Santa Clara University School of Law

Santa Clara High Technology Law Journal

Volume 28 | Issue 3 Article 2

1-2-2012

Why We Need a Strong Patent System and When: Filling the Void Left by the Bilski Case

Richard S. Gruner
John Marshall Law School

WHY WE NEED A STRONG PATENT SYSTEM AND WHEN: FILLING THE VOID LEFT BY THE *BILSKI* CASE

Richard S. Gruner†

If we could first know where we are, and whither we are tending, we could better judge what to do, and how to do it.[1]

Abraham Lincoln

The patent system . . . added the fuel of interest to the fire of genius in the discovery and production of new and useful things.[2]

Abraham Lincoln

Anything that won't sell, I don't want to invent. Its sale is proof of utility, and utility is success.[3]

Thomas Edison

Abstract

Patent law is presently under-theorized. Patents are granted to serve as rewards for certain types of inventive successes,[4] but the

† Professor of Law, John Marshall Law School, Chicago (Ph.D., University of California, Irvine, 2008; L.L.M., Columbia University School of Law, 1982; J.D., U.S.C. Law School, 1978; B.S., California Institute of Technology, 1975).

1. President Abraham Lincoln, Speech Before the Republican State Convention: House Divided (June 16, 1858) *in* I COMPLETE WORKS OF ABRAHAM LINCOLN 1 (John G. Nicolay & John Hay eds., 1894).

2. President Abraham Lincoln, Second Lecture on Discoveries and Inventions (Feb. 1859) *in* I COMPLETE WORKS OF ABRAHAM LINCOLN 113 (John G. Nicolay & John Hay eds., 1894).

3. Thomas Edison, *quoted in* Farha Abdol Chapar, *The Future of Innovation is Inspired by the Intellectual Property System*, *in* THE FUTURE OF INNOVATION 454 (Bettina von Stamm & Anna Trifilova eds., 2009).

4. The reward logic and technology enhancement goals of the patent system are embedded in the Constitutional provision that forms the basis for the United States patent system. The Constitution provides that Congress may establish a patent system "[t]o promote the Progress of . . . useful Arts" by granting inventors exclusive control over their discoveries

nature of the successes to be rewarded, the circumstances that should trigger rewards, and the size of the rewards that will best serve the public remain in substantial dispute. One of the primary reasons for these uncertainties is the incompleteness of underlying theories explaining why patented inventions deserve special treatment and rewards. The lack of good understanding of the theoretical justifications for patent rewards (and the limitations of those justifications) means that patent law standards are being reconsidered and revised without a clear sense of patent law goals. If we do not know where we are trying to go in setting patent standards, how can we know how to proceed?

This article seeks to overcome these limitations on present analyses of patent policies and standards. The article attempts to better define the innovation reward goals of patent laws and to set directions for future thinking about a variety of patent law standards. It describes a model of patent rewards based on Harold Demsetz's analyses of property controls limiting access to public goods. The article argues that patented invention designs should be limited to excludable designs for which workable access controls and access payment systems are feasible. Access to these sorts of inventions—as with other types of excludable public goods—should be regulated and controlled by patent rights to attract resources to the production of patentable inventions and to ensure that these inventions are produced at socially desirable levels.

The article considers the merit of providing patent rewards conforming to this model. It also considers at what points the model breaks down and under what circumstances patent rights should be withheld accordingly. By considering both the justifications for strong patent rewards and the limitations of those justifications, the contours of desirable patent standards in several areas of patent law become clearer. The concluding section of this article offers preliminary thoughts on how the model developed here can be used to shape several important patent law standards.

---

for a limited time. U.S. CONST., art. I, § 8, cl. 8. This provision envisions grants of exclusive rights in discoveries as means to reward discoverers and to thereby promote more discoveries and the expansion of knowledge about useful advances. Through this greater knowledge, the patent system was expected to achieve progress in the accumulation and use of specialized knowledge in the "useful arts."

I. INTRODUCTION

*A. The Post-*Bilski *Void in Patentable Subject Matter Standards*

Over the past two decades, federal courts have struggled to define the boundaries of the patent system.[5] The struggle has focused on attempts to define generally applicable standards for identifying patentable subject matter.[6] Such standards specify the types of advances that can qualify for patents if other patent law tests are met. Since an advance that is not patentable subject matter can never qualify for a patent no matter what other features the invention may have,[7] this standard places outer boundaries on the patent system. It specifies ranges of useful advances for which patents will never issue and for which patent incentives and rewards will never be factors.[8] Because patentable subject matter criteria limit the outer boundaries of the patent system and its effects in this way, these criteria are among our most important patent law standards.[9]

The clarity of patentable subject matter standards is a critically important feature of the patent system since these standards signal when potential innovators can look to patent rewards as incentives. Doubts about whether the patent system will apply reduce and weaken the incentives for innovation that the system provides. If a potential innovator is uncertain about whether a given type of advance will qualify for a patent, she will discount the importance of patent rewards below the levels that the innovator would perceive if she thought that the issuance of a patent was more certain. Hence, uncertainty regarding the range of patentable subject matters risks undercutting the incentives and impacts of the patent system in the areas of uncertainty. The judicial struggle over patentable subject matter standards—and the highly unresolved state of this struggle at present—have impaired the operation of the patent system and

5. The particular standards adopted and rejected by courts in these struggles are recounted in detail at a later point in this article. *See infra* text accompanying notes 57-94.

6. *See infra* text accompanying notes 57-94.

7. *See infra* text accompanying notes 57-94.

8. *See infra* text accompanying notes 57-94.

9. Subject matter standards for determining the range of items and processes that are governed by property rights are among the bedrock standards of any property system. *See generally* Abraham Bell & Gideon Parchomovsky, *A Theory of Property*, 90 CORNELL L. REV. 531, 575-76 (2005) (arguing that any property law must address four interlocking fundamental questions: "(1) what things property law protects; (2) vis-à-vis whom; (3) with what rights; and (4) by what enforcement mechanism").

potentially curtailed the development of valuable technologies by creating doubts about the applicability and scope of patent rewards for many socially valuable advances.

