate outside the scope of other rights. Suppose, in fact, that each potential innovator faces an R&D cost k plus additional costs $e(z)$ that reflect the burdens imposed by other intellectual property rights. The creation is profitable if

$$p^*(z)\,[q(\,p^*(z))- y(\,p^*(z)\,,z)]-k-e(z) > 0.$$

If potential creations differ in their markets (for example, if we introduce a quality variable s into the demand function q), then the more valuable ideas will call forth investment, while the less valuable ones will not. Without formalizing this idea, we will let $N(z|k)$ describe the number of profitable creations with cost k.

The supply of new creations $N(z|k)$ is not monotonic in the strength of protection z because raising z increases the creator's costs. The punch line of this model is that too much protection can be bad for creators as well as for imitators.

The point we would like to make, however, is that the punch line is largely reversed if firms can license to avoid conflicting property rights, rather than being forced into the costly activity of avoiding them. Following the perspective in O'Donoghue, Scotchmer and Thisse (1998), assume that each innovating firm will initially be in the position of paying license fees on the discoveries of its predecessors, and then in the position of collecting license fees from its followers. The effect of strong rights, z, is to increase the licensing obligations, rather than to increase the real resource cost of avoiding prior rights. The essential point is that licensing also creates claims against future innovators.

Suppose, in fact, that all innovators are in symmetric positions: they pay for the same number of licenses, and are paid by the same number of licensees. Then, since all the money must go somewhere in the end, symmetry means each innovator pays as much in licensing fees as it earns in licensing fees, say, $R(z)$ on both sides of the ledger. The profit-maximizing decision is then to invest in the potential creation if

$$p^*(z)\,[q(\,p^*(z))-\,y(\,p^*(z)\,,z)]+R(z)-k-R(z) > 0$$

or equivalently,

$$p^*(z)\,[q(\,p^*(z))-\,y(\,p^*(z),z)]\,-k> 0$$

Hence (assuming that the revenue $p^*(z)\,[q(\,p^*(z))-\,y(\,p^*(z),z)]$ is increasing in z), the ambiguous effect of strong protection on innovation disappears. A strengthening of protection leads to more creations.

Thus, with licensing, we are cast back to the same consideration as with stand-alone inventions, namely that there is a tradeoff between deadweight loss and innovation, but no tension between protecting early innovators and protecting the later innovators who use the knowledge they create. Licensing will largely resolve that tension, to everyone's benefit.

These points about the salutary effects of licensing have mostly been made in models that distinguish between the policy levers of intellectual property, rather than lumping the policy levers into a single variable called the "strength" of the right. We now turn to that more disaggregated discussion.

I.C.2.b. Duration. Although the length of protection has an obvious effect on the overall level of profit, the *statutory* length may be irrelevant. When innovations are under threat of being supplanted by improved innovations, market incumbency only lasts until that happens. O'Donoghue, Scotchmer and Thisse (1998) define a notion of "effective patent life" that focuses on the rate of market turnover, and argue that the effective life of the patent may be determined by the breadth of the right, rather than its statutory length. This is because breadth determines how long it will take before the product is supplanted (see below).

In fact, there is considerable evidence that the "effective" lives of most patents are shorter than their statutory lives. Mansfield (1986) reported, using survey evidence, that in some industries 60% of patents are effectively terminated within four years. The literature on patent renewals (Pakes and Schankerman 1984; Schankerman and Pakes 1986; Pakes 1986; Schankerman 1998; Lanjouw 1998) carries a similar message. For example, Schankerman (1998) reports that half of patents in France are not renewed beyond the tenth year, even though renewal fees are very low.

Besen and Maskin (2002) present a model of sequentialness where this endogeneity of patent life is absent (because the products do not compete in the market) but argue that statutory life should be shorter when innovators learn from previous innovators. Their model has sequentialness in innovation (because innovators learn from each other) but the resulting products are "stand-alone" and live out their statutory lives. They argue that the optimal statutory life should be shorter if innovators learn from each other than if not, because the loss from impeding future innovation is greater. (This result depends sensitively on the absence of licensing; see below for an argument that all results in this arena turn on what is assumed about licensing.)

For the case of basic and applied research, Green and Scotchmer (1995) argue that patent lives must last longer if the research is divided between sequential innovators rather than concentrated in a single firm, because of the problem of dividing profit in a way that respects the costs of both parties.

To the extent that transaction costs may impede licensing and first stage inventors do not need large ex ante rewards to induce innovation, then a shorter duration of intellectual property protection promotes cumulative innovation. Legal protection for computer software fits this profile (Menell 1987). There are relatively strong non-intellectual property incentives for developing operating system and other platform technology for many product markets. Interoperability with widely adopted platforms is often critical to secondary innovation, such as application programs and peripheral devices. Owners of the intellectual property rights in widely adopted proprietary platform technologies can exercise tremendous market power due to network effects and consumer lock-in. Shortening the duration of protection for such technologies is one mechanism for constraining such market power and better equilibrating the incentives of first and second generation innovators.

I.C.2.c. Threshold Requirements and Breadth. The status of an invention that builds on other inventions can be: (1) protected and noninfringing, (2) unprotected and noninfringing, (3) protected and infringing, or (4) unprotected and infringing (Scotchmer 1996; Denicolo 2002a). Thus, the economic effects of threshold (patentability) and breadth are hard to disentangle. Scenario (1) gives the best incentive for second generation innovators, but does not force the second-generation innovator to share the profit with the prior innovator. Scenario (2) will clearly stymie second-generation innovation unless there is a mechanism other than intellectual property to protect the innovator. In scenario (3) – which is possible under patent law – the works are considered "blocking": the later work infringes the prior innovation and cannot be exploited without a license; and the later work is protected and cannot be exploited by the pioneering inventor without license. Such a scenario encourages the inventors to share profit from the

subsequent invention. Scenario (4), which approximates the treatment of derivative works under copyright law, discourages improvements or adaptations by subsequent inventors in the absence of ex ante bargaining. However, there is less difference between scheme (3) and scheme (4) than meets the eye. Even if the later product is not protectable, it can be protected by an exclusive license on the innovation it infringes.

The literature draws widely differing conclusions about the optimal way to organize the rights of sequential innovations, largely because authors make different assumptions about when and whether licenses will be made, and who can be party to the negotiation. Kitch (1977) was the earliest, and perhaps most extreme, licensing optimist.

Licenses can be made either *ex ante*, before the follow-on innovator invests in his project, or *ex post*. If licenses must be negotiated *ex post*, after both innovations have been achieved, scheme (4) may stifle innovation, since the second innovator will fear that the first innovator will simply appropriate it (Green and Scotchmer 1995). On the other hand, if the second innovator can approach the first innovator for an *ex ante* license before investing in his idea, the second innovation is not in jeopardy under either of scheme (3) or (4). However, the first innovator will typically collect more of the profit in scheme (4) because the second innovator will have less bargaining power (Scotchmer 1996).

Denicolo (2002a) considered a model where ideas are common knowledge, and asked how the various scenarios affect patent races, assuming that there will be *ex post* licenses, but no *ex ante* licenses. He finds that the choice should depend on the relative costs of the innovators. If, for example, the cost of the first innovation is low and the cost of the second is relatively high, it may be better not to let the first innovator share in the second innovator's profit. Of course, this also depends on whether the first innovation can earn profit in the market, or only through licensing. Chang (1995) also considered a context where the firms could make *ex post* licenses, but not *ex ante* licenses, and concluded that the choice between (1) and (3) should depend on the relative costs of the innovators.

The worst situation is when licensing may fail entirely, and when, in any case, the earlier innovator does not need to profit from licensing in order to cover his costs. Merges and Nelson (1990) draw on many actual examples to argue that such circumstances are quite plausible. A defect of the one-size-fits-all intellectual property regime is that it cannot distinguish cases where blocking rights are unnecessary for cost recovery from cases where earlier innovators would not invest unless they can profit from licensing. In part on this basis, Burk and Lemley (2002) argue that it would be better to make intellectual property protection more finely attuned to industrial contexts.

But even in those cases where earlier innovators should be allowed to profit from the later innovations they enable, blocking rights are a blunt instrument for dividing profit. Profit shares will not necessarily reflect the cost shares. This is especially true if the licenses will be negotiated *ex post*, after all innovators' costs are sunk (Green and Scotchmer 1995), although Chang (1995) argues that the problem can be mitigated by making the infringement (breadth) responsive to cost. In copyright law, blocking rights

are not available as a tool at all. Follow-on creators may not prepare infringing derivative works without permission of the owner of the copyright in the underlying work. Copyright law therefore has less flexibility than patent law in how it balances the incentives for sequential innovators (Lemley 1997).

In the case of product improvements, there is a question of whether a patent's breadth can extend to slightly better products that have not yet been invented, or only to inferior products. O'Donoghue, Scotchmer and Thisse (1998) define a notion of "leading breadth" that determines when there will be blocking rights in the cumulative context, and also establishes the "effective life" of the patent right. To see why leading breadth is useful as a policy lever, suppose to the contrary that every trivial infringement is noninfringing, even if patentable. A potential improver may discard ideas for small improvements because they lead to price-eroding competition between close, vertical substitutes. It is only the relatively big ideas that will become innovations. This problem can be solved by making the small improvements infringing. Firms may then be willing to invest in them, since control of the improvement and its predecessor can then be consolidated through licensing in the same firm. Instead of competing, both will be marketed together. Further, if the small improvements are infringing, and if it takes a relatively long time for large ideas to come along, the "effective life" of each patent is prolonged. These effects cannot be achieved by choosing the patentability standard alone; the opportunity to consolidate successive improvements in the hands of a single firm arises because the patents are infringing.

In the "ideas" model, there is not much role for a nuanced patentability standard. However, in a "production function" model, a patentability standard can make each successive innovator more ambitious in the size of improvement he invests in. This may be socially beneficial (O'Donoghue 1998).

Hunt (1999, 2004) presents a production-function model motivated by the Semiconductor Chip Protection Act, in which eligibility for protection coincides with noninfringement of previous innovations (as with copyright), He argues that the standard for protection should be increasing in the "dynamicness" of the industry. Since noninfringement coincides with protection, it is hard to sort out their respective roles. Whereas the effective patent life is determined by breadth in the model of O'Donoghue, Scotchmer, and Thisse (1998), it is determined in the Hunt model by both.

Finally, we return to the idea of Kitch (1977), who argued that strong patent rights should be given to pioneers, as he calls them, so that the pioneers can coordinate the subsequent development of the technology. These are called "prospect" patents. The theory is not focused on the reward purpose of the patent, since it would apply even if the pioneer innovation were costless to achieve, and no incentive for R&D were required. The theory thus rejects the line of reasoning that says intellectual property is at best a necessary evil, due to deadweight loss.

The prospecting theory rests on the premise that social interests and the private interests of the patentholder are aligned. Scotchmer (2005b, section 5.6) shows that this may be true in some ways, but is not true in other ways, and in particular, that strong pioneer patents can pre-empt competition policy. As in later theories of cumulativeness, the prospector's profit comes from getting the intellectual property into

use. For this reason the pioneer has an incentive to encourage use at a fee. Further, the pioneer can profit from delegating research effort to the most productive researchers, and avoiding bad projects, as are socially efficient.[3] These are ways in which the pioneer's interest is aligned with the broader public interest.

However, the pioneer can preempt competition policy in two ways: by avoiding competition in the "innovation market" for second-generation products, and by avoiding competition among second-generation innovators once the second-generation innovations exist. The first of these may or may not be socially harmful (see below, the section on the patent/antitrust conflict), but the second is clearly harmful to competition, assuming that patent law is well designed in the first place. If these harms to competition are important, it might be better to avoid pioneer patents even if second-generation innovators must then duplicate the cost of achieving the pioneer innovation. As with all intellectual property, the case for pioneer patents is strongest if the pioneer innovation is costly to achieve, and the patent is actually needed as a reward.

For radical improvements that read on existing patents, Merges (1994) suggests that it may be socially advantageous to exempt an improver from infringement under the "reverse doctrine of equivalents." Under this doctrine, a radical improvement may be deemed non-infringing even though it literally reads upon an existing patent. The doctrine would allow the radical improver to avoid a holdup by the underlying patent holder and to avoid a potential bargaining breakdown. See also Lemley (1997). This is again an argument that relies on difficulties in licensing. Such a doctrine avoids over-rewarding a first-generation inventor (who did not foresee a much greater advance) while providing strong encouragement to visionary subsequent inventors. The rule has rarely been applied in actual cases, although the possibility of its application may well have fostered licensing.

Evolving Doctrines Regarding Subject Matter. We turn now to how these economic arguments have been reflected, if at all, in legal doctrines. Notwithstanding the general applicability of the patent system to "anything under the sun made by man,"(Diamond v. Chakrabarty (1980)), courts have, since the early history of patent system, barred inventors from claiming patents on natural physical phenomena (e.g., the properties of lightening), laws of nature (e.g., the theory of general relativity), mental processes, and abstract intellectual concepts (e.g., algorithms). (Gottschalk v. Benson (1972)) Courts have noted that allowing exclusive rights for such fundamental discoveries would unduly impede future inventors – a cumulative innovation rationale. (O'Reilly v. Morse (1854)). Implicit in this justification is the notion that transaction costs could impede licensing. The courts have thus realized, as is more explicit in the economic models discussed above, that licensing plays an important role in balancing the rights of sequential innovators.

Court decisions over the past 25 years scaling back if not effectively removing the traditional

---

[3] In fact this is always true if the prospector can make deals with consumers as well as follow-on inventors, for example, if consumers can pay the prospector not to retard progress.

jurisprudential limitations on patentable subject matter have led scholars to consider the merits of imposing categorical subject matter exclusions, the development of sui generis protective regimes tailored to particular technological fields, as well as technology-specific rules within the patent system in order to better promote cumulative innovation – particularly in the software, bioscience, and business method fields (Menell 1987). Samuelson et al. (1994), and Cohen and Lemley (2001) have argued that in environments of rapid turnover where costs are relatively low, for example, computer software, strong intellectual property rights can impede the rate of technological advance. As we have seen, this depends on both the design of the rights and the ease of licensing. Heller and Eisenberg (1998) have argued that patenting of gene sequences generates a tragedy of the anticommons, a fragmentation of rights which vastly increases transaction costs, thereby impeding downstream research for medical advances. This is also, at root, an argument about the ease of licensing (see Walsh et al. 2003 (reporting survey data revealing indicating that university research has not been impeded by concerns about patents on research tools as a result of licenses, inventing around patents, infringement (often informally invoking a research exemption), developing and using public tools, and challenging patents in court)).

The opening of the "business method" patent flood gates has raised concerns about whether patent protection is needed at all to promote such innovation and, more troubling, whether such protection is chilling innovation and competition. The idea of protecting business plans runs counter to a core premise of the free market system by offering a form of antitrust immunity for business models. As the rising tide of prior art raises the threshold for protection, such adverse effects may abate and innovation could well produce valuable new business methods. Nonetheless, several scholars find that the costs of extending patent protection to business methods as a class significantly outweigh the benefits (Dratler 2003; Dreyfuss 2000), especially when pro-patent biases and patent quality considerations are factored into the analysis (Meurer 2002; Merges 1999). Drawing upon the approach of foreign patent systems, Thomas (1999) calls for limiting the subject matter of patent law to fields of industrial applicability.

Copyright law applies a different rights structure than patent law, with significant implication for the direction of cumulative innovation (Lemley 1997). Unlike patent law, which allows anyone to patent improvements to patented technology, copyright owners have the exclusive right to prepare derivative works. Therefore, a novelist can prevent others from translating their work into another language, adapting the story for the stage or a motion picture, selling the story in narrated form (e.g., books on tape), and developing sequels that draw extensively on the protectable elements (including possibly character names and attributes – e.g., Rocky IV or the next James Bond film). As we will see below, some borrowing is tolerated under the fair use doctrine, but pioneers generally have exclusive authority to pursue the further development of their expressive work. Copyright law wholly excludes protection for ideas and functional attributes of a work but protects creators against direct or near exact copying of even a significant fragment of the whole for a tremendously long duration (life of the author plus 70 years), reflecting the notion that society prefers to have one hundred different war novels embodying similar themes, ideas, and facts than one hundred versions of "War and Peace" that differ only in their final chapter. Consequently, copyright protection for an author's *expression* of ideas and the relatively long period of its duration effectuates a different balance than patent law. Patent law encourages cumulative innovation by allowing follow-on

inventors to secure rights on improvements and to enable any competitor to build upon the innovation in its entirety within a comparatively short period of time (20 years from the time of the application). By contrast, copyright law, with its narrow scope of protection, allows subsequent creators to pursue competing works using the same ideas as the "pioneer," but allows the pioneer exclusive rights over the development of the expressed ideas (Karjala and Menell 1995).

Adequacy of Disclosure Requirements. Both patent law and copyright law provide for the disclosure and dissemination of knowledge, which promotes cumulative innovation. Patent law requires disclosure as the quid pro quo for patent protection. In the software field, however, disclosure of source code is not required under the best mode requirement – the inventor need only disclose the functions of the software on the grounds that a person of ordinary skill in the art could write software to perform the functions without undue experimentation. (Fonar Corp. v. General Electric, 107 F.3d 1543 (Fed. Cir. 1997)) In practice, however, the knowledge protected by many software patents would be difficult and costly to decipher without access to the source code, which is usually maintained as a trade secret (Burke 1994; Burk and Lemley 2002). This slows the process by which follow-on inventors can build upon earlier generations.

In the case of most copyrighted works, the knowledge contained in the work may be comprehended from direct inspection. Therefore, the publication of a work discloses and disseminates. Furthermore, the Copyright Act requires those who register a work, which is optional, to deposit a copy with the Library of Congress.

Some copyrighted works, however, do not lend themselves to visual inspection and comprehension. Piano rolls, for example are not human readable, although even that technology can be deciphered by the trained ear or a mechanical translation system. Musical or audiovisual works stored on magnetic tape and digital media can also be perceived directly. Computer software, however, cannot typically be perceived unless it is available in source code. Copyright Office regulations, however, do not require disclosure of the entirety of source code in order to secure copyright registration. Thus, as with patent law, copyright law allows protection of software without providing access to the underlying knowledge. As a result, follow-on invention is stifled, although such rules may deter infringement that would otherwise be difficult to detect.

I.C.2.d. Rights of Others (and Defenses). Several doctrines provide safety valves, beyond the limitations embodied in scope of protection, for promoting cumulative innovation. These include the experimental use doctrine (patent law), the fair use doctrine (copyright law), and the reverse engineering doctrines (of trade secret and copyright law). In addition, copyright law provides for several exemptions for educational and related purposes which can be viewed as promoting basic education for new authors and artists.

Experimental Use. A subsequent inventor who wants to improve a patented technology may benefit from experimenting with it. U.S. patent law has had a common law exemption for "philosophical

experiments" and research to ascertain "the sufficiency of [a] machine to produce its described effects." (Whittemore v. Cutter, 29 F. Cas. 1120 (C.C.D. Mass. 1813)[4] Subsequent cases, however, have declared the defense to be "truly narrow" and applicable solely to activities "for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry," (Roche Prods., Inc. v. Bolar Pharm. Co., 733 F.2d 858, 863 (Fed. Cir. 1984); Madey v. Duke Univ., 307 F.3d 1351 (2002), *cert. denied* 123 S. Ct. 2639 (2003)), prompting several scholars to warn that patent law unduly hinders academic and basic research and unduly supplants academic and scientific norms promoting progress, disclosure, and cumulative innovation.

Eisenberg (1989) proposed to exempt research that could potentially lead to improvements or design-arounds of patented technology, but she also pointed out the inherent contradiction that arises when further research is the main use of the patented invention. A broad exemption could entirely undermine the profitability of the patent. However, the effect of a research exemption depends on an ancillary doctrinal question, namely, whether the invention that is achieved by using the prior invention will infringe the prior patent (Scotchmer 2005b, chapter 5). If so, the exemption may (counterintuitively) increase the profit holder's profit. Exercising the research exemption can put the improver in the position of bargaining for a license ex post (after he has sunk his costs) rather than ex ante. This strengthens the bargaining position of the first patentholder.

Kieff (2001a) contends that the exclusive rights provided by patents promote university research by increasing private investment in research and improving the efficiency of academic research environments. His analysis does not, however, directly address whether a more permissive experimental use doctrine would adversely affect the flow of private research investment into universities. Based on survey research, Walsh et al. (2003) find little evidence that patents on research tools have significantly impeded university research. There is one notable exception: the field of genetic diagnostics. Cai (2004) notes that the chilling effect of a narrow experimental use defense may not be very significant due to patent holders' rational forbearance in enforcing against universities as well as legal constraints (sovereign immunity) on suing state actors. Cf. Menell (2000).

Many legal scholars in addition to Eisenberg have proposed alterations to existing law. Mueller (2001) endorsed a broadened experimental use defense along the lines of the European system as well as compulsory licensing. Feit (1989) proposed a compulsory license for patents infringed during experimentation to improve patented technology. These authors, like Eisenberg, raise particular concerns about patents on upstream research tools, particularly in the bioscience field. O'Rourke (2000) proposes a fair-use doctrine for patents that would go beyond the similar doctrine for copyright by allowing courts to judge permissible conduct and impose compulsory royalties. Dreyfuss (2003) proposes to allow experimental use if the investigator's institution promptly waives patents on subsequent discoveries, subject to a "buyout" provision. In cases where a patentee has refused to license to a non-profit on reasonable

[4] Article 27(b) of the European Patent Convention exempts "acts done for experimental purposes relating to the subject matter of the patented invention."

terms, Nelson (2003) proposes a research exemption, provided the researcher agrees to publish his results and agrees either not to patent his own results or to license them on nonexclusive and reasonable terms. Strandburg (2004) proposed to exempt improvement patents and to provide for compulsory licensing of research tools. Other authors (Epstein 2003; Merges 1996) have countered that such proposals would entail significant administrative costs and have complex effects upon licensing markets and the formation of licensing institutions.

Fair Use. The fair use doctrine in copyright law exempts a user from liability for infringement when copyrighted works are used for criticism, comment, news reporting, teaching, scholarship, and research. (17 U.S.C. §107) In applying this doctrine, courts balance the purpose and character of the use (including whether such use is of a commercial nature or is for nonprofit educational purposes), the nature of the copyrighted work, the amount of copying, and, most importantly, the effect of the use upon the potential market for or value of the copyrighted work. Transformative (or more radical "improvements") are more likely to be permissible, whereas uses that merely supplant the underlying work are disfavored. In this way, the fair use doctrine promotes significant creative advances while protecting the pioneer from direct market competition. The fair use doctrine can also be seen as an efficient means for permitting uncompensated use of copyrighted material where the transactions costs of licensing or other means of exchange would prevent a transfer through the market (Gordon 1982).

Courts have applied the fair use doctrine to enable software developers to make copies of protected programs for purposes of learning how such software functions (Samuelson and Scotchmer 2002). With such code deciphered, the rivals could discern unprotectable elements (e.g., interoperability specifications) which they were then free to incorporate in their own commercial products. In this way, the fair use doctrine operates as a form of "experimental use" exemption.

Reverse Engineering. As with copyright law, trade secret law allows others not bound by contractual constraints to reverse engineer technology in order to determine how it functions. To the extent that they decipher trade secrets, they undermine the inventor's advantage. By disclosing the information, they destroy the trade secret. The reverse engineering limitation on trade secret protection thus exposes the trade secret owner to free riding by others. Nonetheless, most commentators believe that it strikes a salutary balance between protection on the one hand and competition and the dissemination of knowledge on the other (Landes and Posner 2003; Samuelson and Scotchmer 2002). The trade secret owner can "purchase" greater protection against this risk by investing in higher levels of security (e.g., more effective encryption for software-encoded technology). The inventor can also pursue patent protection, which proscribes reverse engineering, although only for the limited duration of the patent, and mandates disclosure of the invention to the public. By declining to pursue patent protection (or failing to satisfy the requirements thereof), however, inventors should not be able to secure potentially perpetual rights in technologies merely by encrypting them or otherwise obscuring how they function. To do so would undermine the larger balance of the federal intellectual property system.