The Supreme Court's brief and largely superficial opinion in *Bilski v. Kappos*[10] has brought the need for greater analytic underpinnings and definition in patentable subject matter standards to the fore. In *Bilski*, the Court rejected a generally applicable subject matter test previously articulated by the Federal Circuit court, while refusing to provide a substitute of its own.[11] The Supreme Court's analysis of the invention before it (a method of managing risk in commodities transactions) involved pointing out that patentable inventions must not be just "abstract" ideas (as distinguished from more practical inventions) and then finding that the invention before it was too abstract.[12] The Court did not indicate why it thought that abstract ideas were not properly subject to patent rewards, nor did it describe what makes a design too abstract for patenting.[13] The Court's analysis seemed to reflect no underlying theory of patentable subject matter, which precluded a reasoned explanation of why the abstractness of an advance was important. Nor did the Court indicate what features of a useful advance would bear upon whether the advance constitutes patentable subject matter as distinct from an abstract idea.[14]

The lack of a viable standard articulated in *Bilski*—or even a viable approach to stating or implying such a standard—leaves future determinations of patentable subject matter largely unstructured and susceptible to wide variation. Lower federal courts will need to define new approaches to determining patentable subject matter that are general enough (and forward looking enough) to provide meaningful answers regarding the applicability of the patent system to technologies with characteristics that we cannot even imagine today. Ideally, in order for the incentives of the patent system to encourage and regulate the production of new advances, potential inventors of new technologies should have confidence that their advances will probably be inside or outside the patent system when they contemplate the development of their advances. Such confidence will

10. Bilski v. Kappos, 130 S. Ct. 3218 (2010).
11. *Id.* at 3225-29.
12. *Id.* at 3229-31.
13. *See id.*
14. *See id.*

clarify the rewards they can expect for successful development efforts. Absent some predictability that the patent system and associated patent rights and rewards will probably apply to a given type of advance, it is hard to see how the patent system can encourage and regulate the production of that type of advance. Hence, there is a particularly strong need for clearly articulated principles determining patent system scope and specifying (in a predictable manner) the outcomes of patentable subject matter analyses.

B. In Search of Theoretical Underpinnings

The Supreme Court's lack of a clear approach in defining the outer boundaries of the patent system stems from the Court's incomplete specification of the goals of the system. If it had a clearer picture of the incentive functions Congress intended to be served by the patent system, the Court would have had an easier time in *Bilski* in specifying criteria for determining when the system should apply. Without these theoretical underpinnings, the Court was at a loss to articulate generally applicable tests or principles for assessing whether patentable subject matter was present, and was instead reduced to repeatedly uttering the largely unhelpful notion that patentable advances must not be too "abstract."[15]

Had the Court considered the goals of the patent system in its analyses, it would probably have concluded that these goals remain poorly defined in both prior cases and academic literature. While there is some general sentiment—stated in both cases and articles—that the patent system is aimed at rewarding inventors for their work, there is little detail on how this reward system is to work and when.[16] Hence, the incomplete understanding of the reward logic of the patent system has curtailed the use of reward objectives as criteria for evaluating the patent system and for constructing and interpreting vague patent law standards, such as patentable subject matter tests. This article aims to improve this situation by providing a greater understanding of how and when the patent system may provide socially desirable rewards to inventors, creating incentives that both

15. *See id.* at 3223-25. The court also noted that there are two other well recognized grounds for excluding useful items from patentable subject matter. These involve advances that are laws of nature or physical phenomena. *Id.* Neither of these grounds for exclusion was addressed in the Supreme Court's analysis of the invention at issue in *Bilski*. *Id.*

16. *See, e.g.*, Diamond v. Chakrabarty, 447 U.S. 303, 307 (1980); Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 480 (1974); Donald F. Turner, *The Patent System and Competitive Policy*, 44 N.Y.U. L. REV. 450, 450-51 (1969).

encourage and regulate the production of valuable inventions.

Academic commentators have described a number of theories addressing the impacts of patent rights and patent enforcement.[17] This subsection summarizes these descriptions of patent impacts and assesses how these descriptions relate to the production of new advances by inventors. Overall, these patent theories provide a remarkably incomplete picture of how patents can beneficially influence the production of new advances and the allocation of scarce resources to such production.

C. A Brief Overview of Patent Law Theories

Theories describing the impacts of patents fall into three broad categories: 1) theories describing the impact of patents on the production of patent-eligible advances, 2) theories describing the impact of patents on the disclosure and use of advances after they are invented, and 3) theories describing the role of patents in dividing up and organizing activities surrounding the production of new advances.

This subsection summarizes these theories and their relation to the production of new innovations.

1. Reward Theory: Describing Patent Influences on Invention Production

Reward theory provides the oldest descriptions of the intended impacts of patents on inventions. This type of theory treats patents as means to influence the production of new advances.[18] The special rewards of patent rights are attached to non-obvious advances that are intellectual outliers in their respective technical fields in order to encourage more such advances and to diversify the technological approaches used in various fields.[19] Under this view, patents are

17. The discussion here focuses on instrumentalist theories of patents, derived from similar instrumentalist theories of property rights more generally. Instrumentalist theories of property justify property rights as means to serve some valuable end, with differences in the specific theories focusing mostly on the different end that can be served by various property rights. *See* Bell & Parchomovsky, *supra* note 9, at 534-35. Instrumentalist theories stand in contrast with formalistic theories, which hold that property rights are natural rights gained through some relationship of persons or their actions to specific items of property. *Id.* at 534-36. Under this view, the recognition of property interests is part of the law's efforts to maintain moral order. *Id.* at 541-42. In recent years, most property analyses have moved away from formalism and almost all property theories are based on some version of instrumentalism. *Id.* at 546-57.

18. *See, e.g.*, Turner, *supra* note 16, at 450-51.

19.

instrumental means to promote greater numbers of patent-eligible inventions.[20] Not just every type of invention is encouraged, however. Only the narrow category of non-obvious advances that are beyond the commonly held knowledge and skill in a particular field are specially promoted by the promise of patent rewards.[21]

Under this view, patents are utilitarian tools employed via government action to enlist potential inventors in serving societal needs.[22] Patent rights are recognized, according to reward theory, as means to induce increased attention to inventive efforts that address societal needs and desires.[23] Professor Donald F. Turner summarized these patent law ends and means as follows:

> The basic rationale of the patent system can be simply put. The economic case rests upon two propositions: first, that we should have more invention and innovation than our economic system would provide in the absence of special inducement; and second, that the granting of a statutory monopoly to inventors for a period of years is the best method of providing such special inducement.[24]

Using exclusive control over patented advances (via patent rights and enforcement) to generate economic rewards to inventors has several advantages over alternative means of encouraging inventive efforts.

First, the size of patent-influenced rewards for inventions is scaled to the value that the new inventions provide to invention users. Users will tend to pay amounts for access to new inventions up to (but not more than) the new value that they receive from using the inventions;[25] hence, the payments and gains that inventors can expect to realize from patent controls over access to inventions are scaled to the value that users perceive in the inventions. As noted by John

---

> The authority of Congress is exercised [in enacting patent laws] in the hope that "[t]he productive effort thereby fostered will have a positive effect on society through the introduction of new products and processes of manufacture into the economy, and the emanations by way of increased employment and better lives for our citizens."

Chakrabarty, 447 U.S. at 307 (quoting *Kewanee Oil*, 416 U.S. at 480).

20. *See id.*

21. *See* 35 U.S.C. § 103 (2006).

22. *See, e.g.*, Turner, *supra* note 16, at 450-51.

23. *Id.*

24. *Id.*

25. *See* JOHN STUART MILL, PRINCIPLES OF POLITICAL ECONOMY 451-52 (Prometheus Books 2004) (1900).