Compulsory Licenses. Copyright law uses a statutory compulsory license mechanism for cover

versions of musical compositions, which spurred cumulative innovation in sound recordings. Once a musical composition has been released (with consent of the copyright owner) as a sound recording, any sound recording artist may record and distribute copies of that composition without consent of any copyright owner. This privilege is made possible by limiting the scope of the sound recording copyright to exact duplication[5] and establishing a compulsory license rate for copies made of the underlying musical composition (currently 8.5 cents per copy). (17 U.S.C. §115) The creative freedom associated with this privilege, however, is constrained by the statute – the follow-on recording artist may make "a musical arrangement of the work to the extent necessary to conform it to the style or manner of interpretation of the performance involved, but the arrangement shall not change the basic melody or fundamental character of the work." (17 U.S.C. §115(a)(2)) This privilege also does not extend to the use of prior sound recordings – as, for example, in digital sampling – without the consent of both the musical composition copyright owner and the sound recording copyright owner. Nonetheless, this compulsory license has likely fostered a wider body of interpretations of musical compositions than would have occurred if musical composition copyright owners held exclusive (blocking) rights. Due in part to this privilege, more than 100 cover versions of the popular song "Mack the Knife" are available, with performances ranging from Louis Armstrong to Bobby Darin and instrumental artists.

I.C.2.e. Remedies. As noted above, the main remedies for patent and copyright infringement are injunctions and compensation for past injury, possibly compounded to treble damages in case of willfulness. As these laws are interpreted, courts operate from a baseline of prospective injunctive relief and compensatory damages for past injury. Hence, they do not generally adjust the level of damages as a policy lever, except in the context of enhanced damages, which we discuss below.

With regard to awarding damages for past infringement, the court will often be in the position of having to decide whether, absent the infringement, the rightholder would have licensed. If not, the lost profit may be the value lost to the patentholder because the follow-on product was preempted by the infringer. If licensing would (should) have occurred, the lost profit is lost royalty. These are two rather different inquiries.

Lost royalty is even more speculative in the cumulative context than in the stand-alone context, for a sound theoretical reason (Schankerman and Scotchmer 2001). On one hand, the potential damage award establishes the maximum license fee. On the other hand, the equilibrium license fee establishes the damage award. Hence there is an inherent circularity that leads to multiple equilibria. Because of multiple

---

[5] "The exclusive right of the owner of copyright in a sound recording . . . is limited to the right to prepare a derivative work in which the actual sounds fixed in the sound recording are rearranged, remixed, or otherwise altered in sequence or quality. The exclusive rights of the owner of copyright in a sound recording . . . do not extend to the making or duplication of another sound recording that consists entirely of an independent fixation of other sounds, even though such sounds imitate or simulate those in the copyrighted sound recording . . . " 17 U.S.C. §114(b).

equilibria, the profitability of the patent is unknowable in advance to a researcher investing in it. This problem is especially acute for research tools, and will be less acute for inventions where infringement leads to competition and dissipates total profit. Because of the dissipation, infringement is its own punishment, and infringement is more easily deterred.

The awarding of enhanced damages in patent law (up to treble) and statutory damages in copyright law can be a policy lever, although it is restricted to penalizing willful infringement. It is not generally seen as a way of addressing the cumulative innovation problem. The patent law standard for awarding enhanced damages produces a deleterious effect upon cumulative innovation. To reduce exposure for treble damages (which is based upon a finding of willfulness), patent attorneys routinely advise their clients (including research engineers and scientists) to avoid reading patent prior art, in effect negating a valuable aspect of the disclosure function of the patent system (Lemley and Tangri 2003). This is a form of overdeterrence of socially beneficial behavior – learning from prior discoveries. The inability to consult patent prior art undoubtedly results in duplication of research and may lead a researcher to overlook valuable potential solutions to scientific and technical problems. In order to alleviate this undesirable effect, the Federal Trade Commission (2003, pp. 28-31) recommends that the legal standard be tightened to require either actual, written notice of infringement from the patentee or deliberate copying of the patentee's invention, knowing it to be patented, as a predicate for willful infringement. Similarly, the National Research Council (2004, pp. 118-20) recommends elimination of the use of a subjective test for determining willfulness.

I.C.2.f. Channeling Doctrines. In the context of cumulative innovation, patent law serves a particularly important antidote to the non-disclosure aspect of trade secret law. By penalizing inventors that rely upon trade secret law – by subjecting them to the risk that a later inventor will obtain a patent on what they hold as a secret and thereby be able to block further use of the invention by the trade secret owner – patent law promotes disclosure.

I.D. Administration

Apart from the substantive rules, the administration of intellectual property law significantly affects the efficacy of the overall system. Administration plays the most significant role in patent law, which affords protection only for those inventions judged by a patent examiner to clear the validity hurdles. Therefore, we will focus our attention there. Two sets of issues have attracted substantial economic inquiry – the quality of patent analysis conducted by the Patent Office and administration of judicial decisionmaking. Although optional, registration of copyrights plays some role in the licensing. There are no formal requirements for trade secrets.

I.D.1. Registration/Examination

Registration and examination processes serve several potential functions in systems aimed at promoting innovation. They can screen out undeserving or defective applications, disclose knowledge to the public at large (published patents, deposit of copyrighted works with the Library of Congress), create

a public record of intellectual property titles that can be valuable for licensing and using intellectual property as collateral for debt, and, through the use of application fees, impose some of the system-wide costs of administration on those the most significant beneficiaries.

Due to the relatively high standards for patent protection, ex ante screening of patents through examination by experts trained in technology fields plays a vital role in minimizing disputes as to the validity of patents. An elaborate system of reexamination, reissuance, and interferences (for determining priority among inventions making the same or similar claims) uses skilled and specialized examiners and administrative law judges to resolve disputes.

In view of the low threshold for obtaining copyright protection, examination serves a rather modest role in the overall system. The copyright registration, which is optional, serves primarily as a title registry. Copyright registration also augments the public availability of knowledge through the deposit function. With the move away from formal requirements for obtaining copyright protection (including copyright registration and renewal) to an unconditional system, it is more difficult to identify copyright owners, thereby increasing the transaction costs of licensing (Sprigman 2004; Landes and Posner 2003, ch. 8) (suggesting a system of periodically renewable copyrights to reduce ownership tracing costs).

Renewal requirements and maintenance fees also function as policy levers. Such fees have traditionally been relatively low, being designed to cover the costs of administration, but nevertheless can cause rightholders to let their rights lapse (Lemley 1999). A large literature, based partly on European data (where such fees have been in effect longer) indicates that about half of patents lapse by the tenth year. (See, in particular, Schankerman (1998); see also Lanjouw 1998; Pakes and Schankerman 1984; Schankerman and Pakes 1986; Pakes 1985, 1986). Cornelli and Schankerman (1999) show how renewal fees can be used to give higher incentive for investment to higher-ability inventors, and Scotchmer (1999) shows how a renewal system can be used as a screening device to give more rewards to inventors who, although they may have higher costs, also have much higher value innovations.

I.D.2. Quality

Most commentators take as a measure of quality the probability that the patent would survive a legal challenge (National Research Council 2004; Shapiro 2004), although in a legal environment with unresolved patentability issues, issues of quality shade into issues of design (Scotchmer 2004c). Poor quality patents may result from inadequate review of prior art during examination, poorly drafted claims, or lax standards (the height of the non-obviousness threshold). They may undermine economic efficiency by restraining competition, raising transaction costs, and increasing litigation without promoting innovation. A proliferation of poor quality patents can choke entry and cumulative innovation.

Ensuring quality patents, however, comes at a cost. Emphasizing the large number of patent applications (up to 300,000 per year), the administrative and human resource costs of comprehensive patent examinations, and the relatively small number of patents having significant economic value (under

2,000 patent suits filed per year), Lemley (2001) asserts that the costs of improving patent quality ex ante (through more careful examination) exceed the benefits – the Patent Office is, in his view, rationally ignorant. Kieff (2001b) takes Lemley's insight quite a bit further, arguing for abandonment of patent examination in favor of a pure registration. The United States experimented with such a system in its early history (1793-1836), producing chaotic results. The Senate report accompanying the 1836 legislation reinstating formal examination commented that the registration system left the nation "flooded with patent monopolies, embarrassing to bona fide patentees, whose rights are thus invaded on all sides . . . Out of this interference and collision of patents and privileges, a great number of lawsuits arise . . . It is not uncommon for persons to copy patented machines in the model-room; and, having made some slight immaterial alterations, then apply in the next room for patents." S. Rep. No. 338, 24th Cong.., 1st Sess. 2 (1836). See Landes and Posner (2003, p. 309); Merges (1999).

But what about Lemley's more modest suggestion of shifting more resources toward screening at the litigation stage rather than the examination stage? As others emphasize, low patent standards raise litigation and licensing costs, impede cumulative innovation where dubious patents stand in the way of improvements, deter entry into promising markets, especially by small entrants that cannot easily withstand costly patent litigation, and encourage more filings, which may hamper the operation of the PTO (Ghosh and Kesan 2004; Rai 2003 (benefits of certainty in patent examination); Thomas 2002). The relative merits of screening at the examination versus the enforcement stage likely turns on the field of innovation. For example, given the large expenses needed to bring a drug from the testing laboratory to the market, the pharmaceutical industry cares deeply about knowing that patents are screened thoroughly prior to such investments. Uncertainty about patent validity could undermine investment in that sector. By contrast, software and business method patents likely do not require large up-front investments. The number of such patents is quite large and growing. Therefore, post-issuance screening – where selection is based upon which patents are litigated – may make more sense (so long as the chilling effects on new ventures are too great).

Ayres and Klemperer (1999) offer a less conventional argument for preserving uncertainty in the patent system. Based upon the fact that marginal price increases near the monopoly price only benefit the monopolist a small amount while producing disproportionately large deadweight losses to consumers, they show that even a small amount of uncertainty regarding the enforceability of a patent can alleviate monopoly pricing ex post without materially reducing incentives to innovate. They suggest that such uncertainty can be introduced into the system by having more lax patent examination (as well as other policy instruments, such as the standard for granting preliminary injunctions and application of the doctrine of equivalents). Their model, however, works only where there is uncertainty as to valid patents. Uncertainty as to invalid patents increases deadweight loss (Rai 2003). More generally, introducing uncertainty by loosening patent quality at the examination stage would impose significant additional costs upon third parties for prior art searches, opinion letters, and transaction costs (Merges 1999).

Several considerations suggest that patent quality has declined during the past several decades. The expansion of patentable subject matter to include software and business method patents has been cited a

particular concern. Since much of the technological knowledge in these fields subsists in the form of trade secret, products, and services, as opposed to more readily searchable sources (such as patents and published scientific journals), it is far more difficult to search these fields than the traditional patent fields (Galler 1990; Samuelson 1990; Merges 1999). Furthermore, as patenting of software and business methods took off, the Patent Office lacked examiners adequately trained in these fields. A second reason why patent quality may have declined is as a result of relaxation of substantive validity requirements. Since the consolidation of federal patent appeals in one court in 1982, observers of the patent system have discerned a clear lessening of the non-obviousness requirement (Lunney 2000-01, 2004; Barton 2000; Desmond 1993). The incentive structure and mission statement of the Patent Office may also have contributed to a lessening of standards. In general, compensation levels at the PTO may not be high enough to retain experienced examiners. More specifically, patent examiners are undertrained, overworked, and subject to distorted incentives that bias the system toward allowances. The bonus system – by which examiners get additional compensation for "dispositions" – favors allowances over rejections. Since rejected applications are easily revived through the continuation process, an examiner can more quickly and confidently secure a disposition through a grant than a rejection (Merges 1999; Rai 2003 (noting further that detailed explanations need to be provided for rejections but not allowances); cf. Lemley and Moore 2004 (advocating substantial limitations upon continuation practice); Quillen and Webster 2001-02). By contrast, there are no systematic compensation penalties for errors. On a macro level, the Patent Office is supported by fees paid by applicants. The Patent Office has over the past decade shifted its mission toward "customer service" – with the "customer" being the patent applicant, not the public (Corcoran 1999) .

Several sources of evidence support the concern over patent quality. Lunney (2000-01) documents the easing of the non-obviousness requirement by the Federal Circuit, finding that non-obviousness is far less likely to be cited as the basis for invalidating a patent. Several frequently cited patents – such as those issued for one-click ordering, a system for restroom queuing on airplanes, and crustless peanut butter and jelly sandwiches – reinforce the perception that inventions not need be particularly "inventive" to be patentable. U.S. allowance rates are believed to be in the 70 to 80 percent range, substantially higher than allowance rates at the Japanese and European patent offices (National Research Council 2003, pp. 43-45; Clarke 2003; Quillen and Webster 2001). These discrepancies, however, may be explained by the higher costs of obtaining protection in the United States. U.S. applicants may engage in more pre-screening of applications.

Empirical studies of patent quality, however, have not discerned a significant decline in patent quality. According to US PTO quality assurance audits, error rates have fluctuated between 3.6 and 7 percent since 1980, trending upward through the 1990s, but declining since that time (National Research Council 2003, p. 40). This data, however, is quite limited and has been questioned by external inspectors (National Research Council 2003, p. 40). Cockburn and Henderson (2003) find that resources devoted to patent examination (as reflected in examiner hours and actions) have kept pace with the increased number of applications during the 1985-97 time frame. But see National Research Council (2003, pp. 41-43) (reporting a 20 percent drop in the number of examiners per 1,000 applications from 1999-2002)

Using a data set of 182 patents litigated at the Federal Circuit between 1997 and 2000, Cockburn and Henderson (2003) do not find evidence that examiner experience or workload predict invalidity decisions. Several commentators have suggested that patent quality can be proxied by the number of prior art citations disclosed in the application: fewer citations indicate that the applicant has not provided the examiner with the range of existing knowledge against which to evaluate non-obviousness. Focusing on the subject matter area that has received the most heavy criticism on quality (as well as other) grounds – business methods – Allison and Tiller (2003a, 2003b) found that business method patents examined through the end of 1999 cited more patent and nonpatent prior art references than a large, contemporaneous, random sample of all patents. Focusing on the success ratio of patent applications as a measure of patent quality, Lesser and Lybbert (2004) find that patent standards have been relatively stable over the 1965-97 time horizon.

Even if patent quality has not declined, there could be substantial benefits from reforming the patent system to improve patent quality. Proposals have focused on different stages of the patent review pipeline: (1) initial patent examination; (2) incentives for disclosure of relevant information (including oppositions); (3) substantive standards; and (4) litigation-related reforms.

Patent Examination. Applying insights from the field of personnel economics to the design of the Patent Office, Merges (1999) advocates reforming of Office procedures and practices to emphasize examiner training, employee, and more systematic quality review of applications. He also recommends changes in the bonus system to balance rewarding application processing efficiency with quality review. Responding to public criticism, the Patent Office adopted a second stage of review for business methods. PTO (March 2000).

Eliciting Information. Several scholars have focused on the problem of eliciting good prior art information early during initial examination and possible later review of patent validity.
Although good prior art research data bases are available in many fields, patent applicants and their competitors often have the best access to the most current and relevant prior art. The duty of candor imposed upon applicants encourages disclosure of such knowledge, but it may not be adequate. Various proposals aim to elicit more and better prior art disclosure.

Under current practice, the duty of candor requires only that the applicant disclose all material prior art of which they are aware to the Patent Office. They need not explain how such art relates to the patent. As a means of encouraging more expansive applicant disclosure, Kesan (2002) advocates limiting the presumption of validity in subsequent litigation over the patent to art that the patentee disclosed and explained during prosecution.

Another approach to the information problem focuses on third party disclosure – mechanisms to enable competitors to participate in the examination or post-issuance review process. As currently constituted, the U.S. patent examination is largely ex parte – i.e., involving only the applicant and the Patent Office. Often, competitors of the applicant are in a good position to identify weaknesses in an application. For this reason, many observers recommend expanded use of inter partes (opposition) proceedings in

order to bring forward stronger resistance to weak patents, as occurs in the European Patent Office (National Research Council 2004; Federal Trade Commission 2003; Levin and Levin 2003; Rai 2003; Kesan 2002; Merges 1999; Soobert 1998). Publication of applications can also generate third party input to the examination system (Duffy 1998). The primary motivation for opposition is to prevent a competitor from gaining an unfair advantage in the marketplace through obtaining a questionable patent. This can prevent or limit commercial advantages in the marketplace as well as avoid subsequent litigation costs.

The efficacy of an opposition may well be hindered by a collective action problem. Even though competitors of a patent applicant collectively may gain more from defeating the patent than the costs of pursuing an opposition, each individual competitor may not have sufficient incentive to initiate the process (Thomas 2001). An alternative approach to eliciting pertinent prior art provides a financial reward to those who come forward with information to invalidate unwarranted patents (Thomas 2001 (examination stage bounty); Kesan 2002 (one-way fee shifting rule); Miller 2004 (litigation stage bounty)). The threat of such a bounty being exacted from the patentee creates a further deterrent effect on filing questionable patents.

Substantive Standards. A relatively direct means of increasing patent quality is to raise patentability standards. Some have proposed narrowing the scope of patentable subject matter – for example, by requiring that an invention have "industrial applicability," which would exclude some business methods claims (Thomas 1999); or excluding business methods or DNA sequences as a class. In 2001, the PTO issued written description and utility guidelines intended to signal that it would be applying uniform and stringent guidelines for DNA-related patents. (United States Patent and Trademark Office 2001). Both the FTC and NAS reform studies recommend reinvigorating the non-obviousness bar for all inventions (Federal Trade Commission 2003; National Research Council 2004).

Litigation-Relation Reforms. As an alternative to increasing the costs of patent examination, Lemley (2001) advocates that patent litigation rules be altered to allow better ex post quality review of those relatively few economically significant patents that become the subject of litigation. In particular, he recommends that the presumption of validity be removed. Rai (2000, 2003) comes at the problem from another direction, recommending greater deference to PTO rejections made on grounds of obviousness.

I.D.3. Judicial Administration

The patent system relies heavily upon the judiciary to interpret the Patent Act and adjudicate infringement cases. Although Congress has written several quite specific exceptions into the statute over the past two decades, the courts have played a more central and active role in determining validity requirements and infringement standards. The expansion of the patent domain into the software, business methods, and biotechnology has been driven almost entirely by judicial interpretation as opposed to new legislation. The non-obviousness standard, written description requirement, utility threshold, aspects of novelty (e.g., inherency doctrine), claim construction, and infringement analysis continue to be calibrated by the judiciary.

The design of judicial institutions governing patents can have significant effects on the operation of the patent system. It is useful to distinguish between the trial and appellate levels. In some nations, patent trials are heard by specialized and technically trained tribunals. By contrast, patent cases are heard in the first instance in the United States by courts of general jurisdiction, i.e., non-specialist courts. The location of technology industries (e.g., Silicon Valley (Northern District of California)), incorporation patterns (favoring Delaware), and attorney preferences for plaintiff-friendly juries (Eastern District of Texas) among other factors affect the geographical incidence of patent cases. Thus, several district courts have become relatively more specialized in the handling of patent cases. No systematic comparative international study has yet been made analyzing the advantages of specialization and technically trained jurists on patent adjudication. Several empirical studies highlight the high reversal rate for claim construction decision by district courts (Chu 2001 (finding that the Federal Circuit modified claim construction decisions by district courts in 44% of cases); see also Moore 2001 (finding 33% modification rate for a somewhat different sample period). Based upon institutional analysis, Rai (2003) advocates the use of specialized trial courts for patent cases in order improve the quality of technologically oriented fact-finding within the judicial system. See also Rai (2002); Wiley (2003) (advocating greater use by courts of technology experts as special masters).

Specialization at the appellate level has garnered substantial scholarly attention. Prior to 1982, appeals of patent infringement cases in the United States were heard in the regional circuit courts in which the district courts were located. This system produced inconsistency in patent law as well as high rates of invalidation in some courts of appeals. Not surprisingly, it also produced a good deal of forum shopping (Dreyfuss 1989; Commission on Revision of the Federal Court Appellate System 1975, pp. 217-21). In an effort to improve administrative efficiency, Congress centralized appeals of all patent cases (from both the Patent Office and district courts) at the Court of Appeals for the Federal Circuit in 1982. As some observers at the time surmised, such a move would likely go beyond harmonizing the law. Institutional considerations – such as tunnel vision, political influence in selection of jurists for the Federal Circuit, and socialization effects among the members of the court – would likely produce a pro-patent bias (Posner 1985).

Several studies have borne out this prediction. Since 1982, patent law has become both more unified and more favorable to patentees. Cf. Wagner and Petherbridge (2004) (reporting mixed results on unity). Federal Circuit decisionmaking has generally resulted in an expansive interpretation of the subject matter of the Patent Act, narrow interpretation of limitations (e.g., experimental use), lower thresholds for protection, higher infringement damage awards, and greater average patent scope (Lunney 2004 (finding that prior to 1982, courts were more likely to reject claims on invalidity grounds than non-infringement by a ratio of nearly three to one; since 1982, the ratio nearly reversed, with non-infringement becoming the dominant (68.1%) explanation for cases being dismissed and invalidity becoming comparatively rare); Landes and Posner 2003, 2004 (using regression analysis, finding that Federal Circuit has had a positive and significant impact on the number of patent applications, the number of patents issues, the success rate of patent applications, the amount of patent litigation, and possibly the level of research and development); Lanjouw and Lerner 1997; Kortum and Lerner 1998 (finding that, from 1982 to 1990, the Federal Circuit

affirmed 90% of district court decisions holding patents to be valid and infringed, and reversed 28% of judgments of invalidity and non-infringement); see also Allison and Lemley 2000 (finding that appellate judges appointed to Federal Circuit since 1982 have been significantly more likely to uphold patent validity)).

The normative implications of these effects are complex. The harmonization of patent law has reduced uncertainty about the law, discouraged forum shopping, and possibly promoted research and development spending in some sectors (Landes and Posner 2003, pp. 345-57). The shift away of validity-based policy levers, however, has made the patent system less sensitive to the diversity across the range of technological fields. Lunney (2004) concludes that the Federal Circuit has shifted the patent system toward a more uniform, one size-fits-all regime in which validity has become more routine and scope more narrow. In effect, the court has dampened several critical validity policy levers, limiting the versatility of the patent system to promote the diverse range of new technologies. Several scholars advocate a shift in the Federal Circuit's role, viewing it as the best situated institutions for producing a patent system that responds to the heterogeneity of inventive activity across the growing range of technological fields (Burk and Lemley 2002, 2003; Rai 2003; but cf. Wagner 2003).

I.E. Enforcement

In our discussion of policy levers, we implicitly assumed that the rightholder has little difficulty identifying, pursuing, and excluding unauthorized users. The design conclusions of the literature depend on that assumption. However, enforcement of intellectual property laws in the real world is far more complex than this stylized caricature. The profitability of rights can be changed by uncompensated infringement or by the terms of license on which rightholders are induced to license in the absence of strong rights.

We have already discussed the main remedies to infringement, damages and injunctions, and whether they are likely to deter infringement. We now augment that discussion by saying what is known about the costliness and effectiveness of enforcing intellectual property rights, drawing attention to some additional legal mechanisms.

Evidence on Patent Litigation. Comprehensive evidence on patent litigation can be found in Lanjouw and Schankerman (2001, 2004), based on litigation data assembled by the PTO (to whom patent litigation is supposed to be reported), patents themselves (which contain information on the technology and characteristics of applicants), the Federal Judicial Center (which assembles information on the disposition of cases, such as whether they settle and when), and Standard and Poor's database on companies that are publicly traded (which contains information on characteristics of the company such as size). See also Allison et al. (2004). Based on this data, the overall litigation rate is about 2 cases per 100 patents, concentrated on high-value patents. An earlier study by Lerner (1994), restricted to biotechnology patents, estimated that 6 in 100 patents were litigated. Litigation increased substantially over the 1978-1999 period, but the increase is attributable to the changing composition of patents, and to the overall increase in patenting. There was a 71 percent increase in patent grants from 1978 to 1995. Most of the increase in

patent suits has been in drugs, biotechnology, and computers and other electronics, which have always been highly litigated and have been increasing as a percentage of total patent grants. Thus, litigation has grown faster than patent grants.