Stuart Mill, the grant of an "exclusive privilege" to control an invention under a patent tends to provide a reward to the inventor of an advance that is proportionate to the "usefulness" of the advance.[26]

Second, patent-induced rewards to inventors have the added advantage that the rewards are paid by the specific users of inventions. These rewards are paid via patent-elevated prices charged for patented goods and services or via royalties paid for licensed technologies.[27] These sorts of payments place the costs of invention incentives administered through patent rights upon invention users, the parties who benefit from the inventive efforts being fostered. This matching of burdens to the benefited parties was recognized by Jeremy Bentham in his early descriptions of patent system dynamics. He saw patent-influenced payments to inventors as the equivalents of bonuses paid by invention users to successful inventors who have aided the users' activities. The bonuses are paid out of the public gains and benefits achieved by the discoveries. Because exclusive patent rights produce incentives to inventors from the gains the inventors achieve, Bentham saw grants of exclusive patent privileges as "the best proportioned, the most natural, and the least burdensome [means of] produc[ing] an infinite effect and cost[ing] nothing."[28]

Third, rewards to inventors paid through patent-influenced purchase prices or licensing royalties have the added advantage of providing self-executing mechanisms for paying rewards to inventors. These rewards flow directly from the demand for new inventions coupled with patent rights, which ensure that patent rights holders are the sole parties capable of providing certain technological solutions to societal needs (by providing products or services based on a patented advance).[29] The rewards system is implemented through private market processes and does not depend on government reward-implementing actions or the discretion of government officials as would be the case in a system rewarding inventors though governmental bonuses or payments for the creation of useful inventions.[30]

Reward theory contemplates impacts of future patent rights on

---

26. *Id.*

27. 3 JEREMY BENTHAM, *A Manual of Political Economy, in* THE WORKS OF JEREMY BENTHAM 31, 71 (1962).

28. *Id.*

29. MILL, *supra* note 25, at 451-52.

30. *Id.*

inventors before inventions are made, with the promise of patent rights and rewards attracting potential inventors to work on new, inventive projects. However, the details of how this should occur—and the extent of invention production that should be encouraged—are not specified by the basic forms of reward theory that have been articulated to date.

The enhanced reward theory described in this article explains patent rights and their impacts in terms of invention production decision making. Patent rights are treated here as means for reconciling invention production with user demands for new inventions and for allocating scarce resources to invention projects rather than to other competing uses of the same resources. The ways that patent rights should be shaped to achieve these ends are addressed in Part III of this article.

2. Theories Describing the Influence of Patents on Completed Inventions

Other patent law theories seek to describe the influence of patents on actions regarding inventions after the inventions are already in existence. As such, they are not concerned with rates of production of inventions, but rather with seeing that society receives the maximum gains from already realized inventions. Disclosure theory sees patent rights as sources of rewards encouraging parties who have already made useful inventions to disclose those inventions rather than only using the inventions on a narrow scale for personal advantage or under secrecy constraints.[31] Rent dissipation theory treats patents as means to discourage duplicative inventive efforts once one party has completed the discovery of an invention.[32] Prospect theory views patents as means to encourage patent holders to maximize their efforts to find applications for inventions and to thereby extend the societal uses of the inventions.[33] Commercialization theory sees patents as means to encourage patent holders to engage in commercially effective product design, manufacturing, and marketing efforts that spread the use of products and services incorporating patented advances.[34]

---

31. *See infra* Part I.C.2.a.
32. *See infra* Part I.C.2.b.
33. *See infra* Part I.C.2.c.
34. *See infra* Part I.C.2.d.

a. Disclosure Theory

Disclosure theory views patent law as a tool for encouraging disclosures of useful inventions by successful inventors who would, at least in many cases, otherwise keep their inventions secret.[35] A successful inventor might, in the absence of patent rewards, keep an invention secret for a number of reasons, including personal disinterest in the useful applications of the invention (for example, because the inventor is simply not in the type of business where the invention would be useful and sees no benefit in efforts to disclose and popularize the invention), a desire to withhold the invention from competitors while using it in secret to the inventor's personal advantage, or a goal of spreading the use of the invention under secrecy constraints (perhaps implemented through trade secret controls) and commercializing the advance by charging for access in this way without granting complete public access to the invention.

Disclosure theory sees patent rights as part of a bargain aimed at overcoming these considerations and encouraging full public disclosures about such advances.[36] An inventor is required, in order to obtain a patent, to describe not only the functional features of his invention, but also how to make and use the invention.[37] These features must be included in a patent application and will be disclosed to the public either when the application is published or when a patent issues. In exchange for these public disclosures the inventor is granted valuable patent rights. In essence, the inventor is paid through these rights for giving up the advantages of secrecy about his invention.[38]

The increased disclosures that result from this type of patent-influenced bargain are several fold. First, information disclosed in a patent application or issued patent may enhance subsequent research and innovation.[39] Understanding the workings of the patented invention may provide insights into how to approach other innovations that do not incorporate the patented design and that therefore may be used immediately without infringing the patent on

35. *See* Kevin Emerson Collins, *Propertizing Thought*, 60 SMU L. REV. 317, 357-58 (2007).

36. *See id.*

37. *See* 35 U.S.C. § 112 (2006).

38. *See* Pfaff v. Wells Elecs., Inc., 525 U.S. 55, 63-64 (1998); Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 150-51 (1989).

39. *See* 3 DONALD S. CHISUM, CHISUM ON PATENTS: A TREATISE ON THE LAW OF PATENTABILITY, VALIDITY AND INFRINGEMENT § 7.01 (2010).

that design. Second, disclosures of a patented design may provide the starting point for further research to improve the patented design (although commercialization of these improvements will require a license from the patent holder). Third, upon expiration of the patent, free availability of the disclosed invention will allow widespread productive use of the advance involved, as well as free incorporation of this advance in further research and additional inventions.

Disclosure theory treats patent rights as a means to overcome secrecy barriers to the use of inventions without asserting that patent rights have any influence over the numbers of inventions created. It is doubtlessly true that the public generally will gain from widespread information about new advances and (at least eventually) the free availability for use of those advances. Hence, the disclosures about completed advances that patent law promotes are undoubtedly important.

b. Rent Dissipation Theory

Rent dissipation theory describes the impact of patents in helping to discourage duplicative efforts to develop similar inventions.[40] When a new advance is developed, society frequently gains increased utility (through the use of the advance) over and above the cost of developing the invention.[41] This net gain from an invention is sometimes referred to as the "monopoly rent" associated with the invention.[42] Society gains most where this net gain is maximized—that is, where the gain realized from an invention exceeds the costs of developing the invention by as much as possible.[43] Where several parties work on the same invention (or similarly functional inventions) in parallel, resulting in only one commercially successful, widely used invention, monopoly rents (that is, net societal gains from an invention) can be squandered or "dissipated" because the costs of the multiple inventive efforts only result in one socially valuable invention.