The role of small entities (including independent inventors and firms that acquire patent portfolios for purposes of licensing), and particularly firms that do not themselves practice their inventions, in patent litigation has been the subject of growing interest. At least traditionally, small firms were at a disadvantage due to the magnitude of litigation and enforcement costs. Lanjouw and Schankerman showed that patents held by small firms are more likely to be litigated. Lerner (1995) concluded that small firms avoid technology areas where litigation is prevalent, and Lanjouw and Lerner (2001) showed that, in the litigation process itself, preliminary injunctions are used strategically by large firms against small firms. This pattern, however, appears to be shifting. In the dot com age, a proliferation of software and business method patenting has spawned a plaintiff's patent bar that aggressively enforces patents (Federal Trade Commission 2003 (referring pejoratively to a new class of "patent trolls"); Sandburg 2001; Meurer 2003; Allison et al. 2004 (noting the high percentage of patent purchasers (as opposed to inventors or original assignees) instituting patent litigation)). The asymmetric stakes of such litigation may in fact favor small enterprises, which have little to lose and much to gain by asserting patents against large enterprises.

Large entities with sizable patent portfolios often prefer to resolve their differences with cross-licenses (often royalty-free) rather than risk mutually assured destruction that can result from high stakes patent battles (Parchomovsky and Wagner 2005). Hall and Ziedonis (2001) found that between 1979 and 1995, semiconductor firms amassed large patent portfolios in order to deter litigation and to negotiate more favorable access to technology owned by competitors. A follow-up study indicates a spike in semiconductor patent litigation relative to R&D activity (Ziedonis 2003). More specialized semiconductor design firms – lacking complementary manufacturing assets – have a higher propensity to litigate.

Indirect Liability and Least-Cost Enforcement. Courts have long recognized liability for acts that contribute to infringement by others. Congress codified liability for contributory infringement, with limitations, in the 1952 Patent Act (35 U.S.C. §271). Similarly, copyright law extends liability to those who contribute to and vicariously benefit from copyright infringement.

Indirect liability can reduce enforcement costs by allowing an intellectual property holder to cut off infringement at a higher level in the chain of potentially responsible actors – such as suppliers of the means for infringement. It can also provide a more effective sanction when direct infringers are difficult to identify. Of course, the act which contributes to enforcement may also have a lawful purpose, e.g., the sale of a component part used in practicing a patented invention. For this reason, the law does not recognize contributory infringement if the acts or product sales have "substantial non-infringing uses."

Copyright Enforcement in the Digital Age. Above we cited arguments along the lines that limited sharing does not impinge on rightholders provided that they anticipate the sharing in their pricing behavior. These arguments were more suited to the analog age (e.g., photocopying) where a form of

"natural" encryption – the lack of availability of reproduction technologies, the degradation of quality of second generation copies, and the relatively high cost of making copies – limited unauthorized reproduction. Furthermore, anyone seeking to mass produce and distribute copies could be easily detected. Although copyright enforcement has long been a problem in some foreign markets (Ryan 1998), copyright enforcement was not a major worry in the United States during the analog age.

Modern digital technology has brought enforcement to the forefront of copyright policy throughout the world (Menell 2003). Such technology allows rich media content to be flawlessly copied and redistributed through largely anonymous peer-to-peer digital networks. In this environment, it is less likely that degradation or the cost of copying will protect proprietors, or that sharing groups will be limited in size.

Where infringement is particularly difficult to detect, preventive measures may be a second best means of preventing unauthorized use of intellectual property. In response, film, music, computer software, and computer games proprietors are turning to technical protections, such as encryption and copy controls. The economic consequences depend on how effective the technical protections are, a matter which is still evolving. As a means of enhancing the effectiveness of such technologies, Congress enacted a set of anti-circumvention provisions as part of the Digital Millennium Copyright Act (DMCA) which largely prohibit decryption of digital locks placed on content. Such preemptive protections, however, have the undesired consequence of preventing some otherwise lawful uses – e.g., fair use of an encrypted work. To reduce such effects and balance both under and over-enforcement, the DMCA contains numerous exceptions, such as for reverse engineering of software products for purposes of creating interoperable programs, security testing, encryption research, etc. The Act also empowers the Librarian of Congress to grant categorical exceptions.

If users will circumvent the protection system when the cost of circumvention is lower than the price, the threat of circumvention will have a moderating effect on the pricing strategy of vendors, which reduces per-period deadweight loss (Conner and Rumelt 1991). Park and Scotchmer (2004) point out that if the price reductions are achieved through technical protections that can be circumvented at a cost, and that the technical protections continue forever, just as trade secrets can, the net effect can be beneficial for both content providers and consumers. Because the price will be lower than monopoly price, the profit-to-deadweight-loss ratio will be lower. Consumers may be better off due to the lower price, and proprietors may be better off due to the longer protection. Thus, the transition away from the enforcement of legal protections to technical protection has an ambiguous effect on consumer welfare and on the incentives to create.

Because digital sound recording files are widely available (the compact disc encoding technology introduced in 1981 was not encrypted) and relatively small (in comparison to film files), the sound recording industry has been the first content industry to be affected on a large scale by the capabilities of the emerging digital platform (Menell 2003). Surveys and various other forms of empirical evidence suggest that teenagers (a prime target audience for new music and film releases) consider peer-to-peer networks to be

an attractive source for obtaining content. The overall effects on the content industries are complicated to assess, although the most recent studies seem to suggest that peer-to-peer technology is at least partially responsible for the post-2000 decline in record industry revenues. See Liebowitz (2004) (finding that peer-to-peer file sharing has caused harm); but see Oberholzer and Strumpf (2004) (questioning a link between free downloads and CD sales).

In order to combat unauthorized distribution for the purposes of bolstering traditional retail sales and building support for legitimate online distribution (subscription and download services), the music industry initiated a high profile enforcement campaign against distributors of peer-to-peer software. After an initial victory against a centralized peer-to-peer technology (Napster), the record industry has encountered difficulty shutting down more decentralized networks on legal (newer technologies are outside of the software providers' control and have non-infringing uses) and practical (off-shore providers) grounds. As a result, the record industry has begun pursuing individual uploaders directly, although this is a costly process due to the relative anonymity of filesharers.

Economic analysis of copyright enforcement in the digital environment involves several complex considerations. Allowing greater leeway for courts to hold distributors of peer-to-peer software indirectly liable for infringement has the advantage of economizing on enforcement resources, but it produces a chilling effect on legitimate uses of such technology and discourages the development and diffusion on new digital technologies that might have substantial societal benefits. Some loosening of the "substantial non-infringing use" defense may be called for to balance the competing effects on aesthetic creativity on the one hand and technological innovation on the other (Menell 2005; Lichtman and Landes 2003). Several scholars advocate abandoning direct enforcement in favor of a levy system (fees on technology and Internet services that operate as a compulsory license) as a means of supporting creative enterprise, although such approaches cannot price usage efficiently and introduce administrative costs and rent-seeking behavior (Netanel 2003; Fisher 2004; Gervais 2005; but see Merges 2004a). Lemley and Reese (2004) suggest that limiting enforcement to actions against direct infringers through a streamlined and lower cost administrative enforcement process would provide the best compromise between deterrence and compensation on the one hand and freedom to innovate on the other. Their proposal, however, would entail substantial administrative cost.

Another dimension of enforcement policy relates to the choice between public and private enforcers and the penalty structure. Public enforcement can offer advantages where the government has easier access to information about infringing behavior, can realize economies of scale not achievable by private enforcers, or can impose sanctions (e.g., imprisonment) that are more effective than civil penalties. Exclusive government enforcement may be appropriate where there is some benefit to be gained from prosecutorial discretion. The federal government has expanded criminal penalties for unauthorized online distribution of copyrighted works.

I.F. Interaction with Competition Policy

Intellectual property protection can conflict with competition policy. We discuss the principal economic theories bearing on this tension here (for comprehensive analysis of the intellectual property/antitrust interface, see generally Hovenkamp, Janis, and Lemley 2004). The chapter on antitrust also examines this issue.

There are two stages at which antitrust concerns can be raised in the intellectual property context: in the rivalry to achieve inventions in the first place, and in the licensing that takes place *ex post*. Licensing is generally thought pro-competitive, since it increases the use of intellectual property. Further, licensing is common. Among members of the Intellectual Property Owners Association, 17.6 per cent of patents are licensed out, and many innovators invest with the sole objective of licensing rather than practicing or manufacturing their innovations (Cockburn and Henderson 2003). In the content industries, many works are independently produced and distributed by larger companies that finance and/or license the copyrighted products. Because licensing creates alliances that affect production, distribution, and pricing, such transactions inevitably raise competition issues.

Some of the pro-competitive uses of licensing were discussed in the context of cumulative innovation. These include licensing to resolve blocking patents, and to ensure the widespread use of complementary pieces of technology such as research tools. Here we discuss the more traditional context of horizontal substitutes and the special circumstances of licensing complementary intellectual property.

Licensing is pro-competitive whenever the efficient use of the intellectual property is to share it. However, a problem arises when the licensor and licensee are rivals in the market. The terms of license must facilitate the sharing of technology without at the same time facilitating collusion. This is a fine line to walk, and since the firms will not be inclined to walk it, the proper boundary must be established as a matter of law.

For example, suppose the technology reduces the marginal cost of producing a product. A license from the patent holder to a rival creates a social benefit by cutting the rival's costs. But if the royalties are higher than the cost-saving, the license can result in a market price that is even higher than would prevail in the absence of licensing. In that case, consumers do not benefit from the cost reduction. Should this be allowed?

As in other antitrust areas, the governing principle in U.S. law has increasingly become rule of reason; see Gilbert and Shapiro (1997) for a discussion of the per se rules (the "nine no-nos") that have come and gone in U.S. law and policy. We will not give a comprehensive overview of specific licensing rules that have been in and out of favor, but will instead articulate some of the economic principles that have been suggested as a basis for adjudicating licensing practices.

In general, rule of reason is a test that weighs harms to competition against gains in efficiency. But this is not a very practical test in the intellectual property context, since efficiency can be either *ex ante* or *ex post*. Even if a licensing practice seems collusive *ex post*, the parties may argue that the prospect of

using it is what gave them incentive to invest in the first place. It is hard to see what kind of evidence would either contradict or buttress such a claim, especially in a research environment where, *ex ante*, success was not assured. In that case, a firm will only invest if it earns substantial profit (higher than cost) in case of success. Even more importantly, it is not clear how such an inquiry respects the presumed right of Congress to set the incentives for research.

A slightly more practical test, which is at least founded on a sensible and clearly articulated principle, is that of Kaplow (1984), who recommends that a licensing practice be approved if it allows the rightholder to earn profit in a way that increases the profit-to-deadweight-loss ratio. The conceit is that Congress anticipates this efficiency principle in setting the other policy levers, such as length, so that the courts are implementing Congress's will. A problem with the principle, however, is that it has no natural boundary. Into which markets is the rightholder allowed to leverage? For example, if it is efficient to raise money by taxing real estate, then shouldn't the intellectual property be licensed collusively to real estate owners? It is not obvious how the principle mediates between the incentive purpose of the patent grant (the patent should not be lucrative unless it creates value to users) and the problem of raising money through efficient taxation, in whatever market that can be done most efficiently.

Maurer and Scotchmer (2004a) claim that courts have implicitly addressed this problem by applying a principle they call "derived reward," which means that the profit can only be earned by collecting some of the social value created by the invention. In fact they argue that courts have employed (and previous commentators have implicitly endorsed) three principles that jointly constitute a sensible guide for adjudicating license disputes: profit neutrality, derived reward and minimalism. Profit neutrality means that the rightholder should not be penalized for his inability to work the patent efficiently himself. (This principle may, for example, justify price-fixing in the patent context.) Minimalism means that courts should not allow terms that are unnecessary in achieving the first two principles. Unnecessary terms only give opportunity for sham licenses.

The problems are compounded when a user needs many complementary licenses. Both cross-licensing and patent pools can compound the concerns about competition. For a discussion of cross licensing, see Barton (2002); Denicolo (2002b); Merges (1996, 1999); Gilbert (2002); Lerner and Tirole (2004). Patent pools are generally suspect when they contain substitute technologies, but not when they contain complements. Price-fixing by a pool that contains substitute patents will generally raise the joint price relative to individual licensing, whereas price-fixing by a pool that contains complements will generally lower the joint price relative to individual licensing.

Aside from their effect on prices, cross licenses and patent pools affect the incentives to create and improve technologies in other ways. The division of profit among rightholders in the pool determines their rewards. The division of profit has not been the focus of the literature, but there is no reason to think that profit will be divided as necessary to cover the respective innovators' costs. Looking forward instead of backward, if all members of a pool will share equally in the benefits of new knowledge, then any member's incentive to invest in new knowledge is attenuated. Pooling may reduce the incentives to innovate.

The second concern of antitrust policy is how *ex ante* alliances (rather than *ex post* alliances) affect incentives to innovate. The official policy of the antitrust authorities in this regard is articulated in the 1995 *Antitrust Guidelines for Licensing Intellectual Property* (U.S. Department of Justice 1995). The *Guidelines* distinguish between "technology markets" in which firms license intellectual property that already exists, and "innovation markets" in which firms compete to develop new technologies. The policy with respect to innovation markets addresses two fears: that alliances may retard progress by reducing competition to innovate and reduce the number of substitute innovations, and undermine competition *ex post* in a product market.

The *Guidelines* assume that rivalry in innovation generally improves welfare – more rivalry will lead to greater investment, which will in turn produce more rapid innovation. Cf. Loury (1979), Lee and Wilde (1980), Reinganum (1981, 1982, 1989), Merges and Nelson (1990). Rivalry might, however, result in duplication of costs without yielding more innovation, dissipating the value of innovation (Barzel 1968; Kitch 1977; Grady and Alexander 1992; Gilbert and Sunshine 1995a). For this reason, the prospect theory of patent policy favors non-rivalrous exploitation of innovation opportunities, whereby an initial prospector obtains “breathing room” to develop the claim without fear that rivals will preempt or steal the claim and the inventor will be able to coordinate the development process (Kitch 1977). The opportunity to license the technology enables the inventor to contract with entities that may be better able to develop the claim. The prospect theory thus turns importantly upon a smoothly functioning technology licensing market and the capacity, foresight, and rationality of prospectors to coordinate the development and diffusion of the technology.

In theory, therefore, the effects of rivalry on economic welfare are ambiguous. Whether competition promotes innovation better than coordination depends, among other things, on the nature of the innovative process and the innovative environment. Lurking behind the disagreement is Schumpeter's classic 1942 book, arguing that market concentration encourages innovation. The *Guidelines* largely reflect the opposite view, that concentration inhibits innovation. There is a large, but inconclusive empirical and theoretical literature on this question, originating with Arrow (1962).

The *Guidelines* reflect the policy of the antitrust agencies, which is not necessarily the law as interpreted by the courts, which apply a rule of reason standard. Cost efficiencies that might be considered include delegating effort to the more efficient firms (Gandal and Scotchmer 1993), sharing technical information that might be hidden if firms compete (see Bhattacharya et al. 1990, 2000; Brocas 2004), sharing spillovers of the knowledge created (see Katz and Shapiro 1987; d'Aspremont and Jacquemin 1988; Kamien, Muller, and Zang 1992; Suzumura 1992; Aoki and Tauman 2001), or avoiding duplicated costs. See the appendix to Hoerner (1995) for a compendium of early cases in which courts and the agencies have made judgments about the relative merits of various arguments, and also Gilbert and Sunshine (1995a, 1995b).

Turning to the second concern – that alliances might undermine competition *ex post* in product markets – mergers or other alliances can lessen competition where the combined enterprise develops a

single product where the separate entities would have developed competing products. Evaluating the welfare effects of such alliances puts courts in the difficult position of predicting what types of intellectual property the members of a proposed alliance would develop absent the merger. The firms that propose to merge will presumably not announce that they would otherwise develop noninfringing substitute products. Instead they will argue that competition will be wasteful and duplicative, and that only one firm will, in the end, have a viable product. Given this incentive to dissemble, the agencies and the court might rightfully be skeptical. Cf. Shapiro (2003).

Related issues arise in the rules governing standard-setting organizations, patent pools, and cross-licensing agreements. Such agreements can promote consumer welfare by facilitating innovation in network industries and facilitating the development of products incorporating the most advanced technologies or where different entities hold mutually blocking patents (Barton 1997; Lemley 2002 (standard-setting organizations); Shapiro 2001; Lerner and Tirole 2004 (patent pools); Merges 1996 (patent pools; copyright collectives); Besen, Kirby & Salop 1992). Given the transaction costs of licensing (including the costs and delays in resolving disputes about intellectual property rights) and the importance of standardization in many markets, such institutions can play a critical role in promoting innovation and commerce. Nonetheless, like any agreement among competitors that can exclude competitors and potential entrants, such licensing must be scrutinized to ensure that the pro-innovative benefits outweigh the anti-competitive costs.

I.G. Organization of Industry

The organization of industry can affect the incentive to do R&D, and, in reverse, the task of doing R&D can be a reason that industry wants to reorganize. In this section we discuss (1) the Schumpeterian (1942) hypothesis relating to monopoly and innovation, (2) the role of employment relationships, geographic concentration of innovation resources, and contracting patterns in promoting innovation, (3) patent races and alliances such as research joint ventures, (4) systems competition and network industries, and (5) the open source movement.

Much of the research on the role of industry structure on innovation traces back to Schumpeter's (1942) hypothesis, based largely on empirical grounds, that large, monopolistic firms are more innovative than small, competitive firms because of their superior ability to marshal resources for large R&D programs. The hypothesis, if true, has three important implications. First, if large firms have an exaggerated incentive to do R&D, then R&D perpetuates monopolies rather than controlling them. But this is not necessarily bad if more monopoly means more progress. Second, if monopolists have more incentive than rivals to patent close substitutes, as suggested by Gilbert and Newbery (1982), then the analysis of patent breadth summarized in the section on policy levers may be moot. The analysis is based on competition between rival patentholders, which is not relevant if patents on substitutes are likely to be held by a single firm. Third, if size increases the incentive to innovate, then an antitrust analysis based on rule of reason would be less hostile to merger among innovative firms than otherwise.

Subsequent empirical and theoretical work of the Schumpeter hypothesis has proven inconclusive. Survey research by Levin et al. (1987) and Cohen et al. (2000) suggests a much more complicated relationship between market structure and innovation than suggested by Schumpeter. On purely theoretical grounds, Arrow (1962) showed that monopoly can reduce the incentive to invent, while at the same time making invention more valuable. Suppose that the innovation in question is a cost-reducing innovation, and suppose that the cost reduction is so large that the innovator will become a monopolist even if the market was previously competitive. Compare the following two situations: Prior to the innovation, the innovator operates in a perfectly competitive market, or, prior to the innovation, he is already a monopolist. Then, contrary to Schumpeter's hypothesis, the incremental profit that the innovator earns by innovating is larger if he begins as a competitor than if he begins as monopolist. This is because, as a monopolist, he would have earned some profit in any case. On the other hand, the gain in consumers' surplus is larger if the innovator starts as a monopolist, since consumers then started with already high prices. Thus, when the innovator begins as a monopolist, innovation creates less profit and more consumers' surplus than when he begins as a competitor. As a consequence, it may be optimal to offer greater profit incentives, e.g., through patent life, but there is no way to achieve that, since intellectual property rights cannot depend on market structure.

This inquiry relies on the notion that there are commonly known opportunities to produce knowledge (the "production function" model). If ideas are scarce, then patents on substitute technologies are less likely to become concentrated as a consequence of incentives. Since there is only one monopolist and many rivals, a rival is more likely to think of any given competing product than the monopolist.

Employment conditions as well as the geographic concentration of industry can have a strong effect on the pace of innovation. In comparing Northern California's Silicon Valley to Boston's Route 128 corridor, which were comparably positioned to lead the digital technological revolution, Saxenian (1994) found that Silicon Valley's free wheeling culture of encouraging exchange of information and mobility of labor across companies proved more successful than Route 128's more proprietary, staid, and vertically integrated business ethos. California's legal limitations on non-competition agreements as well as its competitive venture financing network fostered sustained rapid technological progress and relatively stable economic growth, defying the predictions of product cycle theory (positing that regions follow a pattern of innovation, growth, maturation and scale production, and ultimate decline as production shifts to other, lower cost regions) and the production-function model of knowledge creation.

Thus, the organization of industry has important impacts on the success of R&D and the discovery of knowledge. In reverse, intellectual property also affects the organization of industry, in sometimes surprising ways. Of these we mention three: the incentive for competitors to collaborate in research, the organization of network industries, and the open source movement.

Much of the earlier economics literature was devoted to studying patent races, e.g., how many firms would enter a race, how intensively they would compete, and at what point there would be a shake-out. Aspects of the intellectual property environment that affect these matters are the private value of the patent

right (Loury 1979; Lee and Wilde 1980; Dasgupta and Stiglitz 1980a, 1980b; Tandon 1983; Reinganum 1982, 1985, 1989; Wright 1983; Denicolo 1996), legal details such as whether interferences are resolved in favor of the first firm to file or first to invent (Scotchmer and Green 1990), the degree of spillover in knowledge that will occur after the invention (Katz and Shapiro 1987; d'Aspremont and Jacquemin 1988; Kamien, Muller and Zang 1992; Suzumura 1992; Aoki and Tauman 2001), whether the firms can learn from observing each other's investment strategies, either about the other firm's research efficiency (Choi 1991) or the other firm's private information about the value of the objective (Minehart and Scotchmer 1999), and whether licensing would occur to prevent it (Gallini 1984; Gallini and Winter 1985; Shapiro 1985; Rockett, 1990).

We mentioned at the outset that one of the defects of intellectual property as an incentive mechanism is that the investments it incites might not be efficient. First, the private return to entering a patent race is different from the social return. Second, the patent race does not aggregate or use the firms' private information about their relative efficiency or the value of the investment (Gallini and Scotchmer 2002).

With respect to the first point, part of the entrant's reward is a transfer from the other participants. When the entrant's probability of winning goes up, the other firms' probabilities of winning go down. To the extent that these effects are offsetting, entry creates a benefit for the entrant, but not for society as a whole. Thus there may be too much entry. There will almost certainly be too much entry if the private value of the right is equal to the social value. There may alternatively be too little entry if the private value of the intellectual property right is low relative to its social value, or if the innovation will create unappropriable, but beneficial, spillovers among firms.

The other inefficiencies that arise in patent races are due to the imperfect sharing of information about cost efficiency or the value of the objective, an unwillingness to disclose intermediate steps of progress (Scotchmer and Green 1990), and an unwillingness to share technical information (Bhattacharya, et al. 1992). Many of these inefficiencies can be solved by forming a joint venture to share information of the various types, and thus allocating R&D effort efficiently. This problem has been studied in various contexts using the methods of mechanism design (Bhattacharya et al. 1992; Gandal and Scotchmer 1993; Brocas 2004).

Collaborations among innovative firms through merger or a joint venture can have the beneficial effect of avoiding the inefficiencies of a patent race. However they can also be anticompetitive, and are therefore a matter for antitrust scrutiny, regardless of whether the firms have market power in a product market.

Systems competition has come to play a critical role in the digital technology field (Matutes and Regibeau 1992; Katz and Shapiro 1986; Farrell and Klemperer 2004). A "system" has complementary pieces, such as a computer operating system and compatible software. The distinguishing aspect of a "system," as opposed to other complementary products, is that the two pieces of the product must be made compatible by some kind of interface. There are three features of a system that might be protected with

intellectual property: the hardware (platform), the interface, and the software (applications). Which, if any, should be protected?

When the interface is itself proprietary, the system is called "closed," and otherwise "open." Whether or not the platform and applications are also protected, open and closed interfaces lead to different market structures. With open interfaces, firms may enter on both sides of the market to create products compatible with the complementary ones. With closed interfaces, the two sides of the market must be supplied by an integrated firm, namely, the firm that controls the interface. This control may be exercised by licensing the right to make compatible applications, perhaps with an exclusive dealing clause.

Especially in the context of network effects, a proprietary interface may become an important determinant of market structure. Network benefits arise when the value of using the system increases with the number of other users. As a consequence of network benefits, the market may "tip" to a single integrated system, such as the Microsoft Windows operating system and applications. The threat of tipping is reinforced if an entrenched platform owner has more incentive to increase the number of applications because he has more customers.

With an open interface, a system is not likely to remain under integrated ownership, due to entry. In contrast, due to the tipping phenomenon, a proprietary interface can create market power and profit far beyond the value of any social value it provides. Indeed, the interface can be entirely arbitrary, and not have any social value at all. Protection of the interface thus serves a very different economic purpose than protection of the intellectual property in operating systems or applications.

It seems natural to protect the innovations on the two sides of the market (platforms and applications), since they represent the costly and creative endeavors for which intellectual property is intended. If both sides are adequately protected, there is no need to protect the interface as well. The resulting open market structure would be similar to any market with complementary goods. However, this outcome may be difficult to achieve. The interface may be protectable under copyright or patent law, although strong economic arguments can be made on the basis of network economics that the thresholds for such protection should be quite high and that rights should not be exclusive. See Menell (1987, 1989, 1994, 1998b, 2003); Cohen and Lemley (2001); O'Rourke (2000). Even if such protection is not available, interfaces may be protected through encryption (and trade secrecy). Depending on the complexity of the encryption, reverse engineering may be an antidote subject to the limitations of anti-circumvention constraints (Samuelson and Scotchmer 2002). For an analysis based on protection of interfaces, rather than the two sides of the market, see Farrell and Katz (1998).

The open source movement has developed in part as a response to the constraints of closed systems (McGowan 2001; Benkler 2002, 2004). The movement developed in the computer software industry around programming efforts to develop the Apache web server and the Unix operating system, under the name Linux (Raymond 2001; Lerner and Tirole 2000). It employs intellectual property protection in an unconventional manner: as a means of precluding innovators building upon the open source

platform from asserting intellectual property right to exclude others. In addition to this precommitment attribute, the decentralized, collaborative innovation process underlying open source development provides advantages in addressing the myriad complex manifestations of flawed ("buggy") computer code. With more eyes and more uses, more bugs will surface, and those who find them can easily rewrite the code to fix them. Users may find it convenient to develop the code for their own idiosyncratic purposes, but the social value is much larger if the same code can also be used more broadly. The open-source movement exploits this potential by ensuring that all innovators in the open-source community make their source code available.

From the point of view of intellectual property, the open source movement is interesting because is uses copyright for a purpose opposite to its customary one, thereby spawning the term "copyleft." Through the use of a form of copyright license, a software developer can make his code available under a license which allows access and does not require royalties, but requires, for example, that the user make his own derivative product available on exactly the same terms. Such licenses are called "viral" -- products are infected with a self-replicating term that cannot be shaken by creating a new product. In this way, the community keeps the code open, observable, and useful to a broad community.

Who, in the end, pays for all this code if no one pays royalties? Lerner and Tirole (2000) stress career concerns coupled with personal uses, but admit that the economic models we currently have are not well adapted to explaining the phenomenon. None of the four models of the creative environment identified at the outset seems suitable. As stressed by von Hippel (2001), the system seems to work best in an environment where at least some developers have in-house uses. It is hard to see how it would work for strictly a mass market. The two most important examples, Apache and Linux, are notably less user friendly than their rivals in the mass market, e.g., the Microsoft server and Microsoft Windows operating system.

Recent developments, however, suggest that traditional industry players may well see support for open source software as a means of dethroning Microsoft from its long-standing monopoly position attributable to its widely adopted proprietary operating system products. The open source community's pre-commitment to non-proprietary software development creates a means for commercial enterprises that can profit from complementary products (hardware) and service businesses (such as consulting and maintenance) – including IBM and Hewlett-Packard, as well as newer companies such as Red Hat which specializes in supporting Linux – to move outside of the shadow of Microsoft's influence and compete more effectively in other computer business sectors (Merges 2004b). Such "property preempting investments" may be a successful commercial business strategy.

Biomedicine is another realm where industry has organized to cut back on the exercise of intellectual property rights. Biomedicine has become heavily reliant on data and databases which contain gene sequences that may be patented or protected as trade secrets. Science as a whole is more efficient if researchers can share the data produced by others. Licensing such data piece by piece would impose prohibitive transactions costs. Instead, researchers are experimenting with collaborative business models to assemble the data into databases. In the SNP consortium, they have renounced intellectual property

rights altogether (Merges 2004b; Maurer 2003).

H. Comparative Analysis: Intellectual Property versus Other Funding Mechanisms

Public and private funding mechanisms for R&D (and for creative works, to a lesser extent) have always existed side by side. In the U.S., the share of R&D funding provided by the public sector has seldom dropped below a third in the last half of the twentieth century, and in 2000 was about 26% (National Science Board 2002). In most OECD countries the public share has been closer to half. Between the 19th century and the late 20th century, public funding has shifted from a system of *ad hoc* initiatives to a routinized system based predominantly on peer review, with researchers competing for large federal budgets allocated before the recipients are named.

Since public sponsorship can reduce the restrictions on use that afflict intellectual property, and perhaps improve the way that R&D is organized, why isn't all R&D funded by the public sector? What accounts for the mix between public and private incentives? We return to those questions after commenting on the variety of public funding mechanisms currently in use.

A public funding mechanism that has been used more or less continuously throughout history is direct government employment of researchers. This is a system with obvious virtues when the sponsor is the only beneficiary of the resulting knowledge, or the benefits cannot be appropriated by a commercial vendor. However the defects of this system are many. Perhaps most importantly, it does not make use of the imagination that is widely dispersed in the population, and does not recognize that, for any given research task, some other party may be better equipped to perform it. It is an odd conception of research that starts from the premise that we know what we want to discover, we know how to discover it, and we know who can achieve it at least cost, namely, our employee. In what sense is that promoting discovery? In-house research will work rather badly in the ideas model, but much better in the model of induced change and the production-function model, since the investment opportunities are commonly known.

Like in-house research, prize systems also have a long history, continuing to the present. Prizes share several important features with patents. On the virtues side, they can attract investments from unexpected quarters, but on the defects side, will not reliably delegate the research effort to the most efficient firms. Prizes avoid deadweight loss, but prize authorities have two challenges that patents automatically avoid: the problem of choosing the value, and the problem of making it credible that they will, in fact, award the prize.

Of course, depending on what is observable, the flexibility to choose prize values can create an improvement over patents. A unifying theme is that, if a prize giver can base the prize on the value of the innovation, then he should do so, and prizes may dominate intellectual property rights (Wright 1983; de Laat 1996; Kremer 1998; Scotchmer 1999; Gallini and Scotchmer 2002; Shavell and Van Ypersle 2001; Abramowicz 2003). Observations of the value can take many forms. Foray and Hilaire-Perez (2000) discuss how the silk-weaving guild in 18th-century Lyons used a prize committee to make judgments about

value direction. Kremer (1998) argues that the value can be observed ex post by auctioning a patent right, and completing deal (transferring the patent right) with small probability, otherwise putting the invention in the public domain and giving a reward. Shavell and Van Ypersle (2001) suggest that the reward giver can link the reward to sales. Of course it may be more sensible to link the prize to the expected costs instead, and in this sense, prizes are more flexible than patents.

Maurer and Scotchmer (2004b) categorize prizes as *targeted* and *blue-sky*. For targeted prizes, such as the ones that were offered for solving the problem of longitude (Sobel 1995), it is natural to assume that the sponsor, being the one to post the prize, has a good idea of the social value, and can link the prize to it, or alternatively to the expected costs of achieving a solution. For blue-sky prizes, it is harder to tailor the prize to either cost or value. Blue-sky prizes are given for achievements that were not anticipated by the sponsor, so the prize cannot be established in advance.

Prizes can only work if the prize giver can commit not to renege, and will work best if the prize, like a patent, can increase with the social value of the invention. However costs are extremely hard to measure, especially when different inventions are supported with the same overhead and research projects have risky outcomes. Three possibilities for how to accomplish this are (i) to offer the prize against a backdrop of patents (Kremer 1998), (ii) to structure the mechanism as a contest (Che and Gale 2004), and (iii) to link the prize to performance standards as was sometimes done at Lyons. In addition, of course, there must be some means to ensure that the prize giver does not renege.

With patents as a fallback option, an inventor would not accept a prize less than the patent value. The prize value will thus be linked, like patents, to the value of the innovation. However, the prize giver must have some means to discover the value. A scheme suggested by Kremer (1998) is for the prize authority to take possession of the patent. The invention is put up for auction, although it is awarded to the highest bidder with only small probability. In most cases, it is dedicated to the public. But the small probability the patent is transferred to the highest bidder provides incentives for honest bidding, which yields the reward amount paid by the prize authority to the inventor (regardless of whether the invention is granted to the highest bidder).

A contest is a prize coupled with a commitment to give away the money, e.g., through the by-laws of a foundation or a trust. The commitment overcomes the problem of reneging. Nobel prizes are in that category. Contests can be structured so that the reward reflects costs instead of value. In the contest described by Che and Gale (2004), the contestants bid against each other before investing, making contingent contracts with the sponsor for what he will pay, conditional on choosing each contestant's invention *ex post*. The price only depends on which invention is chosen, and is thus easy to enforce. Because the firms compete on the contingent contracts, and will only be paid if chosen, they have an incentive to keep the contingent price low in order to be chosen. On the other hand, if a firm delivers a worthless innovation, the innovation will not be chosen even at a low contingent price. Such contests are sometimes called prototype contests, and they have been used by the U.S. Air Force, for example, in procuring fighter jets.

Prizes and contests share with intellectual property the inconvenience that the inventor is rewarded *ex post* rather than *ex ante*, and must therefore find funding. Government grants are a funding system that overcomes this problem. Of the R&D that is funded by the federal government, only about a quarter is performed in government laboratories. More than half of the National Institutes of Health budget and almost all of the National Science Foundation budget is given out as grants. Even the national labs, which used to be funded directly by the Department of Energy, now compete for funds in peer-reviewed grant processes. The government grant process improves on in-house research in that it taps into the scarce ideas likely to be found elsewhere.

As an incentive mechanism, the problem with grants is that applicants may propose research they cannot accomplish or wastes the funds. Since the whole point is to pay the costs of research as they are incurred, grant-giving organizations do not take the money back if the research fails, and have little recourse if the grantee wastes the funds (other than costly monitoring). But despite the limitations on oversight, the repetitive nature of grantsmanship exerts a discipline. A researcher can be cut out of the system if he does not deliver the research results he promised. For highly productive researchers, this threat will keep them honest, although the system will be more costly than if oversight could be exercised directly (see Maurer and Scotchmer (ch 8 of Scotchmer 2005b)).

Finally, again following Maurer and Scotchmer, chapter 8 in Scotchmer (2005b), we turn to the "hybridization" of public and private institutions in the late 20th century. In year 2000 in the U.S., approximately 75% of total R&D was performed by industry, but only about 68% was funded by industry (NSF 2002). Most of the rest was made up by the federal government. For most of the 20th century, more federal funding has gone to private firms than to universities, mostly from the Departments of Energy and Defense.

Not only is the private R&D sector infused with public money, but the public R&D sector is also infused with private money, at least if one includes universities and national labs as part of the public sector. That is, public and private funds are blended both in private industrial laboratories and in laboratories that have traditionally produced knowledge dedicated to the public. Further, the outputs of federally funded research are increasingly patented and exploited by the private sector under legislation enacted in the 1980's (the 1980 Bayh-Dole Act for universities, the 1980 Stevenson-Wydler Act for national labs, and the 1984 Technology Transfer Act, authorizing the creation of cooperative research and development agreements (CRADAs) (Mowery et al. 2001). This has turned out to be highly controversial. What is the rationale for subsidizing research that will ultimately be subject to intellectual property rights? Why give intellectual property rights on publicly funded research? The purpose stated in the Bayh-Dole Act is "to promote utilization of inventions arising from federally supported research or development .... without unduly encumbering future research and discovery" (Section 200). On the basis of this language, one might guess that the rest of the Act prohibits patenting, since patenting gives the right to restrict use. To the contrary, the point of the Act is to authorize patenting. The Bayh-Dole Act rests on the unlikely premise that the best way to diffuse innovations is to allow exclusions on use, subject to limited and rarely exercised "march in" rights (Eisenberg 1996; Mikhail 2000; Kieff 2001b).

This contradiction is usually reconciled by positing that, without protection of the underlying science, firms will not make the collateral investments required to commercialize it. But it is a well established principle of patent law that improvements and new uses are themselves patentable. If so, this argument has no force. In any case, as many have argued, e.g., David (2003), Scotchmer (2003), Lemley (2004), it would be better to fix the thing that is broken (patent law) than to compromise open science.

We know of only one other justification for the policies that authorize private firms to leverage public money toward ownership of valuable intellectual property. Many gifts from the private sector are like matching funds that industry gives in return for intellectual property rights. Because industry can choose what to match, this system selects the projects that are likely to be commercially valuable, and thus serves the two purposes of allowing the public to subsidize expensive research while at the same time getting the benefit of private expertise in screening investments (Maurer and Scotchmer 2004a).

The controversy over patenting discoveries in the university is not new. Such patenting has been going on since the late 19th century, at least for discoveries that were not funded by federal grants (Mowery and Rosenberg 1998).

I. International Treaties

In studying the optimal design of intellectual property, economists typically assume that the objective is to maximize consumers' surplus plus inventors' profit net of development costs. But whose consumers' surplus and whose profits? Externalities and profit flows across borders change the design problem. Domestic inventions create consumers' surplus abroad, and if protection is available abroad, also generate profit; for empirical evidence, see Alston (2002), McCalman (2001). In reverse, a strengthening of domestic protections will create an outflow of profit. How does this change the design problem?

The profit flows and externalities are governed by international treaties. Treaties create two types of obligations: for national treatment of foreign inventors and for certain harmonized protections. National treatment means that foreign inventors receive the same intellectual property protection as national inventors, while harmonization means that the countries have agreed on at least some aspects of what will be protected. Otherwise all countries could have different protections. These reciprocal obligations affect the rewards to innovation, the balance of trade, and foreign direct investment (Maskus 2000a, 2000b). About half of American and European patents are issued to foreign inventors, and about 20% of Japanese patents (European Patent Office 2002a, 2002b). The treaty obligations also extend to copyright, although there are no analogous administrative data that would allow us to asess their importance.

If the only objective is to minimize the deadweight loss of achieving a given amount of profit, then innovations should be protected in markets where the profit-to-deadweight-loss ratio is highest (Scotchmer 2004b, chapter 11). If this ratio is the same everywhere, then it does not matter for total deadweight loss where the profits are earned. However it obviously matters for equity. In any case, there is no policy maker with worldwide authority to make these decisions. Three arrangements have been tried: autarky,

national treatment with independent choices of protections, and harmonized choices.

The first treaties to create reciprocal obligations for national treatment of copyrighted works and patented inventions were the Berne Convention and Paris Convention in the 1880's. It was not for another 100 years that significant strides toward harmonization were made, culminating in the 1994 TRIPs Agreement (Trade Related Aspects of Intellectual Property). In the meantime, the treaties, which had begun with about a dozen members, had grown to about 140 member states.

By autarky, we mean that each country protects only its own inventors. Autarky was the norm prior to the treaties of the 1880's. The main problem with autarky is that the market in any small country may be too small to cover the costs of innovations. If not, however, autarky can be a good system. Because inventors are protected where they are domiciled, and not elsewhere, inventors in different countries create reciprocal externalities. If the countries are more or less commensurate in size, these externalities more or less offset each other. On the other hand, autarky may not provide enough incentive. Reciprocal national treatment is a solution.

However, with national treatment, the fair solution where each country protects its own innovators is no longer possible. A jurisdiction can either protect a subject matter for both domestic and foreign inventors, or it can free ride, letting its inventors be rewarded abroad. There is no intermediate possibility (Scotchmer 2004a). For a subject matter that a country chooses not to protect, its own consumers get the benefit of competitive supply, not only of its own inventions, but also of inventions made abroad. In the meantime, its own inventors collect profit in foreign markets. Free riding eventually led to the harmonization effort of TRIPS.

Suppose then, that the jurisdictions embark on a harmonization effort to coordinate their protections. One possibility is that they will simply harmonize to the efficient regime that a global optimizer would choose (Grossman and Lai 2001). However, since there is no one in charge of a global optimization, it is more likely that individual countries will argue for harmonizations that serve their own interests (Scotchmer 2004a). The harmonization that arises in actuality will be a negotiated solution from these preferred outcomes, and there is no presumption that they will be efficient. These papers conclude that harmonization will generally increase protections and that countries that advocate stronger protection (either more subject matters or longer protection) are those that either have large markets or are more innovative.

In the actual TRIPS negotiation, it was mainly the large, innovative countries like the U.S. that were behind the expansion of protections. This was apparently due to their innovativeness, and not due to their size. In fact there are small, innovative countries like Switzerland that were equally behind the expansion.

Finally, it is worth noticing how the public sector fits into this analysis. Whether domestic R&D is funded by private inventors or public sponsors, domestic discoveries create externalities for foreign users. The externalities are greater with public sponsorship, since foreign users will pay competitive prices rather

than proprietary prices (assuming that intellectual property rights are not asserted by the public sponsor abroad if not asserted at home).

However, in choosing their policies, domestic policy makers are presumably not influenced by the benefits they confer on foreign users. They are more likely to be influenced by the prospect of repatriating some of those benefits as profit. The prospect of profit can shift the political balance in favor of private funding mechanisms, and cause innovative countries to argue for protecting innovations that might otherwise be judged suitable for the public sector (Scotchmer 2004b).

The treaties that have evolved leave scope for national autonomy. The harmonizations generally specify minimum required protections, but do not prohibit stronger ones. But whether stronger domestic protections can survive the international trading arena, especially in the digital age, is unclear. Protected products can typically be stopped at a national border, so that a rightholder can control its distribution domestically, even if not in the international market. However, for other types of intellectual property, such as research tools that can be used abroad to create products patented at home, the absence of foreign protection may undermine domestic protection as well. See Samuelson (2004).

II. Protecting Integrity of the Market

The second principal branch of intellectual property protection – relating to trademarks and unfair competition – focuses upon the quality of information in the marketplace. Quite unlike patent, copyright, and trade secret law, trademark law does not protect innovation or creativity directly. Rather, it aims to protect the integrity of the marketplace by prohibiting the use of marks associated with particular manufacturers in ways that would cause confusion as to the sources of the goods. In so doing, trademark law reduces consumer confusion and enhances incentives for firms to invest in activities (including R&D) that improve brand reputation. This function, however, is part of a larger framework of laws and institutions that regulate the quality of information in the marketplace.

The fact that trademark law does not directly protect technology or works of authorship does not mean that trademarks do not have significant value. The market value of most companies lies predominantly in the goodwill of the brand (e.g., Coca-Cola, Microsoft, Google). Although such goodwill is intertwined with the physical and other intangible assets of the trademark owner, there is little question that trademarks play a critical role in the value of many companies and that licensing of trademarks has become a major business in and of itself.

II.A. The Economic Problem

The efficiency of the marketplace depends critically upon the quality of information available to consumers. In markets in which the quality of goods are uniform or easily inspected at the time of purchase, consumers can determine the attributes themselves and no information problem arises. In many markets, however – such as used automobiles, computers, watches, as well as designer handbags – an information

asymmetry exists: sellers typically have better information about the products or services being offered than buyers can readily inspect (Economides 1998; Akerlof 1970). Unscrupulous sellers will be tempted to make false or misleading product claims or copy the trademark of a rival producer known for superior quality. It is often easier to copy a trademark than to duplicate production techniques, quality assurance programs, and the like. For example, two watches that look the same on the outside may have very different mechanical features, manufacturing quality, and composition of materials used.

Proliferation of unreliable information in the marketplace increases consumers' costs of search and distorts the provision of goods. Consumers will have to spend more time and effort inspecting goods, researching the product market, and actually testing products. Manufacturers will have less incentive to produce quality goods as others will be able to free-ride on such reputations. In markets for products where quality is costly to observe, high quality manufacturers may not be viable in equilibrium without effective mechanisms for policing the source of products and the accuracy of claims regarding unobservable product characteristics.

Several mechanisms are available to provide and regulate market information: (1) deceit and fraud common law causes of action and privately enforced consumer protection statutes; (2) public regulation and public enforcement of unfair competition laws; (3) trademark, false advertising, and deceptive practices/unfair competition laws; (4) industry self-regulation and certification organizations; and (5) consumer information institutions. Since our focus is on intellectual property law, we begin with an overview and analysis of trademark and related private bodies of unfair competition law. In many markets, trademarks provide a simple, quick, and effective means of communicating valuable product information. We conclude by discussing the role of trademark and unfair competition laws within the broader range of mechanisms for protecting the informational integrity of the marketplace.

II.B. An Overview of Trademark Law

Trademarks have been in existence for nearly as long as commerce itself. Once economies progressed to the point where a merchant class specialized in making goods for sale or barter, the people who made and sold clothing and pottery began to "mark" their wares with a word or symbol identifying the maker. These early marks served several functions, including advertising, proof of the sources of goods (of relevance to resolving ownership disputes), and as an indication of the quality of goods. Modern trademark law has retained these functions. Trademarks reduce information and transaction costs in the marketplace by allowing customers to gauge the nature and quality of goods before they purchase them. Consumers rely most on trademarks where it is difficult to inspect a product quickly and cheaply to determine its quality.

Trademark law facilitates and enhances consumer decisions and encourages firms to supply quality products and services by protecting means of designating the source of commercial products and services. Thus, a trademark does not "depend upon novelty, invention, discovery, or any work of the brain. It requires no fancy or imagination, no genius, no laborious thought." (Trade-Mark Cases, 100 U.S. 82, 94

(1879)). Rather, trademark protection is awarded merely to those who were the first to use a distinctive mark in commerce. In trademark parlance, the senior (that is, first) user of a mark may prevent junior (subsequent) users from employing the same or a similar mark in such a manner as to cause a"likelihood of confusion" among consumers as to the source of the goods or services in question.

Traditionally, there has been nothing in trademark law analogous to the desire to encourage invention or creation that underlies patent and copyright law. There is no explicit federal policy to encourage the creation of more trademarks. Rather, the fundamental principles of trademark law have developed from two tort-based causes of action: the tort of misappropriation of the goodwill of the trademark owner and the tort of deception of the consumer. In this sense, trademarks should not be thought of as "property" rights at all. Rather, they are rights which are acquired with use of a trade mark in commerce[6] and derive protection based on the likelihood of indirect harm to potential purchasers of the trademark owner's products.

More recent legislation and several lines of cases, however, have introduced more of a "property" dimension to trademark law. Under the Federal Trademark Dilution Act of 1995, owners of "famous marks" may now prevent others from using their marks even in contexts in which there is no likelihood of consumer confusion. (Several states had previously enacted anti-dilution legislation.) Congress sought to protect such marks from blurring – the erosion of the distinctive quality of a mark through its adoption and use across a variety of product markets unrelated to the one(s) in which it developed fame – and tarnishment – uses that reduce the mark's positive association. The Act exempts uses in comparative advertising, noncommercial settings, and news reporting so as to address First Amendment concerns.

With the rise of the Internet and the establishment of a first-come, first-served registration system for domain names, so-called "cybersquatters" began registering the trademarks of others and either seeking to extort payments in exchange for transfer of the domains, offering such marks to competitors of the trademark holders, or setting up their own websites at these locations as a means of attracting business. The Anticybersquatting Consumer Protection Act, passed in 1999, imposes liability for registering, trafficking in, or using a domain name that is identical or similar to or dilutive of a trademark with bad faith intent to profit.

Chart II summarizes the principal attributes of contemporary trademark law.

Chart II

---

[6] The Trademark Law Revision Act of 1988 changed this general principle in an important respect. Under that Act, it is now possible to register and protect a trademark based on an *intention* to use that mark in commerce within the next three years. 15 U.S.C. § 1051(b) (1988). Filing an Intent to Use application enables the applicant to establish a constructive priority date prior to actual use so long as the applicant proceeds to use the mark in commerce within the prescribed window.