Rent dissipation theory views patent rights as creating incentives to promote early disclosures of inventions and to discourage competing inventors from continuing duplicative efforts at the earliest

40. *See* Mark F. Grady & Jay I. Alexander, *Patent Law and Rent Dissipation*, 78 VA. L. REV. 305, 308 (1992).

41. *See id.*

42. *See id.*

43. *See id.*

possible point.[44] Patents minimize the dissipation of monopoly rents (and the squandering of net societal gains from inventions) because patent rights tend to cut off the efforts of additional inventors once they realize that one party has "won" the race to develop a successful invention and has gained a patent on the invention. By disclosing the features of a particular invention in a patent and gaining exclusive control over the invention through patent rights, a patent holder signals to other potential inventors that continuing their competing efforts to develop the same type of invention will be wasteful because the patent holder will be able to control who can make, use, or sell the patented invention for the life of the patent and will be able to bar competing inventors from using the fruits of their own efforts. An issued patent becomes a means to discourage competing inventive efforts and, thereby, to reduce total invention production costs. Reduction of these costs minimizes monopoly rent dissipation. Society's net benefits from an invention are increased to the extent that duplicative efforts of multiple inventors are minimized.

c. Prospect Theory

Prospect theory focuses on the potential role of patents in encouraging patent holders to explore or "prospect" for applications of an invention once at least one rudimentary version of the invention has been produced.[45] By giving an inventor control over subsequent uses of a patented invention and an interest in the commercial success of later uses, the patent holder is encouraged to prospect for additional uses of the invention much as a miner is encouraged to prospect for ore in a particular plot of ground by being given exclusive control over a mining claim. Patent rights encourage patent holders to prospect for the full range of socially valuable applications of their patented inventions so as to maximize their own commercial stake in the inventions. To the extent that this prospecting for additional applications is successful, the rights holders will expand the use and social value of the patented inventions.

Prospect theory treats the discovery of a patented invention as the starting point for a product development process and patent rights as means to encourage a patent holder to invest additional resources in the efficient and effective development of a raw invention into

44. *See id.* at 316-17; Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 278 (1977).

45. Kitch, *supra* note 44, at 274-80.

useful products across as many socially valuable applications of the invention as possible.[46] A patent holder will also be able to prevent duplicative efforts by multiple parties to prospect for new applications of patented advances, thereby preventing rent dissipation in the prospecting phase of realizing social value from a patented advance.

d. Commercialization Theory

Commercialization theory is similar to prospect theory in that it focuses on encouraging patent holders to take actions after an invention is made that tend to increase the societal use and utility of the invention.[47] While prospect theory sees patents as inducements for patent holders to prospect for additional applications of patented advances,[48] commercialization theory treats patents as inducements for patent holders to take other commercial actions to produce products based on patented designs and to bring the products to market, thereby putting the patented designs into use by more parties and increasing the total societal gain from the advances.[49] Patents, as seen by commercialization theory, ensure that inventors (or their successors in interest in patent ownership) are encouraged to follow through on inventions with effective commercial efforts that popularize products and services based on the inventions.[50]

3. Specialization Theory: Describing the Influence of Patents on the Organization of Innovative Efforts

At least one additional theory sees patents as means to aid parties in dividing and organizing engineering and business activities surrounding patented advances and related products. Specialization theory argues that patents help parties to organize the work of multiple specialists in bringing innovative products and services to the public.[51] This sort of division and specialization of work on patent-eligible advances can produce improvements in innovation processes by realizing specialization gains at various stages of

46. *Id.*

47. *See* F. Scott Kieff, *Property Rights and Property Rules for Commercializing Inventions*, 85 MINN. L. REV. 697, 732-36 (2001).

48. *See* Kitch, *supra* note 44, at 278.

49. *See* Kieff, *supra* note 47, at 732-36.

50. *Id.*

51. *See generally* Jonathan M. Barnett, *Intellectual Property as a Law of Organization* (USC Center in Law, Economics and Organization Research Paper No. C10-10 2010), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1623565.

innovation and commercialization processes.[52]

Specialization theory treats patents as aids in allocating work on patented advances among specialists, while giving each specialist a stake in the success of the overall engineering and commercialization efforts regarding the advances.[53] A patent gives its holder a stake in the ultimate usefulness and commercial value of products based on the patented innovation even if the party holding the patent will not be the party to sell related products to the public.[54] Patents can facilitate the separation and specialization of work in discovering, developing, and commercializing new inventions and in allocating components of this work to businesses that are efficient in scope and effective in operation, thereby maximizing the benefits of specialization effects gained through work experience and research.[55]

Patents aid parties in separating work on different aspects of patented advances by providing means to transfer ownership interests in the commercial success of patented advances and to thereby encourage each of several specialized parties to apply their particular capabilities towards advancing the overall practical success and social propagation of new advances.[56] By aiding the division and distribution of work among separate organizations (each with work content that best suits the skills of those parties involved in a project and the potential specialization economies that are available in a particular invention development setting), patent rights facilitate diverse choices about organizational groupings of work on patentable advances. Patent rights allow parties to pass on the fruits of specialized work in early phases of the development of patented advances to other parties who will work on later phases. At each phase, the ownership of patent rights in an invention gives a party taking specialized actions concerning the invention a direct self-interest in doing that party's best work to promote the eventual commercial success of the invention. In this way, patent rights facilitate the separation of work into efficient work units and organizational contexts where size efficiencies and specialization effects can be used advantageously, while still ensuring that persons working on discovering inventions and on post-invention

52. *See generally id.*
53. *See generally id.*
54. *See generally id.*
55. *See id.*
56. *See generally id.*

commercialization activities are encouraged to do their best work in accomplishing partial steps towards the full commercialization and broad public use of patented advances.

The benefits that patent rights bring to the organization and specialization of work on patentable advances have been summarized by Jonathan Barnett, as follows:

> Patents enable innovators to make efficient selections of firm scope by transacting with least-cost suppliers of commercialization inputs. These expanded transactional opportunities reduce the minimum size of the market into which any innovator—or the supplier of any other technological or production input—can attempt entry. Disaggregation of the innovation and commercialization process then induces the formation of secondary markets in disembodied technology inputs. These organizational effects over transactional, firm and market structure generate specialization economies that minimize innovation and commercialization costs, which in turn exerts incentive effects consistent with the standard thesis and market growth effects that extend beyond it.[57]

4. Reconciling the Multiple Theoretic Approaches: The Fundamental Importance of Reward Theory

The theories of patent rights summarized here differ significantly in their approach to invention production. Only reward theory and specialization theory contemplate and seek to explain potential impacts of patents on numbers and types of inventions. All of the other theories assume that inventions are produced largely outside of the influence of patents and seek to explain how patents can impact the further steps used to achieve societal advantages from the inventions.

Patent theories other than reward theory generally take the presence of a patented advance as a given—that is, as a starting point in their accounts of patent influence—and then explain the possible impacts of patent rights on the subsequent implementation or use of the invention. The types of post-invention benefits purportedly achieved by patent rights are seen differently in the different theories.