	TRADEMARK
Underlying Theory	perpetual protection for distinctive nonfunctional indications of origin of goods and services in order to protect consumers against confusion in the marketplace
Source of Law	Lanham Act (federal); state statutes; common law (unfair competition)
Subject Matter	trademarks (any designation of origin – including words, slogans, symbols, sounds, color); service marks; certification marks (e.g., Good Housekeeping Deal of Approval); collective marks (e.g., Toy Manufacturers of America); trade dress (product configuration and packaging)
Limitations	*no protection for* functional features, descriptive terms or geographic names (unless they have acquired (secondary) meaning – consumer recognition and association with a specific source), misleading aspects of marks, or names that have become "generic" – become associated with a general product category unconnected with any particular source (e.g., thermos)
Reqts for Protection	priority (first to use in commerce); distinctiveness; acquired (secondary) meaning (for descriptive and geographic marks); use in commerce (minimal)
Process for Obtaining Protection	Use in commerce (or filing of intent to use application (establishing priority date) and subsequent use). Registration is optional, but confers various benefits (establishes *prima facie* evidence of validity (i.e., shifts burden of proof to defendant), constructive knowledge of registration, federal jurisdiction, mark becomes incontestable after 5 years of continuous use (i.e., cannot be found to lack secondary meaning), authorizes treble damages and atty fees, and right to bar imports bearing infringing mark) Examination (prior art search, assessment of requirements for protection) conducted by Trademark Office examiners. Full opposition process for federally registered marks. Maintenance fees for registered marks.
Scope of Protection	protection against uses that create a "likelihood of confusion" among an appreciable number of reasonably prudent consumers; dilution of famous marks; registration, with bad faith intent to profit, of domains names that are confusingly similar to trademarks; false advertising
Duration	perpetual subject to abandonment or loss of distinctiveness (genericide)
Marking	Notice (® for federal registration, ™ **K** otherwise) optional, but confers benefits (burden of proof, remedies)
Rights of Others	truthful reflection of source of product; fair and collateral use (e.g., comment, news reporting, comparative advertising)

	TRADEMARK
Costs of Protection	registration search; marking product (optional); litigation costs
Licensing and Assignment	no naked licenses (owner must monitor licensee); no sales of trademark "in gross" (i.e., without accompanying goodwill of associated manufacturer or service provider); licenses cannot be assigned without licensor consent
Remedies	injunction; accounting for profits; damages (potentially treble); attorney fees (in exceptionable cases); seizure and destruction of infringing goods; criminal prosecution for trafficking in counterfeit goods or services

Trademark law thus consists of two principal branches. The traditional and still most important form of trademark protection provides remedies against the use of trademarks in ways that cause confusion in the marketplace as to the source of goods and services. Passing off or counterfeit goods – the marketing of goods displaying another's mark without authorization – represents the classic and most common example of trademark liability. Anti-dilution protection – a second and more recently developed branch of trademark protection – protects famous marks against some forms of non-confusing uses of trademarks. The economic rationales for these forms of trademark protection differ and hence we take them up separately.

II.C. Confusion-Based Protection

II.C.1. Basic Economics

Economic analysis of seller-provided information (advertising and trademarks) grows out of several fields of economic research and has evolved significantly over the past century. Early industrial organization economists were critical of advertising (and hence marking) on the ground that such activities "unnaturally" stimulated demand, thereby fostering and perpetuating oligopoly through "artificial" product differentiation. Reflecting his interest in monopolistic competition, Chamberlin (1933) viewed trademarks as a means for reinforcing monopoly power by differentiating products and thereby excluding others from using the differentiating characteristic, even if only a mark. By generating a downward sloping demand curve for its brand, trademark owners could under this view generate monopoly rents (and resulting deadweight loss) (Robinson 1933, p.89; Comanor and Wilson 1974; McClure 1979, 1996; Lunney 1999). Drawing upon this literature, Brown (1948) tied the analysis of trade symbols to the larger context of commercial advertising, which he considered to serve both useful (informative) and wasteful (persuasive – intended to suggest that one product is superior to a similar if not identical alternative) ends. This led him to approach trademark protection with ambivalence and caution.

The emergence of the modern information economics literature in the 1960s and 1970s offered a

more productive view of the role of advertising in markets (Stigler 1961; Nelson 1970,1974, 1975; Hirshleifer 1973; Nagle 1981). Trademarks, as a concise and unequivocal indicator of the source (e.g., Intel) and nature (e.g., Pentium) of particular goods, counteract the "market for lemons" problem (Ackerlof 1970) by communicating to consumers the enterprise which is responsible for the goods and, in some cases, the specifications of the goods (Landes and Posner 2003) . The brand name Coca-Cola, for example, informs the consumer of the maker of the soft drink beverage as well as the taste that they can expect. If the product lives up to or exceeds expectations, then the trademark owner gains a loyal customer who will be willing to pay a premium in future transactions; if the product disappoints, then the trademark owner will have more difficulty making future sales to that consumer (or will have to offer a discount to attract their business). In this way, trademarks implicitly communicate unobservable characteristics about the quality of branded products, thereby fostering incentives for firms to invest in product quality, even when such attributes are not directly observable prior to a purchasing decision (Klein and Leffler 1981; Hirshleifer 1973; Shapiro 1982, 1983; Milgrom and Roberts 1986; Economides 1998). Sellers who enter the high quality segment of the market must initially invest in building a strong reputation. Only after consumers become acquainted with the attributes of their brand can they recoup these costs. In equilibrium, therefore, high quality items sell for a premium above their costs of production to compensate for the initial investment in reputation (Shapiro 1983). Trademarks also facilitate efficient new business models, such as franchising, which generate economies of scale and scope in marketing and facilitate rapid business diffusion across vast geographic areas (Wilkins 1992; Williamson 1986).

The marking of products also creates incentives for disreputable sellers to pass off their own wares as the goods of better respected manufacturers. Trademark law (as well as false advertising and unfair competition laws more generally) harnesses the incentives of sellers in the marketplace to police the use of marks and advertising claims of competitors. Sellers often have the best information about the quality of products in the marketplace; they also have a direct stake in preventing competitors from free riding on their brand, reputation, and consumer loyalty. By creating private causes of action, trademark and false advertising law take advantage of this informational base and incentive structure as well as the vast decentralized enforcement resources of trademark owners to regulate the informational marketplace, effectively in the name of consumers.

Under this now widely accepted view of consumer information economics, trademarks economize on consumer search costs (McClure 1996; Kratzke 1991; Economides 1998; Landes and Posner 1987). Consumers benefit from concise and effective designations of the source of products. For example, consumers can quickly assess the attributes of a computer made by Sony featuring an Intel Pentium Processor and Microsoft's XP Operating System. If such trademarks were not available or could not be relied upon, the consumer would have to incur substantial additional costs in shopping for a computer. The ability to establish and maintain reliable trademarks reinforces firms' desire to develop and maintain consistent quality standards. It also fosters competition among firms over a wide quality and variety spectrum (Economides 1998).

In general, consumers distinguish among three types of product features: search attributes, such as

color and price, which can be inspected prior to purchase; experience attributes, such as taste, which can only be verified through use of the product (typically after purchase); and credence attributes, such as durability, which can only be verified over time (or through the use of surrogate sources of information – e.g., Consumer Reports) (Nelson 1974; Darby and Karni 1973; Bone 2004). Brands signal experience and credence attributes. In an empirical study of branded and unbranded gasoline service stations, Png and Reitman (1995) found that branded dealers were more likely to carry products for which quality was more difficult to verify and to serve customers who placed a higher value on search.

Some trademarks also serve a more ambiguous function: signaling status or identity for some consumers. Some have referred to such commodities as "Veblen" goods, reflecting Thorstein Veblen's theory of conspicuous consumption. This theory posits that demand for status goods rise with increases in price (Leibenstein 1950; Veblen 1899). Purchasers of such goods may be interested in being associated with a particular brand – such as a Rolex watch, a t-shirt with the name and colors of a particular university, or a corporate brand – possibly apart from whether it is authentic or the quality associated with the authentic good (Kozinski 1993; Dreyfuss 1990; Higgins and Rubin 1986). Some purchasers of such goods may well prefer a less expensive, counterfeit version. They presumably would not be confused when purchasing such goods (e.g., a Rolex watch sold on a street corner for $10).

The marketing of less expensive, lower quality imitations of status goods creates the possibility of separate harm to the sellers and purchasers of authentic goods. The availability of counterfeit articles could well divert some consumers who would otherwise purchase the authentic article, although this effect is likely to be relatively small due to the large price differential and the availability of the authentic goods for those who are interested. The lower quality of the counterfeit goods could, however, erode the goodwill associated with the authentic manufacturer through post-sale confusion – on-lookers who mistake the shoddier counterfeit good for the authentic good and are thereby less inclined to purchase the authentic version, thereby reducing sales by the trademark owner. In addition, due to the proliferation of non-easily recognized "fakes," prior and potential purchasers of the authentic "status" goods may be less interested in owning a much less rare commodity. The value of ownership may be sullied. In essence, status goods exhibit a negative network externality, whereby proliferation of such goods erodes the value to prior purchasers (Higgins and Rubin 1986; Kozinski 1993; Dogan and Lemley 2004). The significance of these harms is considered speculative. See Lunney (1999) (questioning the economic basis for protecting status values).

Notwithstanding the general benefits afforded by trademarks, such protection entails several types of costs. Protection of descriptive terms as trademark can increase search costs and impair competition by raising the marketing costs of competitors. For example, if a cookie manufacturer were to obtain a trademark on the word "cookie," then other companies interested in selling cookies would have a much more difficult time communicating the nature of their goods to consumers. If, however, the trademark was to "Mrs. Fields Cookies" and any protection for "cookies" was disclaimed, then potential competitors would be able to describe their products in the most easily recognized manner and would be able to develop their own marks – such as "ACME Cookies." At a minimum, trademark protection for descriptive

terms significantly reduces the effective range of terms that may be used others.

A complicating factor in the protection of trademarks is the endogeneity of the usage and meaning of terms and symbols over time. Even a distinctive term can become "generic" (common) if consumers come to associate marks with a particular product (as opposed to its manufacturer). The evolution of the use of the term "thermos" illustrates this phenomenon. At the turn of the twentieth century, the original manufacturer of vacuum-insulated flasks selected the term "Thermos" – derived from the word "therme" meaning "heat" – to brand its product. At the time that it was selected (in effect, coined), Thermos was distinctive and not associated with any particular product. The American Thermos Bottle Company, which acquired the U.S. patent rights for this technology, undertook advertising and educational campaigns that tended to make "thermos" a generic term descriptive of vacuum-insulated flasks rather than of its origin. After the patents expired, other manufacturers began using this term to describe their own vacumm-insulated flask products. As we will discuss further below, use of the term became generic in the eyes of consumers, and hence the law, and the original manufacturer of the product (and developer of the mark) lost trademark protection. King-Seeley Thermos Co. v. Aladdin Industries, Inc., 321 F.2d 577 (2d Cir. 1963)

More generally, trademark protection for descriptive terms can impede competition. Gaining control over the most effective term for describing a product raises the costs of potential competitors seeking to sell in that marketplace. By not being able to use a term or means of communication most easily understood by the consuming public, the entrant must bear higher marketing costs. Limitations on the use of trademarked terms for purposes of comparative advertising would also impede vigorous competition. Trademark law is least problematic at its traditional core: protecting inherently distinctive (i.e., non-descriptive) source identifying marks against directly competing uses that confuse consumers. The expansion of trademark protection to encompass non-competing products, dilution (non-confusing uses of famous marks), product configuration and packaging (trade dress), merchandising of trademarks (mere sponsorship), post-sale confusion, and more distant reputation zones have increased the tension between trademark protection on the one hand and competition and innovation policy on the other (Lemley 1999; Lunney 1999; Bone 2004).

Trademark protection can also interfere with both communicative and creative expression. Broad exclusive trademark rights would limit the ability of others (including non-competitors) to comment on and poke fun at trademarks and their owners. As we will see below, various doctrines limit such adverse effects. But as trademark protection has expanded beyond the traditional core – for example, to encompass a broad conception of connection to, sponsorship, and affiliation with a trademark owner – it becomes more difficult to assess the boundaries, leading film and television production companies, for example, to tread carefully (and increasingly incur the costs of licensing transactions) in the use of

trademarks in their works.[7]

As with other modes of intellectual property, trademark protection also involves administrative and maintenance costs. Although the costs of acquiring trademark protection is relatively low, mark owners must police their marks to prevent use of the marks without authorization and supervise licensees to ensure that quality standards are maintained. As a mark enters common parlance and becomes associated in the minds of consumers with a general product category as opposed to the manufacturing source – as in the Thermos example – the owner must invest in advertising to clarify that the mark is associated with a particular supplier in order to prevent "genericide" – the death of a trademark due its becoming generic. For many years, Xerox spent large sums on advertising to discourage generic usage of the term "xerox" as noun or verb for photocopying. Google faces a similar exposure today.

II.C.2. Policy Levers

As summarized in Section II.B, trademark law has evolved into a complex system of administrative rules and judicial doctrines. Such rules and doctrines can be seen as a series of policy levers that may be crafted by legislators and courts and administered by the Trademark Office (through a registration system) and the courts. The economic analysis of particular trademark doctrines (policy levers) focuses on balancing the salutary effects of trademarks with the various costs: constraints on the availability of language and symbols to economize on consumer search, protection costs (administrative, maintenance, and enforcement), anticompetitive effects, and limitations on the freedom of creative and communicative expression. We examine the administration of the trademark system in the following section.

Before turning to the analysis of specific rules governing trademark law, it is useful to re-emphasize one of the observations made at the outset: Although the "real property" metaphor provides some useful insights for understanding intellectual property law, simply extrapolating from economic analysis of real property overlooks important distinctions. Trademarks serve a different set of purposes than real property law and operate in a much more diffuse environment (control of words and symbols). Perhaps most significantly, trademark law standards use the public's perception of the meaning of words and symbols as the touchstone for determining the rights (validity, breadth, infringement, and duration) and limitations of trademark owners. As one jurist (Kozinski 1993) has aptly noted:

---

[7] As a reflection of the growing importance of brand exposure and image creation, film and television production companies view product placements as an advertising revenue source. Nonetheless, to the extent that trademark law requires licensing of trademarks for such works, there is a cost (and potential distortion) imposed on the creative process. The fact that some trademark owners are willing to pay for exposure alleviates but does not eliminate this concern. Although some mark owners would compete for product placements even if the trademark licensing was not required, creators are hampered to the extent that they prefer to use marks that are not available for licensing.

> Words and images do not worm their way into our discourse by accident; they're generally thrust there by well-orchestrated campaigns intended to burn them into our collective consciousness. Having embarked on that endeavor, the originator of the symbol necessarily – and justly – must give up some measure of control. The originator must understand that the mark or symbol or image is no longer entirely its own, and that in some sense it also belongs to all those other minds who have received and integrated it. This does not imply a total loss of control, however, only that the public's right to make use of the word or image must be considered in the balance as we decide what rights the owner is entitled to assert.

II.C.2.i. <u>Threshold Requirements</u>

The three principal requirements for establishing trademark protection – distinctiveness, priority, and use in commerce – can all be understood to reflect economic considerations.

Distinctiveness. Trademark law affords protection to the first enterprise to use fanciful ("Kodak" (photographic products)), arbitrary ("Apple" (computers)), and suggestive (requiring a leap of imagination by the consumer, such "Chicken of the Sea" (tuna) or "cyclone" (braided wire fencing)) marks as soon as they are used in commerce, whether or not they are registered. (We discuss registration of trademarks in section II.E below.) Descriptive terms (such as "Digital" for computers), surnames (McDonalds), and geographical designations (e.g., New York Times) are protectable only upon acquiring secondary meaning (denoting a single seller or source) in the minds of a substantial portion of the relevant consumer marketplace. Generic terms are ineligible for protection, reflecting the idea that search costs to consumers would be greater if new entrants could not use the common meanings to label and advertise their products.

Affording automatic protection to inherently distinctive marks (fanciful, arbitrary, and suggestive terms) rather than awaiting proof of secondary meaning can be justified on process, error cost, and predictability grounds (Bone 2004; Denicola 1999; Landes and Posner 1987) (defending categorization of marks versus case-by-case balancing as saving administrative and dispute resolution costs). Proving secondary meaning requires relatively time consuming and costly consumer surveys. Moreover, inherently distinctive terms are plentiful in supply (an infinite number of fanciful and arbitrary terms are available) and hence potential entrants would not be constrained in any significant way by the removal of such terms from the universe of potential inherently distinctive marks. Providing automatic protection for such terms reduces the costs of entry and enables firms to make investments in developing brand equity secure in the knowledge that their mark will be valid (assuming priority of use, which can be assessed through a relatively quick and inexpensive trademark search).

By contrast, affording protection to descriptive terms (including geographic designations and surnames) before such terms became associated with a particular source would raise consumer search costs and impose undue barriers to entry by competitors. Effective descriptive terms are limited in supply (Carter 1990) and therefore any restriction on the use of such terms by consumers and potential entrants

raises search costs. As Burge (1984, p. 126) notes, "suggestive and descriptive marks tend to be preferred by advertising people because these marks are thought to enhance initial product salability." But once a descriptive term becomes associated with a source – e.g., New York Times, "Bed & Bath" (home products store), "Chap-Stick" (lip balm), "McDonalds" (fast food) – allowing entrants to adopt identical or similar designations risks confusion in the marketplace. Trademark law balances these costs by delaying the time at which such marks can be protected until sufficient consumer recognition has been achieved.

In an interesting judicial use of the distinctiveness threshold as a policy lever, the Supreme Court has ruled that product configurations (as opposed to mere product packaging) can never be inherently distinctive – i.e., acquired meaning must always be established in order to obtain protection. (Wal-Mart Stores, Inc. v. Samara Brothers, Inc., 529 U.S. 205 (2000)) In so doing, the Court expressly used this tool to encourage competition in product markets by requiring express proof that product configurations, even if arbitrary, functioning as trademarks have acquired meaning in the minds of consumer before receiving protection. Note that this requirement is in addition to the separate rule, which we discuss below, that functional elements of products may not be protected as trademarks.

Priority and Use in Commerce. Although registration of trademarks is optional (see section II.E relating to administration), trademark rights are accorded to the first user of a mark in commerce. Such a rule discourages rent seeking, such as the stockpiling of names for subsequent resale or the locking up of a large segment of the useful semiotic domain. Landes & Posner (1987); Carter (1990). Pure registration systems – such as the Japanese trademark system and the domain name registration for the Internet – have produced rent seeking behavior resulting in the warehousing of terms, making it more costly for others to enter markets (Landes and Posner 2003, pp. 179-80). The use requirement also serves a notice function.

The use requirement can be criticized on economic grounds as being both too lax and too strict. Under current rules, even token use suffices to establish priority and with registration merely optional, the notice function may not be adequately served and banking of potential terms is still possible at relatively modest cost. On the other hand, requiring actual use exposes companies planning large product introductions some risk that their mark could be preempted on the eve of the announcement. Such risk adds needless uncertainty. The introduction of the Intent to Use application process addressed this problem by enabling companies to obtain a certain priority date for a trade name in advance of use in commerce so long as use follows within a six month period (with extension possible for a total of up to three years). Carter (1990) has expressed concern that this system provides undue potential for anticompetitive warehousing behavior and calls for imposition of penalties where it appears that a registrant has filed numerous intent-to-use applications without a serious intention to use such marks in commerce.

II.C.2.ii. Duration

Given the primary purpose of trademark law of reducing consumer search costs, there is a strong justification for trademark protection lasting as long as a mark represents a reliable designation of source

of goods and services. Due to the infinite availability of arbitrary and fanciful marks, perpetual protection for trademarks does not prevent others from entering the market. And to the extent consumers connect a descriptive term to a particular source, confusion would result from expiration of that mark while the developer of the mark continues to operate under that name or logo. Unlike with copyrights or patents, there is no concern about perpetual duration hindering cumulative innovation by others because trademark protection does not extend to functionality or creativity per se. Limiting doctrines allow others to make some expressive usage of trademarks – e.g., for comparative advertising and social commentary.

Trademarks will lose protection, however, if they become generic. At the point at which an appreciable number of consumer associate a mark with a product category as opposed to a source, allowing one manufacturer exclusive rights to the mark raises consumer search costs and the marketing costs of competitors. We discuss genericide more fully in the section on rights of others (and defenses).

Trademarks will also lose protection through voluntary abandonment or dramatic changes in product quality. Abandonment occurs through a trademark owner exiting a business (without transferring the mark along with the associated goodwill to another firm). Whereas trademark owners may evolve their products and product quality, they may not so dramatically change the quality or nature of a product sold under the mark (without appropriate warnings) as to constitute fraud upon the consuming public. This prohibition discourages deceptive opportunism. Other consumer protection statutes – protecting against deception and fraud – potentially address this form of deception as well.

Once a mark is abandoned, it becomes fair game for new entrants or existing manufacturers. Such a doctrine could cause confusion in the marketplace to the extent that a new user of a recently abandoned trademark offers goods of different quality than the prior mark owner. For this reason, trademark law requires new users of abandoned marks to take reasonable precautions to prevent confusion until such time as the association with the prior supplier has faded from the public's lexicon. Other consumer protection statutes may also discourage deceptive practices that may occur following a change in trademark ownership.

II.C.2.iii. <u>Ownership and Transfer Rules</u>

The assignment and licensing of trademarks presents a problem for maintaining the integrity of quality standards, and hence the expectations of consumers. To allow free alienability of trademarks – as is permitted for conventional forms of property as well as patents and copyrights – could jeopardize the quality assurance implicit within the nature of a trademark. For this reason, U.S. trademark law prohibits marks to be assigned "in gross" – i.e., without the good will underlying the mark (including the right to produce the goods sold under the mark) – or licensed without ongoing supervision by the trademark owner. Such restrictions on alienability discourage "end game" opportunism – selling the mark at a premium upon exiting the trade (or entering bankruptcy) to a company that intends to reap a premium on the sale of shoddy goods.

The concern with this opportunism scenario appears to be overblown. Cf. McCarthy (2004 §18.10); McDade (1998). If a mark has value in the market, then the assignee/purchaser of the mark will jeopardize the long term value by lowering quality standards. Moreover, as noted above, radical changes in the quality of goods sold under a mark could result in abandonment. In fact, most other nations permit assignments of trademarks in gross,[8] suggesting that a rule regulating transfers may be unnecessary. Similarly in the licensing context, trademark licensors ultimately bear the costs of erosion in brand equity and therefore have strong incentives to put in place efficient supervisory systems to maintain or enhance brand equity without additional legal constraints – i.e., it is not at all clear that there is an externality. More generally, significant changes in products or services following assignment or under licensing agreements could result in liability for fraud or deceit or exposure under publicly enforced consumer protection statutes.

II.C.2.iv. Breadth and Infringement Analysis

As we saw earlier, patent and copyright law afford exclusive rights for purposes of promoting investment in the development of new works. Hence, infringement analysis focuses on a comparison of the elements of the protected work (the patent claims or the copyrighted book, musical composition, or other work) and the allegedly infringing work. By contrast, trademark law does not grant exclusive rights but rather limits protection to the purpose of protecting consumers against confusion as to the source goods or services. The touchstone for trademark protection is *consumer perception* – whether an appreciable number of reasonably prudent consumers perceive the defendant's product or services to be sponsored by, affiliated with, or otherwise connected to the trademark owner. This standard tailors the scope of trademark protection to the consumer search cost rationale, leaving freedom for competitors and others to use marks in ways that are not likely to cause consumer confusion.

The shift in focus from comparing protected and allegedly infringing works (irrespective of locus of use (so long as it is in the United States for patents) and product market) as in patent and copyright law to assessing consumer confusion requires a multi-dimensional framework. The scope of trademark protection can be thought of spatially along semiotic (linguistic and symbolic), product market, and geographic dimensions. To illustrate this framework, consider the trademark of the ACME Bread Company in Berkeley, California. Along the semiotic dimension, we can imagine a spectrum of marks from ACM to ECME to ACMF which bear some resemblance to ACME – all selling bread in Berkeley community. Along the product dimension, we can envision different companies in Berkeley operating under the ACME selling baked goods (pastries as well as bread), groceries, office furniture, as well as fishing supplies. Along the geographic dimension, we can imagine ACME Bread Companies (with different owners) in neighboring Oakland, California, St. Louis, Missouri, or Atlanta, Georgia.

---

[8] Article 21 of the Trade-Related Aspects of International Trade (TRIPS) agreement permits the owner of a registered trademark to assign the mark without transferring the "business" associated with the mark.

Which of these competing businesses, if any – from the ECME Bread Company in Berkeley to the ACME Fishing Supply Company in Berkeley to the ACME Bread Company in Atlanta Georgia – infringes the trademark owned by the ACME Bread Company in Berkeley[9]? Under early trademark law, protection was limited to goods of the same descriptive class – i.e., directly competing goods. Since 1946, however, protection has encompassed confusion as to origin, sponsorship, approval, and connection, whether or not goods are in direct competition. Thus, modern trademark law does not provide categorical answers to the scope of protection. Rather it determines liability and hence scope of protection on the basis of a comprehensive, fact-intensive examination of a wide range of factors bearing on the perceptions of reasonably prudent consumers in the relevant marketplace. Under modern tests, courts look to the following non-exhaustive list of factors:

- characteristics of the trademark (inherent distinctiveness, acquired meaning)
- characteristics of the allegedly infringing mark (similarity to the plaintiff's mark)
- marketplace considerations:
    - strength of the senior user's mark
    - nature of the product market (low cost versus high cost products; care exercised by consumers)
    - proximity of the goods
    - likelihood of expansion of either party into the other's product market
    - channels of trade and methods of distribution
    - advertising and promotion
    - nature and sophistication of consumers
- evidence of actual consumer confusion (e.g., misdirected service calls by consumers, testimonial evidence, surveys)
- evidence of bad faith (e.g., intentional copying of mark) by the junior user.

Over the past several decades, the effective scope of trademark protection has expanded to encompass promotional goods (enabling universities, sports teams, and corporate sponsors to enjoin clothing manufacturers from selling t-shirts emblazoned with trademarks without authorization), initial interest confusion (whereby consumers may be only initially confused as to source, but not at the time of purchase), post-sale confusion, and trade dress (product configuration and packaging). This has led some commentators to believe that trademark law has gone beyond the boundaries necessary to optimize consumer search costs and increasingly threatens competition (Lunney 1999; Lemley 1999; Bone 2004 (suggesting that disclaimers ought to be more readily credited, especially in the case of promotional goods, and that trade dress protection be abolished); Dogan and Lemley 2004a). The application of the trademark law to Internet activities has continued this trend, with courts focusing on a rather limited range

[9] We are assuming here that the ACME Bread Company in Berkeley has priority over these other enterprises – i.e., it was the first to use the ACME trademark in commerce. As an arbitrary mark, ACME would have received protection upon initial use.

of factors for determining infringement (similarity of marks and relatedness of goods) and finding infringement on the basis of initial interest confusion readily, notwithstanding the rather modest costs of redirecting Internet searches (Dogan and Lemley 2004b).

II.C.2.v. Breadth and the Rights of Others

The other principal policy levers affecting the scope of trademark protection relate to exceptions and defenses to liability. Several doctrines limit the scope of protection in order to promote competition, innovation, and freedom of communicative and creative expression.

Functionality. The expansion of trademark law to encompass product configurations bought trademark's regime of perpetual protection on the basis of relatively low validity requirements potentially into conflict with patent law's exacting validity requirements and limited duration. Without appropriate limitations, trademark law could protect sub-patentable technologies as well as extend protection for patented technologies beyond the expiration of the patents. To avoid upsetting the balances of the patent system, courts developed a rule that functional product features – defined as those elements that are essential to the use or purpose of a product that affect its cost or quality – cannot serve as a trademark. The aesthetic functionality doctrine applies a comparable channeling principle with regard to copyrightable product elements, such as pottery and silverware designs.

Parchomovsky and Siegelman (2002) show that the ability to protect distinctive functional features of patented technologies under trademark law beyond the expiration of a patent induces the patentee to moderate its pricing during the term of the patent in order to foster brand loyalty. This effect offsets to some extent the static deadweight loss of patent protection in anticompetitive effects of allowing perpetual trademark protection for functional product features. As they note, however, the optimal level of trademark leveraging will vary across patented technologies and policymakers will often lack the information needed to tailor the balance optimally.

Genericide. Consumer perception of the meaning of words and symbols can change over time, sometimes resulting in trademarks drifting from designating the source of a good to becoming a generic means of describing a category of products. Thermos, yo-yo, escalator, refrigerator, and aspirin have all made this transition. Once a substantial percentage of consumers come to treat a term as a generic product category rather than a brand designation, consumer search costs are raised (costs of having to communicate around a well-known, but protected, term) and undue market power conferred by allowing but one manufacturer to control the use of the term. For example, if the term "plexiglass" could not be freely used, competitors would have to resort to rather prolix expressions such as "unbreakable clear plastic sheets that function as glass" in order to describe their products. (Merges, Menell, and Lemley 2003, p.685; Landes and Posner 1987, p. 292). Such a mouth-full raises the costs of advertising and would likely engender significant consumer confusion as a result of consumers inferring that the purveyor must not mean "plexiglass" because that would obviously have been easier to convey.

In recognition of this phenomenon, the genericide doctrine strips trademark protection from terms whose primary significance in the minds of the consuming public signifies a general product category rather than a particular product sold by a manufacturer, even if the originator of the term put substantial effort into creating it and encouraging its use.. Although commentators differ over the appropriate standard for determining when a term has become generic – with some favoring expressly economic formulations (Folsom and Teply 1980, 1988a, 1988b; Coverdale 1984 (advocating use of antitrust-type cross-elasticities of demand approach for determining the degree of substitutability among terms); Landes and Posner 1987 (proposing cost-benefit test)), and others favoring more conventional formulations on practical grounds (Swann 1980; Swann and Palladino 1988; Oddi 1988) – there is general consensus that the genericide principle economizes consumer search costs.

Fair Use and Nominative Use. Notwithstanding the protectability of descriptive marks, geographic designations, and personal names that acquire secondary meaning, trademark law balances the resulting constraint on the use of commonly understood terms by allowing competitors to make "fair use" of the protected terms to describe their own goods or services, their geographic origin, or the names of people involved in their own business. The nominative use doctrine allows others to use a protected mark to describe the mark owner's product, as, for example, in comparative advertising or in non-trademark usages. Allowing such uses reduces consumer search costs by making it easier to communicate relevant information to consumers, thereby promoting free competition and use of language.

Use in Commerce and Indirect Liability. Trademark liability can only be imposed where a competitor uses a mark in advertising or commerce "causing the public to see the protected mark and associate the infringer's goods or services with those of the mark holder." (DaimlerChrysler AG v, Bloom, 315 F.3d 932, 939 (8th Cir. 2003)) This doctrine has come under scrutiny with the emergence of Internet search technologies and business models. Website developers often insert hidden codes, such as metatags, that are used by web search engines to index web sites based on relevance of search queries. The question arises whether the placement of a competitor's trademark into a metatag constitutes a use in commerce. Similarly, search engine companies, such as Yahoo and Google, that deliver sponsored advertisements based on search queries derive a substantial portion of their revenue by selling keyword advertising placements. Does the sale of such keyword advertising placements constitute use of the terms in commerce? Such keyword advertising placements can be seen as a form of free-riding, seeking to divert web surfers looking for links to an established trademark; they can also be viewed as general inquiry into the commercial marketplace – a proxy for range of relevant sites. Dogan & Lemley (2004b) advocate tying the liability for trademark infringement to the search cost rationale – only those who are using the mark to advertise their own wares or services have the motive and opportunity to interfere with the clarity of the mark's meaning in conveying production information to consumers. Hence, they would permit search engines to escape liability.

Freedom of Expression. Courts recognize a First Amendment defense to trademark infringement where another seeks to use a mark to communicate ideas or express points of view. One court recently held that a song entitled "Barbie Girl," that poked fun at the Mattel Corporation's doll of the same name,

did not infringe the trademark. (Mattel, Inc. v. MCA Records, 296 F.3d 894 (9th Cir. 2002).

II.C.2.vi. Remedies

Courts award injunctive relief as a matter of course upon a showing of likelihood of consumer confusion. There is no requirement of actual confusion (Bone 2004 (suggesting that such rules may be justified by process cost considerations)). Monetary relief (actual damages, lost profits, the defendant's profits attributable to the infringement, punitive damages in cases of willful infringement, and attorney fees) is also available, although damages are often difficult to quantify. In 1984, strong federal criminal sanctions along with public enforcement was put in place in order to stem a growing tide of international trademark counterfeiting. Due to the unique federal role in and resources for policing international borders, public enforcement of trademark counterfeiting offered significant economies of scale and scope over private enforcement by individual trademark owners.

II.D. Dilution-Based Protection

II.D.1. Basic Economics

The economic rationales for dilution grow out of the same considerations applicable to confusion-based trademark liability – reducing consumer search costs and fostering investment in product quality and brand equity – although the concerns are somewhat more attenuated. The principal problem to which dilution protection is addressed concerns blurring (loss of distinctiveness) of brand identity (Schechter 1927). As consumers develop their mental lexicon of brands, they associate both specific products and general attributes with particular trademarks. For example, Rolls Royce connotes both the source of a luxury automobile as well as brand of uncompromising quality and ornate styling (as well as high cost). If another company were to introduce Rolls Royce candy bars, it is unlikely that many (if any) consumers would believe that the automobile manufacturer was the source. Whether intended or not, the candy company may benefit from the particular general attributes that the consuming public associates with the Rolls Royce brand. They may also gain some "status" equity to the extent that consumers value the signal associated with a mark. Thus, adopting the Rolls Royce name enables the newcomer some ability to free-ride on the general brand reputation of the famous trademark owner.

Such use, however, would impose some costs on consumers and the famous trademark owner. As this new use of the Rolls Royce term gained popularity, the association between the mark and a particular source would become blurred. Furthermore, as more companies in unrelated markets adopt this moniker – Rolls Royce tennis racquets, Rolls Royce landscaping, Rolls Royce tacos – the distinctive quality of the mark would become further eroded. Over time, consumers would lose the non-product specific identity (i.e., Rolls Royce as a brand of uncompromising quality and ornate styling) that the original Rolls Royce mark once evoked. This raises consumers' search costs: consumers' mental lexicon has become more difficult to parse. A lack of protection against trademark dilution could weaken the incentives of suppliers to invest in and maintain their brand equity, although this effect is likely to be quite attenuated in

most circumstances. Owners of famous marks have strong incentives to maintain and enhance their brand equity even without formal protection against dilution. Nonetheless, the full benefits of their investment are not internalized due to potential free-riding by others. Protection against blurring of famous marks has some parallels to the prospect and rent dissipation theories of intellectual property protection (Kitch 1977; Grady and Alexander 1992). Upon establishing a famous mark, the owner obtains broad scope for further developing the intellectual property right.

A second form of dilution relates to the tarnishment of a well-known brand. If the maker of pornographic films were to sell their movies under the brand "Disney," it is unlikely that consumers would believe that the Disney Corporation, famous for family oriented entertainment, was the manufacturer of such unwholesome products. Nonetheless, consumers' shopping lexicon would arguably be distorted because the Disney name would trigger associations with both family oriented content and smut. Such a negative association could well injure the Disney Corporation's brand equity. As with blurring, tarnishment interferes with established associations. Perhaps even more so than blurring, it undermines brand equity.

Anti-dilution protection prevents this erosion of the distinctive quality of a mark by prohibiting famous marks from being used by others – even in unrelated product markets and in non-confusing ways. The Rolls Royce automobile manufacturing company can preclude the marketing of Rolls Royce candies without its authorization. Disney can prevent pornographers from adopting the Disney name. This preserves distinctive brands and affords the owners exclusive rights to carry their brand names into wholly new markets (or not). We see examples of such brand migration in many markets. Sony Corporation, for example, which honed its reputation in the consumer electronics marketplace, has now developed products in the sound recording and motion picture marketplaces.[10] Cross-branding, such as the marketing of a Barbie doll adorned with Coca-Cola's logo and a distinctive red ensemble, is also increasingly common.

Expanding trademark law to protect against dilution of marks can impose several costs. Dilution law could operate to keep otherwise generic terms from being available to all. This marginally increases consumer search costs and raises the marketing costs of other companies. In effect, dilution law could conceivably constrain use of the language. Beyond this concern, to the extent that transaction costs discourage some valuable licensing of "dilutive" uses, protection against dilution may well be inefficient. For example, there may be parodic uses of famous trademark that might well be valued by consumers but would not be licensed, notwithstanding minimal effects of brand equity. More generally, broad protection against dilution could chill news reporting (e.g., stories exposing negative information about companies with famous trademarks), comparative advertising, and expressive creativity. Constitutional safeguards of freedom of expression as well as exceptions to trademark dilution liability seek to balance these competing considerations. As suggested above, it is not at all clear that dilution poses significant harm. The economic

---

[10] The expansion of traditional trademark protection has, to some extent, afforded protection against diluting uses of trademarks. Courts consider the likelihood that a trademark owner would expand into a new market in determining infringement.

benefits are attenuated in most circumstances and traditional trademark law addresses the most significant concerns – where likely consumer confusion can be demonstrated. Therefore, dilution protection may be of questionable net value (Port 1994; Lunney 1999).

II.D.2. Policy Levers

Since trademark dilution is an outgrowth of traditional trademark protection, the same validity, duration, and ownership and transfer rules discussed above apply to protection against dilution. The principal critical levers for this cause of action are the additional thresholds (fame and possibly inherent distinctiveness), the standard for determining infringement, and limitations on liability.

Additional Threshold Requirements. Unlike confusion-based trademark protection, federal dilution protection is available only to famous marks. Thus, the threshold for determining fame serves as a policy lever for determining the availability of dilution protection. As noted earlier, some commentators have expressed greater concern about the need for and adverse effects of dilution protection. They would limit dilution protection to the best known national brands. Some courts, however, have recognized that fame can exist even in niche markets. (Times Mirror Magazines, Inc. v. Las Vegas Sports News, L.L.C., 212 F.3d 157 (3rd Cir. 2000). The earliest and most ardent academic proponent of the dilution cause of action suggested that protection should be confined to inherently distinctive marks – e.g., Kodak – and not be available to descriptive marks (including geographical designations and surnames) that have acquired distinctiveness and fame, such as United Airlines and McDonalds (Schechter 1927). Such a constraint on the range of marks eligible for protection allows descriptive terms to remain more available for use in other markets.

Infringement Standard. As we saw in the context of general confusion-based trademark liability, a mark owner need only establish a "likelihood of confusion" in order to prove infringement. Perhaps as a means of cabining dilution protection, the Supreme Court has interpreted the federal dilution statute to require proof of adverse impact on the famous mark as a result of blurring – which will typically require consumer surveys (unless the junior mark is identical to the senior mark). (Moseley v. V. Secret Catalogue, Inc., 537 U.S. 418 (2003)) This relatively high threshold limits the availability of the dilution remedy.

Limitations on Liability. The scope of dilution protection is quite broad, affording the owner of a famous mark broad discretion to enter (or preclude others from entering) any market under the famous mark. The right also protects the owner against tarnishment. Due to this vast potential scope of protection, Congress included several exceptions for comparative advertising, noncommercial uses (e.g., product reviews), and news reporting so as to balance competitive considerations and to address First Amendment concerns.

Remedies. Dilution protection envisions injunctive relief as the principal remedy, although damages and profits are available upon a showing that the defendant "wilfully intended to trade on the owner's

reputation or to cause dilution of the famous mark."

II.E. Administration

Trademarks may be secured under common law without the need for registration or through federal or state registration regimes. In either case, use in commerce is typically required, although the federal protection can be secured for inherently distinctive marks with minimal (token) use. As noted above, it is now possible to reserve a trademark (and establish a priority date) by filing an intent-to-use application. The benefit of using this process is that the initial application is considered "constructive use," entitling the registrant to nationwide priority from the date of the application.

Examination of trademarks falls somewhere between the patent and copyright extremes. Trademarks must overcome greater technical hurdles than copyright law – classification of marks along the distinctiveness spectrum, prior art search, evaluation of evidence bearing on secondary meaning (in the case of non-inherently distinctive marks), statutory bars (immoral, deceptive, scandalous, disparaging, or functional marks are not registrable) – but are typically more straightforward to assess than patents.

Unlike patent law, the trademark system provides for a full inter partes opposition system. Marks eligible for registration are published in the PTO's Official Gazette, after which third parties have a 30 day period during which they may oppose registration. If no opposition is filed or the applicant prevails, then the mark is registered. Five years after registration, the trademark owner may apply for incontestability status, which insulates marks from being invalidated on the grounds that they are merely descriptive (i.e., lack secondary meaning) or lack priority. They may, however, continue to be challenged on several grounds, including abandonment, fraud, functionality, and generic status. Such rules reduce the risks of improvident grants of rights, provide for greater security upon registration, and reduce the costs of litigation.

II.F. Comparative Analysis

As in the context of promoting innovation, trademark law represents but a part of a broad and complex array of legal regimes and public and private institutions that address the problem of ensuring the informational integrity of markets (Best 1985). Therefore, as with patent, copyright and trade secret law, economic analysis of trademark law should take into consideration the full landscape of governance institutions and instruments.

As trademark law has evolved from the common law action for passing off the goods of one manufacturer for another's into relatively broad set of rights, other legal rights and institutions have also developed to police advertising and selling practices. In addition to common law causes of action for fraud and deceit, the federal and state governments have enacted consumer-protection statutes that both create private rights of action and empower public agencies (the Federal Trade Commission as well as state analogs) to investigate deceptive practices and enforce consumer protections (Sovern 1991). Federal and state governmental agencies proactively develop advertising and trade guidelines, field consumer complaints

about deceptive practices, and initiate enforcement actions. Private enforcement of such statutes has become a large legal practice area (Sheldon and Carter 1997). Non-governmental consumer protection organizations have developed to conduct independent product review (such as Consumer's Union), advocate for consumer protection regulation (e.g., Public Citizen), and support private enforcement of consumer protection laws (such as the National Consumer Law Center). Both public and private organizations have developed to provide independent certification of advertising claims and product quality (e.g., Underwriters Laboratories). Industry self-regulation has also emerged, most notably the Better Business Bureau, which processes consumer complaints and provides an alternative dispute resolution process for resolving false advertising complaints among businesses.

Each of these institutions draw upon different enforcers – consumers (and their attorneys), public entities, sellers in the market place, advertising industry organizations, independent certification laboratories, and consumer consortiums – with different sources of information and motivation to provide information and police disreputable sellers. The emergence of this broad range of enforcement resources suggests that trademark law should not be viewed as the sole or even principal means of protecting consumers. Rather, it should be seen as part of the composite mix. As a result, it need not be greatly concerned with the more egregious problems of consumer deception as other institutions focus more directly on such concerns.

Some trademark doctrines – such as the rule prohibiting assignment of trademarks in gross or the licensing of trademarks without supervision – may no longer be particularly important and may in fact raise transaction costs needlessly. These doctrines can also be questioned directly on economic grounds. It is not at all clear why a company which acquires a valuable trademark "in gross" would be any more inclined to engage in opportunism than the original owner or an assignee which took over the underlying business. Similarly, trademark licensors strong incentives to develop efficient supervisory structures even without a rule prohibiting "naked" licenses. In any case, consumer protection laws may provide a better institutional means of confronting the problems that these trademark rules purport to address.

References

Abramowicz, M. (2003), "Perfecting Patent Prizes", Vanderbilt Law Review 56:115-236.

Ackerlof, G. A. (1970), “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism”, Quarterly Journal of Economics 84:488-500.

Akerlof, G.A. et al. (2003), as Amici Curiae in Support of Petitioners in Eldred v. Ashcroft, 537 U.S. 186 (2003).

Allison, J. R. and M.A. Lemley (1998), "Empirical Evidence on the Validity of Litigated Patents", AIPLA Quarterly Journal 26:185-275.

Allison, J.R. and M.A. Lemley (2000), “How Federal Circuit Judges Vote in Patent Validity Cases”, Florida State University Law Review 27:745-766.

Allison, J.R. and M.A. Lemley(2002), "The Growing Complexity of the United States Patent System", Boston University Law Review 82:77-145.

Allison, J.R., M.A. Lemley, K. Moore, and R. D. Trunkey (2004), “Valuable Patents”, Georgetown Law Journal 92:435-479.

Allison, J.R. and E. H. Tiller (2003a), “Internet Business Method Patents”, in: W. Cohen and S. Merrill, eds., Patents in the Knowledge-Based Economy (National Academy Press, Washington, D.C.) 259-285.

Allison, J.R. and E. H. Tiller (2003b), “The Business Method Patent Myth”, Berkeley Technology Law Journal 18:987-1084.

Alston, J. (2002), "Spillovers", Australian Journal of Agricultural and Resource Economics 46:315-346.

Anton, J. and D. Yao (2005), “Finding Lost Profits: An Equilibrium Analysis of Patent Infringement Damages." Fuqua Business School, Duke University, Working paper FRPS04-108.

Aoki, R. and Y. Tauman (2001), “Patent Licensing with Spillovers”, Economic Letters 73:125-130.

Armond, M. (2003), “Introducing the Defense of Independent Invention to Motions for Preliminary Injunctions in Patent Infringement Lawsuits”, California Law Review 91:117-162.

Arora, A., A. Fosfuri, and A. Gambardella (2001), Markets for Technology: The Economics of Innovation and Corporate Strategy (MIT Press, Cambridge).

Arrow, K. (1962), "Economic Welfare and the Allocation of Resources for Invention", in: R.R. Nelson, ed., Universities-National Bureau of Economic Research Conference Series. The Rate and Direction of Economic Activities: Economic and Social Factors (Princeton University Press, New York).

Ayres, I. and P. Klemperer (1999), "Limiting patentees' market power with reducing innovation incentives: The perverse benefits of uncertainty and noninjunctive remedies", Michigan Law Review 97:985-1033.

Bakos, Y., E. Brynjolfsson, and D. Lichtman (1999), "Shared Information Goods", Journal of Law and Economics 42:117-155.

Bar-Gill, O. and G. Parchomovsky (2003), "The Value of Giving Away Secrets." Virginia Law Review 89:1857-1895.

Barton, J.H. (1997), "Patents and Antitrust: A Rethinking in Light of Patent Breadth and Sequential Innovation", Antitrust Law Journal 65:446-449.

Barton, J.H. (2000), "Intellectual Property Rights: Reforming the Patent System", Science 287:1933-1934.

Barton, J.H. (2002), "Antitrust Treatment of Oligopolies with Mutually Blocking Patent Portfolios", Antitrust Law Journal 69:851-882.

Barton, J.H. (2003), "Nonobviousness", IDEA The Journal of Law and Technology 43:475-508.

Barzel, Y. (1968), “Optimal Timing of Innovations”, Review of Economics and Statistics 50:348-355.

Benkler, Y. (2002), “Coase's Penguin,or Linux and the Nature of the Firm”, Yale Law Journal 112:369-446.

Benkler, Y. (2004), "Sharing Nicely: On shareable goods and the emergence of sharing as a modality of economic production”, Yale Law Journal 114:273-358.

Besen, S.M. and S.N. Kirby (1989a). "Private Copying, Appropriability and Optimal Copyright Royalties." Journal of Law and Economics 32:255-280.

Besen, S.M. and S.N. Kirby (1989b), "Compensating Creators of Intellectual Property", (Rand Corporation, Santa Monica, CA), No. R-3751-MF.

Besen, S.M., S.N. Kirby, and S.C. Salop (1992), "An Economic Analysis of Copyright Collectives", Virginia Law Review 78:383-411.

Besen, J. and E. Maskin, E. (2002), "Sequential Innovation, Patents and Imitation" (MIT, Department of

Economics, Cambridge, MA), Working Paper 00-01.

Best, A. (1985), “Controlling False Advertising: A Comparative Study of Public Regualtion, Industry Self-Policing, and Private Litigation”, Georgia Law Review 20:1-72.

Bhattacharya, S., C. d'Aspremont, and J.-L. Gerard-Varet (2000), "Bargaining and Sharing Innovative Knowledge", Review of Economic Studies 67:255-271.

Bhattacharya, S., J. Glazer, and D.E.M. Sappington (1990), "Sharing Productive Knowledge in Internally Financed R&D Contests", Journal of Industrial Economics 39:187-208.