Disclosure theory assumes that some inventions (such as those potentially protected by trade secrets) may be commercialized by inventors in secret absent patent protections. The result of secret

57. *See generally id.*

commercialization (or no commercialization at all) will be that information about the advances will be bottled up with the inventors and will not enable broader uses and understanding of the inventions by all interested parties. In this theory, patent rights are the means to encourage disclosures of useful inventions, resulting in broader distribution of information about the inventions and enhanced availability of the inventions for use (at least after patent rights have expired). However, this account assumes that the relevant inventions for which disclosure is an issue already exist; it says nothing about producing more or different inventions under the influence of patent rights.

Rent dissipation theory sees the merit of patents in preventing post-invention waste due to duplicative inventive efforts that will achieve no additional societal benefits over the results of the first successful effort to produce an invention. This theory treats the first to discover an advance as the "winner" of the race to make this discovery and then seeks to cut off duplicative efforts of other inventors to produce the same results a second time. The aim is to gain social efficiency by preventing these duplicative efforts. There is no aim in this theory to explain why the first inventor discovered his or her invention or to encourage more of these first discoveries.

Prospect theory is aimed at describing how patents can encourage increased efforts to explore and design the full range of applications made possible by an already completed invention. The goal is to realize the full measure of societal gains made possible by a given invention by carrying forward the invention into the broadest possible set of applications and user gains. The aim is to ensure that the full societal potential of a new advance is realized, for the benefit of both the inventor and potential invention users. Prospect theory does not, however, purport to explain how the promise of a patent may produce more patent-eligible advances susceptible to prospecting.

Commercialization theory likewise focuses on maximizing post-invention use of new advances and on avoiding inefficiencies from failing to gain the greatest possible social advantages from already completed inventions. In commercialization theory, the focus is on the impact of patents on commercial efforts to create, manufacture, and market viable products. These efforts put patented advances into the hands of numerous consumers and thereby produce large-scale societal gains through widespread use of the patented advances. Commercialization theory contends that patents are important in

promoting effective commercialization of patented advances, in part because such commercialization may be particularly difficult and uncertain. Difficult commercial challenges concerning patented advances may stem from the new and sometimes very unusual features of these advances and the potential difficulties of new product design, manufacturing, and marketing that popularizing the patented advances may entail. Commercialization theory does not, however, purport to explain how patents affect the making of new advances and, hence, does not say anything about how patents influence the number or nature of new advances. This theory only seeks to describe how patents can improve commercial follow through on patent-eligible advances, to the gain of both patent holders and product users.

Specialization theory is the one exception among the theories discussed here in that it describes some potential impacts of patent rights on the number and nature of patent-eligible advances. Specialization theory argues that patents aid firms in dividing and organizing work on the production and use of patent-eligible advances. Such a process of work division among firms (linked by patent rights giving each firm a stake in the commercial success of the project they are contributing to) can achieve specialization benefits that produce different results than if the same tasks were undertaken without the ability to organize work in this manner. Specialization theory describes a means whereby patents can change the number or nature of advances that are produced under the influence of patents. Inventions will change in number or nature if the specialization effects promoted by the availability of patent rights change inventive processes to produce different invention outputs than would prevail in the absence of patents and the specialization effects patents promote. Certain types of specialization effects enhanced by patents may produce more inventions or different types of inventions than inventive processes without these specialization effects.

Specialization theory is not focused just on these invention production effects, however. Some work divisions and associated specialization effects that patents promote relate to commercialization steps after the creation of a new invention. These sorts of specialization effects will realize benefits in the post-invention phase—that is, in the exploration for applications of the invention or in the further steps needed to commercialize the invention. In these post-invention versions of specialization promoted by patents, the benefits described by specialization theory are more like those

associated with prospect and commercialization theories. In addressing these types of post-invention actions, specialization theory only explains how patents help us to be more efficient in our use of patent-eligible advances, not how to produce more such advances.

Only reward theory focuses primarily on the production of more inventions and is, in this sense, a more fundamentally important type of patent theory than the others. Reward theory describes how patents can encourage more (or more useful) inventions. It explains how patents can produce more of the things—inventions—on which the other theories operate and, in this sense, is a root theory from which other theories build or improve. If aspects of patent law can produce more (or better) inventions, as reward theory posits, other aspects can (perhaps) go further and achieve the additional benefits posited by the other theories of patent law. However, if there are no inventions to operate on, the post-invention benefits described by the other theories of patent law are of little or no importance. These additional theories are, in this sense, supplemental to reward theory and the production of patent-eligible inventions.

Because reward theory attempts to describe how the patent system can produce more inventions, it is in a sense a more fundamental theory of patent rights than these other theories which focus primarily on post-invention efficiency. If patents can, as reward theory suggests, produce more patent-eligible advances, then the further impacts of patents as described by the other theories can be obtained over a wider range of inventions. If an accurate account of how patents influence the production of more patent-eligible advances can be described and used to shape the patent system to this end, the result will be a bigger, more socially responsive, and beneficial set of patent-influenced inventions, towards which the additional advantages described by the other patent theories can also be pursued. Because reward theory and its impacts on invention production have this type of especially fundamental impact on the number and nature of patent-eligible inventions and the associated social benefits such inventions can realize, the reformulation and refinement of reward theory concepts of the sort attempted in this article are especially worth our attention.

II. PLACING PATENT REWARDS IN THE CONTEXT OF PROPERTY LAW THEORY

A. Basic Property Law Goals: Long-Term Perspectives and Stable Ownership

Property laws permit property owners to control the use of property, potentially for extended periods.[58] Property rights generally achieve this by affording property owners the right to exclude parties from their property and to selectively allow access to the property in exchange for payments or other compensation.[59] This right to exclude, coupled with the right to transfer property ownership and the right to exclude to other parties, gives an owner the reassurance that they will control the use of an item of property not only now, but also in the future. This long-term perspective is a key goal of property law.[60] A property owner can take a long-term view of what is an advantageous present use of property, knowing that the owner (or transferees of the owner's property interest) will benefit in the future from the good husbanding of property now, as well as from investments in changes and improvements to that property which will enhance its future value. The encouragement of property owners to adopt long-term perspectives regarding the use of property is a primary goal of property laws in general and, as discussed later in this section, of intellectual property laws in particular.

1. Property Laws and Long-Term Use Perspectives

Robert C. Ellickson has recognized the importance of the long-term perspectives on asset use that property rights can instill. In discussing the implications of property rights in the context of real property ownership, Ellickson emphasized the following perspective-setting impacts of property rights:

> Although the assertion may seem counterintuitive, the key to land conservation is to bestow upon living persons property rights that extend perpetually into the future. The current market value of a fee in Blackacre is the discounted present value of the eternal stream of rights and duties that attach to Blackacre. A rational and

58. *See* Robert C. Ellickson, *Property in Land*, 102 YALE L.J. 1315, 1369 (1993).

59. *See id.* at 1353.

60. *See id.* at 1368.

> self-interested fee owner therefore adopts a[n] infinite planning horizon when considering how to use his parcel, and is spurred to install cost-justified permanent improvements and to avoid premature exploitation of resources. The fee simple in land cleverly harnesses human selfishness to the cause of altruism toward the unborn, a group not noted for its political clout or bargaining power.[61]

The transferability of ownership interests concerning an item of property gives the owner a long-term perspective on how the property should best be used since optimizing this use will maximize the transfer value of the property, even if the present owner will not be the only one to use the property over the long term.[62]

While his description does not separate the two, Ellickson's account addresses the impacts of property rights on both the maximization of the value of property in its present state (referring to "land conservation" and "the avoidance of premature exploitation of resources") and in encouraging improvements in property features (referring to the impact of property rights in "spurr[ing owners] to install cost-justified permanent improvements" to owner-controlled property). Property rights—and the long-term perspectives they encourage—have important impacts in both the present use and future-focused improvement of assets.

2. Additional Advantages of Stable Ownership

More recently, Abraham Bell and Gideon Parchomovsky have expanded on the importance of property laws in creating long-term owner control regarding the use and improvement of items of property. They argue that protecting stable ownership is the unifying theme of property law.[63] In their view:

> [P]roperty law as a legal institution is organized around creating and defending the value inherent in stable ownership. Property law both recognizes and helps create stable relationships between persons and assets, allowing owners to extract utility that is otherwise unavailable. Adopting this focus enables us to recast many of the key insights of the extant property literature and demonstrate that these insights can form a coherent theory of

61. *Id.* at 1369.

62. *See id.* at 1368.

63. *See* Bell & Parchomovsky, *supra* note 9, at 531.

property.[64]

Beyond just promoting the management of an asset so as to maximize the long-term value of the asset, Bell and Parchomovsky see several other advantages in property rights ensuring stable ownership of assets. These additional advantages of legal protections promoting stable ownership include:

1) Facilitating more open and frequent uses of assets, as owners need not rely only on secrecy and physical possession as means to retain control over assets;

2) Enabling the decoupling of ownership and possession (or use/operation), thereby increasing the use value of assets by making possible the temporary transfer of assets to higher-value users who cannot afford to purchase the assets (through actions like rentals or even informal arrangements like cab rides);

3) Increasing value from asset use by encouraging learning about how best to use an asset during long-term association with the asset; and

4) Encouraging parties to count on the availability of an asset over the long term and to assemble complementary assets that are productive in combination with the asset.[65]

Patents have the potential to achieve parallel benefits in all four of these areas. The counterpart benefits from enforcing limitations on the making, using, and selling of patented inventions involve:

1) Encouraging uses of patented advances in contexts that do not rely on secrecy or physical possession of items embodying patented designs as means to retain control over the designs and prevent free rider usage;

2) Enabling the decoupling of invention ownership and various forms of invention use, thereby increasing the ways that subcomponents of rights to use patented inventions may be granted through coordinated and narrowly targeted licenses in economically desirable quanta that match the needs of narrowly focused invention users;

3) Increasing value from invention use by encouraging patent holders to invest resources in learning about how best to use an

64. *Id*. at 538.

65. *Id*. at 556-58.

invention and how to produce, deliver, and market related products; and

4) Encouraging parties to count on the exclusive commercialization of an invention over a long period (that is, the patent term) and to assemble complementary assets that are productive in combination with the patented invention.

3. Reconciling Patent Law Goals with Property Law Goals Generally

As the analyses of property law purposes by scholars like these demonstrate, rights to exclude parties from property are generally granted for one or both of two purposes.

First, rights to control access to property ensure that the present use of the property can be managed to its greatest value by property owners, adopting the use with the greatest value and avoiding wasteful overuse or other value-decreasing uses.[66] This purpose of property controls involves the efficient use of existing property.

Second, rights to control access to property are also imposed to induce parties to create new property and to improve existing property.[67] In this regard, the promise of controls once new property is constructed or improved provides a means to gain value from the newly created or improved property and to encourage the production of such property. As Ellickson noted, a rational owner will be encouraged by property rights controlling use of property to add cost-justified improvements that enhance the value of the property.[68] In this regard, property controls serve an asset production incentivizing function encouraging asset production.

Patent law controls limiting access and use of patented inventions are unusual forms of property controls as they serve the second of these property law purposes, but not the first. Use of patented ideas is not rivalrous and, consequently, not in need of property controls to ensure the efficient use of patented ideas in their most valued ways. However, the production of new inventions is rivalrous in that it involves the use of scarce resources that are devoted to invention projects in rivalry with other potential uses of the same resources. This means that patent controls and rewards for

66. *See* Ellickson, *supra* note 58, at 1369.

67. *See id.*

68. *Id.*

inventions can serve valuable ends in regulating the allocation of scarce resources to the production of new property in the form of new inventions. Patents not only attract resources to socially valuable invention projects in socially desirable ways, they tend to encourage the production of patentable inventions at socially desirable levels. They are, in short, key regulators of the productions of inventions and, as such, serving a traditional property law role with respect to property production and resource usage.

The remainder of this section summarizes why patent controls over the use of patented advances serve some, but not all, of the traditional functions of property law.

B. Property Justifications for Controls to Ensure Efficient Use of Consumable Assets Do Not Apply to Patents

Patents give their owners the ability to exclude other parties from the making, using, and selling of items or services incorporating patented inventions.[69] These rights limit the use of patented inventions by unauthorized parties.[70] Such restrictions on the use of invention ideas are not needed to avoid overconsumption or other inefficient use of the ideas since the use of such ideas by one party does not diminish the value of the ideas if used by another.[71] Nor are restrictions on the use of design ideas needed to ensure that these ideas are used by the parties who value the ideas the most. Since design ideas can be used by multiple parties without limiting each other's use, the parties who value the ideas most should be able to use them along with all other interested parties.

This means that patent controls over access to patented designs—that is, patent controls for ideas about designs of useful items or processes—are not needed to ensure efficient use of the designs. Indeed, by elevating the cost of use for patented designs,

---

69. 35 U.S.C. § 271(a) (2006) ("[W]hoever without authority makes, uses, offers to sell, or sells any patented invention, within the United States or imports into the United States any patented invention during the term of the patent therefor, infringes the patent.").