Bhattacharya, S., J. Glazer, and D.E.M. Sappington (1992), "Licensing and the Sharing of Knowledge in Research Joint Ventures", Journal of Economic Theory 56:43-69.

Blair, R.D. and T.F. Cotter (1998), "An Economic Analysis of Damages Rules in Intellectual Property Law", William and Mary Law Review 39:1585-1694.

Blair, R.D. and T.F. Cotter (2001), "Rethinking Patent Damages", Texas Intellectual Property Law Journal 10:1-93.

Blair, R.D. and T.F. Cotter (2002), “Strict Liability and its Alternatives in Patent Law”, Berkeley Technology Law Journal 17:799-845.

Bone, R.G. (2004), “Enforcement Costs and Trademark Puzzles”, Virginia Law Review 90:2099-2185.

Brown, R.S. Jr. (1948), “Advertising and the Public Interest: Legal Protection of Trade Symbols”, Yale Law Journal 57:1165-____.

Britt, B. (1990), “International Marketing: Disney’s Global Goals,” Marketing, May 17, 1990, 22-26.

Brocas, I. (2004), "Optimal Regulation of Cooperative R&D under Incomplete Information", Journal of Industrial Organization 52:81-120.

Burge, D. (1984), Patent and Trademark Tactics and Practice (John Wiley, New York).

Burk, D.L. and M.A. Lemley (2002), “Is Patent Law Technology-Specific?”, Berkeley Technology Law Journal 17:1155-1206.

Burk, D.L. and M.A. Lemley (2003), “Policy Levers in Patent Law”, Virginia Law Review 89: 1575-1696.

Burke, T.P. (1994), “Software Patent Protection: Debugging the Current System”, 69 Notre Dame Law Review 69:1115-1166.

Cai, M. (2004), “Madey v. Duke University: Shattering the Myth of Universities' Experimental Use Defense”, Berkeley Technology Law Journal 19:175-192.

Calabresi, G. and A.D. Melamed (1972), "Property Rules, Liability Rules, and Inalienability: One view of the Cathedral", Harvard Law Review 85:1089-1126.

Carter, S.L. (1990), “The Trouble with Trademark”, Yale Law Journal 99:759-800.

Chamberlin, E. (1933), “The Theory of Monopolistic Competition”, (Ph.D. Thesis, Harvard), p. 56, p. 204.

Chang, H.F. (1995), “Patent Scope, Antitrust Policy, and Cumulative Innovation”, Rand Journal of Economics 26:34-57.

Chen, Y. and I. P’ng (2003), “Information Goods, Pricing, and Copyright Enforcement: Welfare Analysis”, Information Systems Research 14:107-123

Choi, J.P. (1991), "Dynamic R&D Competition under 'Hazard Rate' Uncertainty", Journal of Industrial Economics 22:596-610.

Chu, C.A., (2001), “Empirical Analysis of the Federal Circuit's Claim Construction Trends”, Berkeley Technology Law Journal16:1075-1164.

Clarke, R. (2003), “U.S. Continuity Law and its Impact on the Comparative Patenting Rates of the U.S., Japan, and European Patent Offices”, Journal of the Patent and Trademark Office Society 8:335-49.

Cockburn, I. and R. Henderson (2003), "Survey Results from the 2003 Intellectual Property Owners Association Survey on Strategic Management of Intellectual Property", (Boston University and MIT, Departments of Economics.), mimeograph.

Cohen, J.E. and M.A. Lemley (2001), “Patent Scope and Innovation in the Software Industry”, California Law Review 89:1-57.

Cohen, W.M., R.R. Nelson, and J.P. Walsh (2000), “Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not), National Bureau of Economic Research, Working Paper No. 7552 (Cambridge, Massachusetts), at: http://papers.nber.org/papers/w7552.pdf.

Comanor, W.S. and T.A. Wilson (1974), “Advertising and Market Power”, (Harvard University Press).

Commission on Revision of the Federal Court Appellate System, Structure and Internal Procedures: Recommendations for Change (1975), reprinted in Federal Register 67:54878-54905.

Conner, K. R. and R. P. Rumelt (1991), "Software Piracy:An Analysis of Protecting Strategies", Management Science 37:125-139.

Cooter, R. (1982), "Economic Analysis of Punitive Damages", Southern California Law Review 56:79-100.

Corcoran, R. (1999), "Quality Review and Control in the PTO: The Historical Evolution", Journal of the Patent and Trademark Office Society 81:7-11.

Cornelli, F. and M. Schankerman (1999), "Patent Renewals and R&D Incentives", The Rand Journal of Economics 30:197-213.

Cotter, T.F. (2004), "An Economic Analysis of Enhanced Damages and Attorneys' Fees for Willful Infringement", Federal Circuit Bar Journal 14:291-331.

Coverdale, J.F. (1984), "Comment, Trademarks and Generic Words: An Effect-on-Competition Test", University of Chicago Law Review 51:868-891.

Darby, M.R. and E. Karni (1973), "Free Competition and the Optimal Amount of Fraud", Journal of Law and Economics 16:67-88.

Dasgupta, P. and J. Stiglitz (1980a), "Industrial Structure and the Nature of Innovative Activity", Economic Journal 90:266-293.

Dasgupta, P. and J. Stiglitz (1980b), "Uncertainty, Market Structure and the Speed of Research", Bell Journal of Economics 11:1-28.

d'Aspremont, C. and A. Jacquemin (1988), "Cooperative and Noncooperative R&D in Duopoly with Spillovers", American Economic Review 5:1133-1137.

David, P. (1985), New Technology, Diffusion, Public Policy, and Industrial Competitiveness (Stanford University, Center for Economic Policy Research).

David, P. (2003), "The Economic Logic of 'Open Science' and the Balance between Private Property Rights and the Public Domain in Scientific Data and Information:A Primer," in: National Research Council, The Role of the Public Domain in Scientific and Technical Data and Information (National Academies Press, Washington, D.C.).

de Laat, E.A. (1996), "Patents or Prizes: Monopolistic R&D and Asymmetric Information", International Journal of Industrial Organization 15:369-390.

Denicola, R.C., (1999), “Freedom to Copy”, Yale Law Journal 108:1661-1686.

Denicolò, V. (1996), "Patent Races and Optimal Patent Breadth and Length." Journal of Industrial Economics 44:249-265.

Denicolò, V. (1997), "Patent Policy with a Finite Sequence of Patent Races", (Department of Economics, University of Bologna), mimeograph.

Denicolò, V. (2002a), "Two-Stage Patent Races and Patent Policy", Journal of Industrial Economics 31:488-501.

Denicolò, V. (2002b), "Sequential Innovation and the Patent-Antitrust Conflict." Oxford Economic Papers 54:649-668.

Desmond, R. (1993), “Nothing Seems ‘Obvious’ to the Court of Appeals for the Federal Circuit: The Federal Circuit, Unchecked by the Supreme Court, Transforms the Standard of Obviousness Under Patent Law”, Loyola of Los Angeles Law Review 26:455-490.

Dogan, S.L. and M.A. Lemley (2004a), “Trademarks and Consumer Search Costs on the Internet”, Houston Law Review 41:777-838.

Dogan, S.L. and M.A. Lemley (2004b), “The Merchandising Right: Fragile Theory or Fait Accompli?”, Stanford Public Law Working Paper No. 105.

Dratler, J. Jr.(2003), “Does Lord Darcy Yet Live? The Case Against Software and Business-Method Patents”, Santa Clara Law Review 43:823-899.

Dreyfuss, R.C. (1989), “The Federal Circuit: A Case Study in Specialized Courts”, New York University Law Review 64:1-77.

Dreyfuss, R.C. (1990), “Expressive Genericity: Trademarks as Language in the Pesi Generation”, Notre Dame Law Review 65:397-424.

Dreyfuss, R.C. (2000), “Are Business Method Patents Bad for Business?”, 16 Santa Clara Computer and High Technology Law Journal16:263-280.

Dreyfuss, R.C. (2003), “Varying the Course”, in: F. Scott Kieff, ed., Perspectives on the Properties of the Human Genome Project (Elsevier, New York).

Duffy, J. (moderator) (1998), "Early Patent Publication: A Boon or Bane? A Discussion on the Legal and Economic Effects of Publishing Patent Applications After Eighteen Months of Filing", Cardozo Arts & Entertainment Law Journal 16:601-633..

Dunner, D., J. Jakes, and J. Karceski (1995), "A Statistical Look at the Federal Circuit's Patent Decisions: 1982-1994", Federal Circuit Bar Journal 5:151-180.

Dyson, J. (2001), A History of Great Inventions (Carroll & Graf, New York).

Economides, N. (1998), "Trademarks", in: Peter Newman, ed., The New Palgrave Dictionary of Economics and the Law (MacMillan, London), 601-603.

Eisenberg, R.S. (1989), "Patents and the Progress of Science: Exclusive Rights and Experimental Use", University of Chicago Law Review 56:1017-1086.

Eisenberg, R S. (1996), "Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research", Virginia Law Review 82:1663-1727.

Eisenberg, R.S. (2000), "The Promise and Perils of Strategic Prior Art Creation Through Publication: A Response to Professor Parchomovsky", Michigan Law Review 98:2358-2370.

Epstein, R.A. (2003), "Steady the Course: Property Rights in Genetic Material", in: F. Scott Kieff ed., Perspectives on Properties of the Human Genome Project (Elsevier, New York), 153-156.

European Patent Office (2002a). "Trilateral Statistical Reports 2001" at: http://www.european-patent-office.org/tws/tsr_2001/index.php

European Patent Office (2002b), "Facts and figures 2002" at: www.european-patent-office.org\TEXTsymbol\backslashepo\TEXTsymbol\backslashfacts_figures\TEXTsymbol\backslashfacts2001\TEXTsymbol\backslashpdf\TEXTsymbol\backslashfacts_figures_01.pdf

Farrell, J. and M. Katz (1998), "The Effects of Antitrust and Intellectual Property Law on Compatibility and Innovation", Antitrust Bulletin 48:609-650.

Farrell, J. and P. Klemperer (2004), "Coordination and Lock-in: Competition with Switching Costs and Network Effects", at: www.paulklemperer.org.(Nuffield College, Oxford). Forthcoming in Armstrong, M. and Porter, R.H., eds, Handbook of Industrial Organization, Vol. 3. (North Holland, Amsterdam).

Feit, I.N., (1989), "Biotechnology Research and the Experimental Use Exception to Patent Infringement", Journal of the Patent and Trademark Office Society 71:819-840.

Fisher, W. (2004), “Promises to Keep: Technology, Law, and the Future of Entertainment”, (Stanford University Press).

Folsom, R.H. and L.L. Teply (1980), “Trademarked Generic Words”, Yale Law Journal 89:1323-1359.

Folsom, R.H. and L.L. Teply (1988a), “Surveying ‘Genericness’ in Trademark Litigation”, The Trademark Reporter 78:1-31

Folsom, R.H. and L.L. Teply, (1988b), “A Reply to Swann and Palladino’s Critique of Folsom and Teply’s Model Survey”, The Trademark Reporter 78:197-208.

Foray, D. and L. Hilaire-Perez (2000), "The Economics of Open Technology:Collective Organization and Individual Claims in the "Fabrique Lyonnaise" during the Old Regime", at: http://emlab.berkeley.edu/users/bhhall/ipconf01.html (conference paper).

Friedman, D.D., W.M. Landes, and R.A. Posner (1991), “Some Economics of Trade Secret Law”, Journal of Economic Perspectives 5:61-72.

Gale, I. and Y-K. Che (2003), "Optimal Design of Research Contests", American Economi Review 93: 646-671.

Galler, B.A. (1990), “A Proposal for a Software Patent Institute”, June 26, 1990.

Gallini, N.T. (1984), "Deterrence by Market Sharing: A Strategic Incentive for Licensing", American Economic Review 74:931-941.

Gallini, N.T. (1992), "Patent Policy and Costly Imitation", Journal of Industrial Economics 44:52-63.

Gallini, N.T and S. Scotchmer (2002), "Intellectual Property: When is it the Best Incentive Mechanism?", Innovation Policy and the Economy 2:51-78.

Gallini, N.T. and R. Winter (1985), "Licensing in the theory of Innovation", Journal of Industrial Economics 16:237-252.

Gandal, N. and S. Scotchmer (1993), "Coordinating Research through Research Joint Ventures", Journal of Public Economics 51:173-193.

Gervais, D. (2005), “The Price of Social Norms: Towards a Licensing Regime for File-Sharing”, Journal of Intellectual Property Law 12:__-__.

Ghosh, S. and J. Kesan (2004), “What Do Patents Purchase? In Search of Optimal Ignorance in the Patent Office”, Houston Law Review 40:1219-1264.

Gilbert, R. (2002), "Patent Pools:100 Years of Law and Economic Solitude", at: http://www.innovationlaw.org/pages/patent_pools.doc.

Gilbert, R. and D. Newbery (1982), "Preemptive Patenting and the Persistence of Monopoly", American Economic Review 72:514-526.

Gilbert, R. and C. Shapiro (1990), "Optimal Patent Length and Breadth", Journal of Industrial Economics 21:106-112.

Gilbert, R. and C. Shapiro (1997), "Antitrust Issues in the Licensing of Intellectual Property: The Nine No-No's Meet the Nineties", in: C. Winston, ed., Brookings Papers on Microeconomic Activity (The Brookings Institution, Washington, D.C.).

Gilbert, R., and G.C. Sunshine (1995a),. "Incorporating Dynamic Efficiency Concerns in Merger Analysis:The Use of Innovation Markets", Antitrust Law Journal 63:569-602.

Gilbert, R., and G.C. Sunshine (1995b), "The Use of Innovation Markets:A Reply to Hay, Rapp and Hoerner", Antitrust Law Journal 64:75-82.

Goldstein, P. (1986), “Infringement of Copyright in Computer Programs”, University of Pittsburgh Law Review 47:1119-1130

Gordon, W. (1982), "Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and its Predecessors", Columbia Law Review 82:1600-1655.

Grady, M.F. and J.I. Alexander (1992), “Patent Law and Rent Dissipation”, Virginia Law Review 78:305-350.

Green, J. and S. Scotchmer (1995), "On the Division of Profit in Sequential Innovation", Journal of Industrial Economics 26:20-33.

Grossman, G. and E. Lai (2001), "International Protection of Intellectual Property", mimeograph (Princeton University, Princeton, NJ).

Grushcow, J. (2004), “Measuring Secrecy: A Cost of the Patent System Revealed”, Journal of Legal Studies 33:59-84.

Hall, B.H. (2003), "Business Method Patents, Innovation and Policy", Working Paper 9717 (National Bureau of Economic Research, Cambridge, MA).

Hall, B.H. and R.H. Ziedonis (2001), "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995", Journal of Industrial Economics 32:101-128.

Heller, M.A. and R.S. Eisenberg (1998), "Can Patents Deter Innovation? The Anticommons in Biomedical Research", Science 280:698-701.

Hicks, D. (1995), “Published Papers, Tacit Competencies, and Corporate Management of the Public/Private Character of Knowledge”, Industrial and Corporate Change 4:401-424.

Hicks, J. (1932), The Theory of Wages (MacMillan, London).

Higgins, R.S. and P.H. Rubin (1986), “Counterfeit Goods”, Journal of Law and Economics 29:211-230.

Hirshleifer, J. (1973), “Where Are We in the Theory of Information?”, American Economic Review. Papers and Proceedings 63:31-39.

Horstmann, I., G.M. MacDonald, and A. Slivinski (1985), "Patents as Information Transfer Mechanisms:To Patent or (Maybe) Not to Patent", Journal of Political Economy 93:837-858.

Hovenkamp, H., M. Janis, and M.A. Lemley (2004), IP and Antitrust: An Analysis of Antitrust Principles Applied to Intellectual Property Law (Aspen Publishers, New York).

Hunt, R. (1999), “Nonobviousness and the Incentive to Innovate:An Economic Analysis of Intellectual Property Reform", Working Paper 99-3 (The Federal Reserve Bank of Philadelphia, Philadelphia).

Hunt, R. (2004), "Patentability, Industry Structure, and Innovation", Journal of Industrial Economics 52:401-425.

Jaffe, A.B. and J. Lerner (2004), Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It (Princeton University Press, New Jersey).

Johnson, W.R. (1985), "The Economics of Copying", Journal of Political Economy 93:158-174.

Kamien, M.I., E. Muller, and I. Zang (1992), “Research Joint Ventures and R&D Cartels”, American Economic Review 85:1293-1306.

Kaplow, L. (1984), “The Patent-Antitrust Intersection:A Reappraisal”, Harvard Law Review 97:1813-1892.

Kaplow, L. and S. Shavell (1996), “Property Rules versus Liability Rules: An Economic Analysis”, Harvard Law Review 109:715-789.

Karjala, D.S. and P.S. Menell (1995), “Applying Fundamental Copyright Principles to Lotus Development Corp. v. Borland International, Inc.”, High Technology Law Journal 10:177-192.

Katz, M.L. and C. Shapiro (1986), "Technology Adoption in the Presence of Network Externalities", Journal of Political Economy 94:822-881.

Katz, M.L. and C. Shapiro (1987), "R&D; Rivalry with Licensing or Imitation", American Economic Review 77:402-420.

Kesan, J.P. (2002), "Carrots and Sticks to Create a Better Patent System", Berkeley Technology Law Journal 17:763-797.

Kieff, F.S. (2001a), "Facilitating Scientific Research: Intellectual Property Rights and the Norms of Science--a Response to Rai and Eisenberg", Northwestern University Law Review 95:691-705.
Kieff, F.S. (2001b), "Property Rights and Property Rules for Commercializing Inventions", Minnesota Law Review 85:697-754.

Kitch, E.W. (1977), "The nature and function of the patent system", The Journal of Law and Economics 20:265-290.

Kitch, E.W. (1980), "The Law and Economics of Rights in Valuable Information", Journal of Legal Studies 9:683-723.

Klein, B. and K.B. Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance", Journal of Political Economy 89:615-41.

Klemperer, P. (1990), "How Broad Should the Scope of Patent Protection Be?", Journal of Industrial Economics 21:113-130.

Kortum, S. and J. Lerner (1998), "Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?", Carnegie-Rochester Conference Series on Public Policy, 48:247-304.

Kozinski, A. (1993), "Trademarks Unplugged", New York University Law Review 68:960-978.

Kratzke, W.P. (1991), "Normative Economic Analysis of Trademark Law", Memphis State University Law Review 21:199-286.

Kremer, M. (1998), "Patent Buyouts: A mechanism for encouraging innovation", Quarterly Journal of Economics 113:1137-1168.

La Manna, M., R. MacLewod, and D. de Meza (1989), "The Case for Permissive Patents", European Economic Review 33:1427-1443.

Landes, W.M. and R.A. Posner (1987), "Trademark Law: An Economic Perspective", Journal of Law and Economics, 30:265-309.

Landes, W.M. and R.A. Posner (2003), The Economic Structure of Intellectual Property Law (Harvard University Press, Cambridge).

Landes, W.M. and R.A. Posner (2004), "Intellectual Property, an Empirical Analysis of the Patent Court", University of Chicago Law Review 71:111-128.

Lanjouw, J.O. (1998), "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value", The Review of Economic Studies 65:761-710.

Lanjouw, J.O. and J. Lerner (1997), "The Enforcement of Intellectual Property Rights: A Survey of the Empirical Literature", National Bureau of Economic Research Working Paper No. 6296.

Lanjouw, J.O. and J. Lerner (2001), "Tilting the Table? The Predatory Use of Preliminary Injunctions", Journal of Law and Economics 44:573-603.

Lanjouw, J.O. and M. Schankerman (2001), "Characteristics of Patent Litigation: A Window on Competition", The Rand Journal of Economics 32:129-151.

Lanjouw, J.O., and M. Schankerman (2004), "Protecting Intellectual Property Rights: Are Small Firms Handicapped?", Journal of Law and Economics 47:45-74.

Lee, T. and L. Wilde (1980), "Market Structure and Innovation: A Reformulation", Quarterly Journal of Economics 94:429-436.

Leibenstein, H. (1950), "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand", Quarterly Journal of Economics, 64, 183-207.

Leibovitz, J.S. (2002), "Inventing a Nonexclusive Patent System", Yale Law Journal111:2251-2287.

Lemley, M.A. (1997), "The Economics of Improvement in Intellectual Property Law", Texas Law Review. 75:989-1083.

Lemley, M.A. (1999), "The Modern Lanham Act and the Death of Common Sense", Yale Law Journal 108:1687-1715.

Lemley, M.A. (2001), "Rational Ignorance at the Patent Office", Northwestern University Law Review 95:1495-1532.

Lemley, M.A. (2002), "Intellectual Property Rights and Standard-Setting Organizations", California Law Review 90:1889-1979.

Lemley, M.A. (2004), "Ex Ante versus Ex Post Justifications for Intellectual Property Law", University of

Chicago Law Review 71:129-149.

Lemley, M.A. and C.V. Chien (2003), "Are the U.S. Patent Priority Rules Really Necessary?", Hastings Law Journal 54:1299-1333.

Lemley, M.A. and K.A. Moore (2004), "Ending Abuse of Patent Continuations", Boston University Law Review 84:63-126 (2004).

Lemley, M.A. and D.W. O'Brien (1997), "Encouraging Software Reuse", Stanford Law Review 49:255-304.

Lemley, M.A. and R. Reese (2004), "Reducing Digital Copyright Infringement Without Restricting Innovation", Stanford Law Review 56:1345-1434.

Lemley, M.A. and R.K. Tangri (2003), "Ending Patent Law's Willfulness Game", Berkeley Technology Law Journal 18:1085-1125.

Lerner, J. (1994), "The Importance of Patent Scope: An Empirical Analysis", Rand Journal of Economics 25:319-33.

Lerner, J. (1995), "Patenting in the Shadow of Competitors", Journal of Law and Economics 38:563-595.

Lerner, J. and J. Tirole (2000), "The Simple Economics of Open Source", Working Paper 7600 (National Bureau of Economic Research, Cambridge).

Lerner, J. and J. Tirole (2004), "Efficient Patent Pools", American Economic Review 94:691-711.

Lesser, W. and T. Lybbert (2004), "Do Patents Come too Easy?", IDEA: The Journal of Law and Technology 44:381-407.

Lessig, L. (2004), Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity (Penguin Press, New York).

Levin, J. and R. Levin (2003), "Benefits and Costs of an Opposition Process" in: W. Cohen and S. Merrill, eds., Patents in the Knowledge-Based Economy (National Academy Press, Washington, D.C.)

Levin, R.C., W.M. Cohen, and D.C. Mowery (1994), "R&D Appropriability, Opportunity and Market Structure: New Evidence on some Schumpeterian Hypotheses", Papers and Proceedings of the American Economic Association 75:20-24.

Levin, R.C., A.K. Klevorick, R.R. Nelson, and S.G. Winter (1987), "Appropriating the Returns from Industrial R&D", Brookings Papers on Economic Activity: Microeconomics 3:783-820.

Lichtman, D. (1997), "The Economics of Innovation: Protecting Unpatentable Goods", Minnesota Law Review 81:693-734.

Lichtman, D., S. Baker, and K. Kraus (2000), "Strategic Disclosure in the Patent System", Vanderbilt Law Review 53:2175-2217.

Lichtman, D. and W. Landes (2003)., "Indirect Liability for Copyright Infringement: An Economic Perspective", Harvard Journal of Law & Technology 16: 395-410.

Liebowitz, S.J. (1982), "Durability, Market Structure, and New-Used Goods Models", American Economic Review 72:816-824.

Liebowitz, S.J. (1985), "Copying and Indirect Appropriability: Photocopying of Journals", Journal of Political Economy 93:945-957.

Liebowitz, S.J. (2004), "File Sharing: Creative Destruction or Just Plain Destruction?", at: http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=59984.

Liebowitz, S J and S.E. Margolis, "Journals as Shared Goods: Comment", American Economic Review 72:597-602.

Long, C. (2002), "Patent Signals", University of Chicago Law Review 69:625-679.

Loury, G.C. (1979), "Market Structure and Innovation", Quarterly Journal of Economics 93:395-410.

Lunney, G.S. Jr. (1999), "Trademark Monopolies", Emory Law Journal 48:367-487.

Lunney, G.S. Jr. (2000-01), "E-Obviousness", Michigan Telecommunication and Technology Law Review 7:363-422.