70. *Id.*

71. The use of an idea does not diminish its value to another because an idea is shared, not consumed. This is also true of the ideas in literary works. For example, Tom G. Palmer has observed that the literature is of value to numerous readers without any consumption or diminishment effect because "there is one *Othello* for all of us, rather than one *Othello* for each of us, or even one for each of our separate readings or viewings of the play." Tom G. Palmer, *Are Patents and Copyrights Morally Justified? The Philosophy of Property Rights and Ideal Objects*, 13 HARV. J.L & PUB. POL'Y 817, 846 (1990).

patent controls may diminish the overall use of patented designs below levels that would be socially desirable in the absence of patent-imposed costs. Hence, focusing on uses of designs for useful items and processes, property controls over patented designs would appear to be without justification and potentially harmful as undesirable burdens on valuable applications of useful designs. At least, patented ideas lack the beneficial impacts on efficient use of assets that provide key justifications for property controls in other property domains.[72]

C. Property Justifications for Controls to Ensure Asset Improvements Do Apply to Patents

However, control over use of patented designs may be justified as means to encourage the production of new and improved designs and to attract the allocation of scarce resources to this type of design production. In this respect, the production-enhancing ends served by property controls in the patent field are similar to the asset-enhancing benefits of property rights generally. The focus of property rights in promoting improvements to assets (in both the physical world of personal property and the intellectual world of patents) is on mediating the allocation of scarce resources to asset improvements.

Many accounts of property rights (such as Professor Ellickson's account quoted earlier in this article[73]) describe the potential impacts of patent rights in encouraging improvements to existing properties, such as the encouragement that real property ownership gives to property owners to consider and advance building projects on their land as means to improve the property characteristics and to maximize the long-term value of the property. However, in personal property and intellectual property settings, the promise of long-term property controls creates incentives to "improve" property by creating it. The production of new goods or patent-eligible advances is an act of property creation which is encouraged and mediated by the expectation that the party producing a new item will have long-term control over it and will be able to realize gains from use of the item over a substantial period or from the sale of the item to another who will make such long-term use. This promise of future controls

72. *See* Peter S. Menell, *The Property Rights Movement's Embrace of Intellectual Property: True Love or Doomed Relationship?*, 34 ECOLOGY L.Q. 713, 726 (2007) ("Unlike tangible goods, knowledge and creative works are public goods in the sense that their use is nonrivalrous. One agent's use does not limit another agent's use.").

73. *See supra* text accompanying note 61.

mediates the production of new goods and new patent-eligible advances, just as it mediates the construction of new improvements to parcels of real property. In all of these instances, the promise of future gains from the new property or improvements establish criteria for determining whether it is worth devoting scarce resources to the creation of the new property or improvements.

Because scarce resources are significant in the production of both patented advances and other types of assets, property controls are important in promoting production in all of these areas. The purposes served by patent controls in this respect are much like the purposes served by property interests more generally. Some of our knowledge about how property interests serve to promote asset production in other settings can aid us in analyzing the impacts of patent policies on allocations of scarce resources and in proposing desirable patent law reforms.

III. THE PRIVATE PRODUCTION OF PUBLIC GOODS: A DEMSETZIAN FRAMEWORK FOR EVALUATING PATENT RIGHTS

A. Property Controls to Attract Resources to Public Goods Production

The analysis here builds on the work of Harold Demsetz regarding the private production of public goods and the potential role of property controls over access to public goods as means to encourage such production. Demsetz's conclusions about the desirability of property controls for public goods (including public goods like patented ideas) incorporate three key insights: 1) previously defined economic principles for analyzing the production of joint supply goods (that is, goods for which there are at least two purposes and two different types of consumer demands) also provide valuable tools for analyzing the production of public goods, 2) the production of public goods through private means (including the production of ideas defining patent-eligible inventions) can be promoted by restricting access to public goods and precluding non-purchasers from using the goods, and 3) such a system of exclusion and charging for access can create a form of market for the production of public goods that both attracts scarce inputs to the production of public goods and helps to ensure that pubic goods are produced in

socially optimal amounts.[74] These conclusions were all described in Demsetz's landmark article "The Private Production of Public Goods,"[75] first published in 1970, but now somewhat neglected, particularly in intellectual property studies.[76]

This section describes the major tenets of Demsetz's analysis in this area. The next section shows how these views about public goods production can be extended to define an improved patent reward theory—that is, a theory describing how patent rights incentivize and regulate the private production of public goods in the form of additional patent-eligible advances.

1. Constructs for Evaluating Joint Supply Goods Production

Demsetz's first insight is stated clearly at the outset of his analysis: "The allocation of resources to the production of public goods can be understood with the aid of the model formulated long ago by Alfred Marshall for the analysis of joint supply."[77] A joint supply good is one for which there are two or more uses (usually of different parts or features) and, hence, two or more demands.[78]

---

74. *See* Harold Demsetz, *The Private Production of Public Goods*, 13 J.L. & ECON. 293 (1970).

75. *Id.*

76. A Westlaw search showed only 107 articles in the JLR database (addressing all areas of law, not just patent law or intellectual property) that have cited Demsetz's article in the 40 years that it was written. One key exception to the general neglect of Demsetz's views on public goods production among intellectual property scholars is the work of Brett M. Frischmann, who has analyzed (and criticized) Demsetz's views in the context of copyright law. *See* Brett M. Frischmann, *Evaluating the Demsetzian Trend in Copyright Law*, 3 REV. L. & ECON. 649 (2007). *See also* Harold Demsetz, *Frischmann's View of "Toward a Theory of Property Rights"*, 4 REV. L. & ECON. 127 (2008) (responding to Frischmann's criticisms).

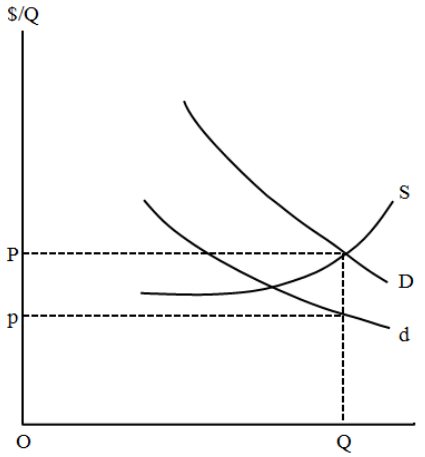
77. Demsetz, *supra* note 74, at 293. The insights of Alfred Marshall referred to by Demsetz were described in Marshall's book, PRINCIPLES OF ECONOMICS. I ALFRED MARSHALL, PRINCIPLES OF ECONOMICS 436 (1890). Demsetz recognized that he was not the only analyst to see the relationship between the analysis of joint supply goods production and the analysis of public goods production. He notes that some earlier analysts had mentioned the similarity of these types of production in passing, but had not explored the implications of this equivalency for public goods production to the degree that Demsetz pursued in his work described in this article. *See* Demsetz, *supra* note 74, at 293 n.1 (describing the work of other analysts who had, as of 1970, recognized the similarities between joint supply goods production and public goods production).

78. *See* Demsetz, *supra* note 74, at 293-94. The separately desired components or features of a joint supply good or item are typically produced simultaneously in one act of production. For example, the construction of a dam creating a reservoir—a single production process creating a joint supply item—can satisfy separate demands for irrigation water, hydropower, flood control benefits, and recreation. *See* A. ALLAN SCHMID, PROPERTY, POWER,

Marshall's insights were that these two demands would operate somewhat independently in sales of components of a joint supply good, but would combine to produce a hybrid demand function affecting and incentivizing the upstream production of the joint supply good.[79]

A simple example used by Demsetz in his analysis illustrates Marshall's earlier ideas about the production of joint supply goods. Steers are produced with the expectation that a particular steer will ultimately be used to supply both meat and a hide.[80] The production of steers for slaughter is therefore influenced by the demand for both meat and hides.[81] Under competitive conditions, the rate at which steers are raised and slaughtered is determined by an aggregate demand function (and associated marginal returns for slaughtering one more steer) which reflects the sum of the demands and returns for meat and hides.[82]

Figure 1.