Lunney, G.S. Jr. (2004), "Patent Law, the Federal Circuit and the Supreme Court:A Quiet Revolution", Supreme Court Economic Review 11:1-80.

Macedo, C. (1990), "First-to-File: Is American Adoption of the International Standard in Patent Law Worth the Price?", AIPLA Quarterly Journal 18:193-234.

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

Mak, J. and G.M. Walton (1972), "Steamboats and the Great Productivity Surge in River Transportation", Journal of Economic History 32:619-640.

Mansfield, E. (1986), "Patents and Innovation: An Empirical Study". Management Science 32:173–181.

Maskus, K. (2000a),. Intellectual Property Rights in the Global Economy (Institute for International Economics, Washington, D.C.).

Maskus, K. (2000b), "Lessons from Studying the International Economics of Intellectual Property Rights", Vanderbilt Law Review 53:2219-2240.

Matutes, C. and P. Regibeau, (1992), "Compatibility and Bundling of Complementary Goods in a Duopoly", Journal of Industrial Economics 40:37-54.

Maurer, S.M. (2003), "New Institutions for Doing science: From Databases to Open Source Biology", at: http://www.merit.unimaas.nl/epip/papers/maurer_paper.pdf.

Maurer, S.M. and S. Scotchmer (2002), "The Independent Invention Defense in Intellectual Property", Economica. 69:535-547.

Maurer, S.M. and S. Scotchmer (2004a), "Profit Neutrality in Licensing: The Boundary between Antitrust Law and Patent Law", Working Paper 10546 (National Bureau of Economic Research, Cambridge).

Maurer, S.M. and S. Scotchmer (2004b), "Procuring Knowledge", in G. Libecap, ed., Advances in the Study of Entrepreneurship, Innovation and Growth, Vol 15 (JAI Press (Elsevier), Amsterdam), 1-31.

McCalman, P. (2001), "Reaping What you Sow: An Empirical Analysis of International Patent Harmonization", Journal of International Economics 55:161-185.

McCarthy, J.T. (2004), McCarthy on Trademarks and Unfair Competition (Westlaw).

McClure, D.M. (1979), "Trademarks and Unfair Competition: A Critical History of Legal Thought", The Trademark Reporter 69:305-356.

McClure, D.M. (1996), "Trademarks and Competition: The Recent History", Law and Contemporary Problems, 59:13-43.

McGowan, D. (2001), "The Legal Implications of Open Source Software", University of Illinois Law Review 2001:241-304.

McDade, A.S. (1998), "Trading in Trademarks – Why the Anti-Assignment in Gross Doctrine Should Be Abolished When Trademarks Are Used as Collateral", Texas Law Review 77:465-492.

Menell, P.S. (1987), "Tailoring Legal Protection for Computer Software", Stanford Law Review 39:1329-1372.

Menell, P.S. (1989), "An Analysis of the Scope of Copyright Protection for Application Programs", Stanford Law Review 41:1045-1104.

Menell, P.S. (1994), "The Challenges of Reforming Intellectual Property Protection for Computer Software, Columbia Law Review 94:2644-2654.

Menell, P.S. (1998a), "Intellectual Property: General Theories", in: B. Bouckaert and Gerrit De Geest, eds., Encyclopedia of Law and Economics Volume II (Edward Elgar, Cheltenham).

Menell, P.S. (1998b), "An Epitaph for Traditional Copyright Protection of Network Features of Computer Software", Antitrust Bulletin 43:651-713.

Menell, P.S. (2000), "Economic Implications of State Sovereign Immunity from Infringement of Federal Intellectual Property Rights", Loyola of Los Angeles Law Review 33:1399-1466.

Menell, P.S. (2003), "Envisioning Copyright Law's Digital Future", New York Law Review 46:63-193.

Menell, P.S. (2005), "Indirect Copyright Liability: A Re-examination of Sony's Staple Article of Commerce Doctrine", UC Berkeley Public Law Research Paper No. 682051.

Merges, R.P. (1988), "Commercial Success and Patent Standards: Economic Perspectives on Innovation", California Law Review 76:803-876.

Merges, R.P. (1992), "Uncertainty and the Standard of Patentability", High Technology Law Journal 7:1-70.

Merges, R.P. (1994), "Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents", Tennessee Law Review. 62:75-76.

Merges, R.P. (1996), "Contracting into Liability Rules: Intellectual Property Rights and Collective Rights Organizations", California Law Review 84:1293-1393.

Merges, R.P. (1999a), "As Many as Six Impossible Patents Before Breakfast:Property Rights for Business Concepts and Patent System Reform", Berkeley Technology Law Journal 14:577-615.

Merges, R.P. (1999b), "Institutions for Intellectual Property Transactions: The Case of Patent Pools", in: R. Dreyfuss, ed., Intellectual Products: Novel Claims to Protection and their Boundaries (Oxford University Press, Oxford).

Merges, R.P. (2004a), "Compulsory Licensing vs. the Three 'Golden Oldies' Property, Contract Rights and Markets", CATO Institute Policy Analysis No 508, at: http://www.cato.org/pubs/pas/pa508.pdf.

Merges, R.P. (2004b), "A New Dynamism in the Public Domain", University of Chicago Law Review 71:183-203.

Merges, R.P. and J. Duffy (2002), Patent Law and Policy, (Matthew Bender, San Francisco).

Merges, R.P., P.S. Menell, and M.A. Lemley. (2003), Intellectual Property in the New Technological Age (Aspen, New York).

Merges, R.P. and R. Nelson (1990), "On the Complex Economics of Patent Scope", Columbia Law Review 90:839-916.

Meurer, M.J. (2002), "Business Methods and Patent Floods", Washington University Journal of Law and Policy 8:309-339.

Meurer, M.J. (2003), "Controlling Opportunistic and Anti-competitive Intellectual Property Litigation", Boston College Law Review 44:509-544.

Mikhail, P. (2000), "Hopkins v. CellPro: An Illustration that Patenting and Exclusive Licensing of Fundamental Science is Not Always in the Public Interest", Harvard Journal of Law & Technology 13:375-394.

Milgrom, P. and J. Roberts (1986), "Price and Advertising Signals of Product Quality", Journal of Political Economy 94:796-821

Miller, J.S. (2004), "Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents", Berkeley Technology Law Journal 19:667-739.

Minehart, D. and S. Scotchmer (1999), "Ex Post Regret and the Decentralized Sharing of Information", Games and Economic Behavior 17:114-131.

Mokyr, J. (1990), The Lever of Riches: Technological Creativity and Economic Progress (Oxford University Press, New York).

Moore, K.A. (2000), "Judges, Juries, and Patent Cases – An Empirical Peek Inside the Black Box", Michigan Law Review 99:365-409.

Moore, K.A. (2001), "Are District Court Judges Equipped to Resolve Patent Cases?", Harvard Journal of Law & Technology 15:1-39.

Mossinghoff, G. (2002), "The First-to-Invent System Has Provided No Advantage to Small Entities", Journal of the Patent and Trademark Office Society 84:425-431.

Mowery, D., R. Nelson, B.N. Sampat, and A.A. Ziedonis (2001), "The growth of patenting and licensing by U.S. universities: an assessment of the effects of the Bayh-Dole act of 1980", Research Policy 30:99-119.

Mowery, D. and N. Rosenberg (1998), Paths of Innovation: Technological Change in 20th-Century America. (Cambridge, MA: Cambridge Univ. Press).

Mueller, J. (2001), "No 'Dilettante Affair': Rethinking the Experimental Use Exception to Patent Infringement for Biomedical Research Tools", Washington Law Review 76:1-66.

Nagle, T.T. (1981), "Do Advertising-Profitability Studies Really Show that Advertising Creates a Barrier to Entry?", Journal of Law and Economics 24:333-344.

National Research Council (2004), A Patent System for the 21st Century (National Academies Press, Washington, D.C.), Stephen A. Merrill, Richard C. Levin, and Mark B. Myers, eds.

Nelson, P. (1970), "Information and Consumer Behavior", Journal of Political Economy 78:311-329.

Nelson, P, (1974), "Advertising as Information", Journal of Political Economy 82:729-754.

Nelson, P. (1975), "The Economic Consequences of Advertising", Journal of Business 48:213-241.

Nelson, R.R. (1959), "The Simple Economics of Basic Scientific Research", Journal of Political Economy 67:297-306.

Nelson, R.R. (2003), "The Market Economy and the Scientific Commons", National Bureau of Economic Research, Summer Institute Working Paper, at: http://www.nber.org/~confer/2003/si2003/papers/pripe/nelson.pdf.

Nelson, R.R. and S. Winter (1982), An Evolutionary Theory of Economic Change (Harvard University Press, Cambridge).

Netanel, N. (2003), "Impose a Noncommercial Use Levy to Allow Free Peer-to-Peer File Sharing", Harvard Journal of Law & Technology 17:1-84.

Newell, R.G., A.B. Jaffe, and R.N. Stavins (1999), "The Induced Innovation Hypothesis And Energy-Saving Technological Change", The Quarterly Journal of Economics, 114:941-975.

Nordhaus, W. (1969), Invention, Growth and Welfare (MIT Press, Cambridge).

Novos, I. and M. Waldman (1984), "The Effects of Increasing Copyright Protection:An Analytic Approach", Journal of Political Economy 92:236-246.

Oberholzer, F. and K. Strumpf, K. (2004), "The Effects of Filesharing on Record Sales: An Empirical Analysis", at: http://www.unc.edu/~cigar/papers/FileSharing_March2004.pdf.

Oddi, A.S. (1988), "Assessing 'Genericness': Another View", The Trademark Reporter 78:560-578.

O'Donoghue, T. (1998), "A Patentability Requirement for Sequential Innovation", Journal of Industrial Economics 29:654-679.

O'Donoghue, T., S. Scotchmer and J.F. Thisse (1998), "Patent Breadth, Patent Life and the Pace of Technological Progress", Journal of Economics and Management Strategy 7:1-32.

Ordover, J.A. and R.D. Willig (1978), "On the Optimal Provisions of Journals qua Sometimes Shared Goods", American Economic Review 68:324-339.

O'Rourke, M. (2000), "Toward a Doctrine of Fair Use in Patent Law", Columbia Law Review 100:1177-1250.

Ottoz, E., and F. Cugno (2004), "The Independent Invention Defence in a Cournot-Duopoly Model", Economics Bulletin 12:1-7.

Pakes, A. (1985), "On Patents, R&D and the Stock Market Rates of Return", Journal of Political Economy 93:390-409.

Pakes, A. (1986), "Patents as options: Some estimates of the value of holding European patent stocks", Econometrica 54:755-784.

Pakes, A. and M. Schankerman (1984), "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources", in: Zvi Griliches, ed., R&D, Patents, and Productivity (University of Chicago Press, Chicago).

Parchomovsky, G. (2000), "Publish or Perish", Michigan Law Review 98:926-952.

Parchomovsky, G. and P. Siegelman (2002), "Toward an Integrated Theory of Intellectual Property", Virginia Law Review 88:1455-1428.

Parchomovsky, G. and R.P. Wagner (2005), "Patent Portfolios", University of Pennsylvania Law School, Public Law Working Paper 56, at http://ssrn.com/abstract=582201.

Park, Y. and S. Scotchmer (2004), Digital Rights Management and the Pricing of Digital Content, (NET Institute, New York University).

Png, I.P.L. and D. Reitman (1995), "What Are Some Products Branded and Others Not?", The Journal

of Law & Economics 38:207-224.

Polinsky, A.M. (1980), "Resolving Nuisance Disputes: The Simple Economics of Injunctive and Damage Remedies, Stanford Law Review 32:1075-____.

Polinsky A.M. and S. Shavell, "Punitive Damages: An Economic Analysis", Harvard Law Review 111:869-962,

Port, K.L. (1994), "The 'Unnatural' Expansion of Trademark Rights: Is a Federal Dilution Statute Necessary?", Seton Hall Legislative Journal 18:433-488.

R. Posner, (1980), The Federal Courts, (Harvard University Press, Cambridge).

Quillen, C. Jr., and O. Webster (2001-02),. "Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office", Federal Circuit Bar Journal 1:1-21.

Rai, A.K. (2000), "Addressing the Patent Gold Rush: The Role of Deference to PTO Patent Denials", Washington University Journal of Law & Policy 2:199-227.

Rai, A.K. (2002), "Specialized Trial Courts: Concentrating Expertise on Fact", Berkeley Technology Law Journal 17:877-897.

Rai, A.K. (2003), "Engaging Facts and Policy: A Multi-Institutional Approach to Patent System Reform", Columbia Law Review 103:1035-1135.

Raymond, E. (2001), The Cathedral and the Bazaar (O'Riley Press, New York).

Reinganum, J. (1981), "Dynamic Games of Innovation", Journal of Economic Theory 25:21-41.

Reinganum, J. (1982), "A Dynamic Game of R&D:Patent Protection and Competitive Behaviour", Econometrica 50:671-688.

Reinganum, J. (1985), "Innovation and Industry Evolution", Quarterly Journal of Economics 100:81-99.

Reinganum, J. (1989), "The Timing of Innovation:Research, Development and Diffusion", in: R. Schmalensee, and R.D. Willig, eds., Handbook of Industrial Organization (Elsevier, Amsterdam),. 849-908.

Robinson, J. (1933), The Economics of Imperfect Competition (Macmillan, London).

Rockett, K.E. (1990), "Choosing the Competition and Patent Licensing", The Rand Journal of Economics 21:161-171.

Rosenberg, N. (1972), "Factors Affecting the Diffusion of Technology", Explorations in Economic History 10:3-33.

Ruttan, V.W. (2001), Technology, Growth and Development: An Induced Innovation Perspective (Oxford University Press, Oxford).

Ryan, M.P. (1998), Knowledge Diplomacy (Brookings Institution Press, Washington, D.C.).

Samuelson, P. (1990), "Benson Revisited: The Case Against Patent Protection For Algorithms And Other Computer Program-related Inventions", Emory Law Journal 39:1025-1154.

Samuelson, P. (2004), "Intellectual Property Arbitrage: How Foreign Rules Can Affect Domestic Protections", University of Chicago Law Review 71:223-239.

Samuelson, P., R. David, M. Kapor, and J.H. Reichman (1994), "A Manifesto Concerning the Legal Protection of Computer Programs", Columbia Law Review 94:2308-2431.

Samuelson, P. and S. Scotchmer (2002), "The Law and Economics of Reverse Engineering", Yale Law Journal 111:1575-1663.

Sandburg, B. (2001), "Battling the Patent Trolls", The Recorder, July 30, 2001.

Saxenian, A. (1994) Regional Advantage Culture and Competition in Silicon Valley and Route 128 (Harvard University Press, Cambridge).

Schankerman, M. (1998), "How Valuable Is Patent Protection? Estimates by Technology Field", The Rand Journal of Economics 29:77-107.

Schankerman, M. and A. Pakes (1986), "Estimates of the value of patent rights in European Countries during the post-1950 period", Economic Journal 97:1052-1076.

Schankerman, M. and S. Scotchmer (2001), "Damages and Injunctions in Protecting Intellectual Property", Journal of Industrial Economics 32:199-220.

Schechter, F.I., (1927), "The Rational Basis of Trademark Protection", Harvard Law Review 40:813-___.

Scherer, F.M. and D.Ross (1990), Industrial Market Structure and Economic Performance (Third Edition) (Houghton Mifflin,Boston).

Schumpeter, J.A. (1942), Capitalism, Socialism and Democracy (Harper & Row, New York).

Scotchmer, S. (1991), "Standing on the Shoulders of Giants:Cumulative Research and the Patent Law",

Journal of Economic Perspectives 5:29-41.

Scotchmer, S. (1996), "Protecting Early Innovators:Should Second-Generation Products be Patentable?", Journal of Industrial Economics 27:322-331.

Scotchmer, S. (1999), "On the Optimality of the Patent System", The Rand Journal of Economics 30:181-196.

Scotchmer, S. (2003), "Intellectual Property:When is it the Best Incentive Mechanism for S&T Data?", in: National Research Council, The Role of the Public Domain in Scientific and Technical Data and Information (National Academies Press, Washington, D.C.).

Scotchmer, S. (2004a), "The Political Economy of Intellectual Property Treaties", Journal of Law, Economics and Organizations 20:415-437.

Scotchmer, S. (2004b), Innovation and Incentives (MIT Press, Cambridge).

Scotchmer, S. (2004c), "Patent Design, Patent Quality, and Patent Politics", Remarks before the European Patent Office, Munich, Germany.

Scotchmer, S. (2005a), "Consumption Externalities, Rental Markets and Purchase Clubs", Economic Theory 25:235-253.

Scotchmer, S. (2005b), Innovation and Incentives (MIT Press, Cambridge).

Scotchmer, S. and J. Green (1990), "Novelty and Disclosure in Patent Law", Journal of Industrial Economics 21:131-146.

Shapiro, C. (1982), "Consumer Information, Product Quality, and Seller Reputation", Bell Journal of Economics 13:20-35.

Shapiro, C. (1983), "Premiums for High-Quality Products as Returns to Reputations", Quarterly Journal of Economics 98:659-679.

Shapiro, C. (1985), "Patent Licensing and R&D Rivalry", American Economic Review 75:25-30.

Shapiro, C. (2001), "Navigating the Patent Thicket:Cross Licenses, Patent Pools, and Standard-Setting", Innovation Policy and the Economy 1:119-150.

Shapiro, C. (2003), "Antitrust Limits to Patent Settlements", Journal of Industrial Economics 34:391-411.

Shapiro, C. (2004), "Patent System Reform: Economic Analysis and Critique", Berkeley Technology Law

Journal 19:1017-1047.

Shavell, S. and T. van Ypserele (2001), "Rewards versus Intellectual Property Rights", Journal of Law and Economics 44:525-547.

Sheldon, J. and C.L. Carter (1997), Unfair and Deceptive Acts and Practices (National Consumer Law Center).

Shy, O. and J.-F. Thisse (1999), "A Strategic Approach to Software Protection." Journal of Economics and Management Strategy 8:163-190.

Sidak, J. (2004), "Trade Secrets and the Option Value of Involuntary Exchange", at: http://ssrn.com/abstract=577244.

Solow, R.M. (1957), "Technical Change and the Aggregate Production Function", Review of Economics and Statistics 39:312-320.

Soobert, A.M. (1998), "Breaking New Grounds in Administrative Revocation of U.S. Patents: a Proposition for Opposition--and Beyond", Santa Clara Computer and High Technology Law Journal 14:63-187.

Sovern, J. (1991), "Private Actions under the Deceptive Trade Practices Acts: Reconsidering the FTC Act as Rule Model", Ohio State Law Journal 52:437-467.

Sprigman, C. (2004), "Reform(aliz)ing Copyright", Stanford Law Review 57:485-568.

Stigler, G.J. (1961), "The Economics of Information", Journal of Political Economy 69:213-225.

Strandburg, K. (2004), "What Does the Public Get? Experimental Use and the Patent Bargain", Wisconsin Law Review 2004:81-153

Suzumura, K. (1992), "Cooperative and Noncooperative R&D in an Oligopoly with Spillovers", American Economic Review 5:1307-1320.

Swann, J.B. (1980), "The Economic Approach to Genericism: A Reply to Folsom and Teply", The Trademark Reporter 70:243-252.

Swann, J.B. and V. Palladino (1988), "Surveying 'Genericness': A Critique of Folsom and Teply", The Trademark Reporter 78:179-196.

Swiss Reinsurance Company (2000), "The significance of intellectual property assets, risks and insurance", at: http://www.swissre.com.

Tandon, P. (1982), "Optimal Patents with Compulsory Licensing", Journal of Political Economy 90:470-486.

Tandon, P. (1983), "Rivalry and the Excessive Allocation of Resources to Research", Bell Journal of Economics 14:152-165.

Taylor, C. and Z.A. Silberston (1973), The Economic Impact of the Patent System (Cambridge University Press, London).

Thomas, J.R. (1999), "The Patenting of the Liberal Professions", Boston College Law Review 40:1139-1185.

Thomas, J.R. (2001), "Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties", University of Illinois Law Review 2001:305-353.

Thomas, J.R. (2002), "The Responsibility of the Rulemaker: Comparative Approaches to Patent Administration Reform", Berkeley Technology Law Journal 17:727-761.

United States Department of Justice and Federal Trade Commission (1995), Antitrust Guidelines for Licensing of Intellectual Property.

United States Federal Trade Commission (2003), To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy.

United States Patent and Trademark Office (2001), Guidelines for Written Description and Utility Requirements.

Varian, H. (2000), "Buying, Renting, Sharing Information Goods", Journal of Industrial Economics 48:473-488.

Veblen, T. (1899), The Theory of the Leisure Class: An Economic Study of Institutions (Macmillan, New York).

von Hippel, E. (1988), The Sources of Innovation (Oxford University Press, Oxford).

von Hippel, E. (2001), "Innovation by User Communities:Learning from Open Source Software", Sloan Management Review 42:82-86.

Wagner, R.P. (2003), "Of Patents and Path Dependency: A Comment on Burk and Lemley", Berkeley Technology Law Journal 18:1341-1360.

Wagner, R.P. and L. Petherbridge (2004) "Is the Federal Circuit Succeeding? An Empirical Assessment

of Judicial Performance", University of Pennsylvania Law Review 152:1105-80.

Walsh, J.P., A. Arora, and W.M. Cohen (2003), "Effects of Research Tool Patents and Licensing on Biomedical Innovation", in: Wesley M. Cohen and Stephen A. Merrill, eds., Patents in the Knowledge-Based Economy (National Academy Press, Washington, D.C.), 285-340.

Wilkins, M. (1992), "The Neglected Intangible Asset: The Influence of the Trade Mark on the Rise of the Modern Corporation", Business History 34:66-95.

Wiley, J.S. Jr. (2003), "Taming Patent: Six Steps for Surviving Scary Patent Cases", UCLA Law Review 50:1413-1483.

Williamson O. (1986), Economic Organization: Firms, Markets, and Policy Control (New York University Press, New York).

Wright, B. (1983), "The Economics of Invention Incentives: Patents, Prizes and Research Contracts", American Economic Review 73:691-707.

Ziedonis, R.M. (2003), "Patent Litigation in the U.S. Semiconductor Industry", in: Wesley M. Cohen and Stephen A. Merrill, eds., Patents in the Knowledge-Based Economy (National Academy Press, Washington, D.C.), 180-212.

Cases

Cuno Engineering Corp. v. Automatic Devices Corp., 314 U.S. 84 (1941).

DaimlerChrysler AG v, Bloom, 315 F.3d 932 (8th Cir. 2003).

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

Fonar Corp. v. General Electric, 107 F.3d 1543 (Fed. Cir. 1997).

Gottschalk v. Benson, 409 U.S. 63 (1972).

Graver Tank & Mfg. Co. v. Linde Air Products Co., 339 U.S. 605 (1950).

Johns Hopkins Univ. v. Cellpro, Inc., 152 F.3d 1342 (Fed. Cir. 1998).

King-Seeley Thermos Co. v. Aladdin Industries, Inc., 321 F.2d 577 (2d Cir. 1963).

Madey v. Duke Univ., 307 F.3d 1351 (Fed. Cir. 2002), *cert. denied* 539 U.S. 958 (2003).

Mattel, Inc. v. MCA Records, 296 F.3d 894 (9th Cir. 2002), *cert. denied* 537 U.S. 1171 (2003).

Moseley v. V. Secret Catalogue, Inc., 537 U.S. 418 (2003).

O'Reilly v. Morse, 56 U.S. 62 (1854).

Roche Prods., Inc. v. Bolar Pharmaceutical Co., 733 F.2d 858 (Fed. Cir. 1984).

Times Mirror Magazines, Inc. v. Las Vegas Sports News, 212 F.3d 157 (3d Cir. 2000).

Trade-Mark Cases, 100 U.S. 82 (1879).

Wal-Mart Stores, Inc. v. Samara Brothers, Inc., 529 U.S. 205 (2000).

Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17 (1997).

Whittemore v. Cutter, 29 F. Cas. 1120 (C.C.D. Mass. 1813).

W.L. Gore & Assocs. v. Garlock, Inc., 721 F.2d 1540 (Fed. Cir. 1983).