In standard format, the supply and demand characteristics for consumption of meat and hides and the supply of steers are as shown in Figure 1.[83] The x-axis in this figure reflects the quantity of items

AND PUBLIC CHOICE: AN INQUIRY INTO LAW AND ECONOMICS 85 (2d ed. 1987).

79. *See* INGRID RIMA, DEVELOPMENT OF ECONOMIC ANALYSIS 335 (7th ed. 2009).

80. *See* Demsetz, *supra* note 74, at 293-94.

81. *Id.*

82. *Id.*

83. This figure is a reproduction of a graph from Demsetz's article *The Private*

and the y-axis reflects the price per unit quantity. The curve d reflects the demand for hides and the curve D is the sum of the demands for hides and meat (the demand curve for meat alone is not shown, but would equal D-d at every quantity Q). Curve S reflects the number of steers that suppliers are willing to produce at various price points. The equilibrium price-quantity combination is determined by both the aggregate demand (reflecting the component demands for both hides and meat) and the marginal costs of producing additional steers. The equilibrium point is represented by the intersection of the S and D curves at price P and quantity Q.

In general, this analysis reflects standard (and fairly basic) economic analysis of supply and demand under circumstances of competition. The one key addition that is special to joint supply goods is that the separate market demands for components of the joint supply good can be translated into a single demand function influencing production and supply of the joint supply good. Competition in these circumstances maximizes the sum of welfare for users of the component parts (although it may not provide optimal levels of hides or meat looking at either of these alone).

Another way to look at this is that the compensative system economizes on opportunity costs. A competitive system of the sort shown in Figure 1 ensures that net societal costs (taking into account opportunity costs) are minimized in the production of steers as joint supply goods. If production is at levels below the P-Q point, the aggregate opportunity costs to hide and meat consumers of forgoing one more unit of hides and meat will be greater than the marginal cost of producing one more steer and avoiding these opportunity costs. Where this is the case, competitive processes will tend to ensure that the opportunity costs will be avoided because an additional steer will tend to be produced to fill this demand at a price that is on the S curve and acceptable to steer producers. In this manner, the aggregate demand for a joint supply good tends to discourage undersupply of that good, with the optimal supply level determined by the minimization of total opportunity costs to users of the components of the joint supply goods.

Competitive processes can ensure against oversupply of joint supply goods in a similar fashion. The equilibrium reached at P-Q is one where the missed opportunity value of having one more hide and

---

Production of Public Goods, but is reinterpreted here in terms of its implications for the production of patented advances. *Id.* at 294.

meat unit (taken together) is less to hide and meat users than the marginal costs of producing that additional steer. A socially advantageous equilibrium is achieved because the opportunity costs of forgoing additional production are less than the marginal costs of proceeding with additional production.

2. Extension of the Joint Supply Model to Public Goods Production

a. Promoting Public Goods Production Through Access Charges

Demsetz saw the production of public goods as having key similarities to the production of joint supply goods.[84] Provided that potential users of public goods can be excluded from use unless they pay for access to the public goods, a system of use payments (or use-related payments) can establish demand functions for the supply of public goods.[85] These demands will aggregate across multiple users of the goods to produce an overall combined demand for the goods.[86] The combined demand (reflecting the willingness of all interested parties to pay for access to the goods) will encourage and incentivize the production of public goods at optimal levels, much as the aggregate demand for joint supply goods encourages the production of joint supply goods at socially optimal levels.[87]

A word on nomenclature is in order here. The term "public goods" is sometimes used in two different senses. Some analysts (including Demsetz) use the term to refer to non-rivalrous items (which are not diminished in value with additional use) without regard to whether the goods are also excludable (in that non-purchasers or other unauthorized parties *can* be excluded from use).[88] In this usage, the term "public goods" refers to goods that are non-rivalrous, but possibly excludable. To other analysts, the term "public goods" is used in a stricter sense to refer only to goods that are both non-rivalrous (meaning that they can be used by additional parties

84. The key similarity is that both public goods and joint supply goods have multiple and different uses by different parties that produce multiple demands for the same item. *See id.* at 294-95.

85. *Id.* at 295.

86. *Id.*

87. *Id.* at 293-95.

88. *Id.* at 295.

without additional costs) and non-excludable (meaning that persons cannot be excluded from using the items).[89] To Demsetz, a public good that satisfies the additional condition of non-excludability is better treated as a distinct type of good, which he terms a "collective good."[90]

This article adopts Demsetz's nomenclature and refers to designs of patent-eligible advances as excludable public goods.[91] Patent-eligible inventions are public goods because they are non-rivalrous. There are no marginal costs associated with additional use of these advances, meaning that their use by multiple parties is non-rivalrous. However, these same advances are excludable (at least to a degree) because limits on access to the advances can be imposed, usually by limiting access to items or processes that implement the advances.

The potential for effective access controls over implementations of patent-eligible advances creates opportunities to charge for access to the advances. This in turn provides the basis for payment schemes that reward and incentivize producers of the advances. The key point for our purposes is that patented designs are non-rivalrous in use, but are subject to access controls imposed through the patent system (with some imperfections in access controls that are discussed at a later point in this article). Persons who do not reach an arrangement with a patent holder and pay that party for access (or pay someone else who has already paid for and gained access to the patented invention) can be excluded from access to the invention through injunctions and other relief. These limitations on access are not costless, but are recognized under patent laws to establish an access payment and reward scheme that is designed to encourage production of patentable inventions and to regulate the allocation of resources to such production.

b. Parallels in Joint Supply Goods and Public Goods Production

The parallels between the dynamics of the production of joint supply goods and public goods lead to parallels in the equilibriums between supply and demand that tend to define optimal production of both types of goods. These parallels in production equilibriums are

89. *See* Bell & Parchomovsky, *supra* note 9, at 578 n.250.

90. Demsetz, *supra* note 74, at 295.

91. If one wishes to adopt a nomenclature other than the term "public goods" for nonrivalrous, but excludable items, the analysis in this article is the same.

apparent if we reconsider the supply and demand functions in Figure 1 as descriptions of public goods production relationships, rather than as plots of relationships for joint supply goods production. Assume now that we are interested in the production of an excludable public good, say television shows to be made available through cable television systems.[92] This type of intellectual product is an excludable public good. Television program are public goods in that one party's viewing of a program does not interfere with or diminish the experience of viewing by an additional party. However, in systems like cable television, access to these programs is excludable and a system of charging for access can be implemented.

The advantage of analyzing the production of this sort of public good in relation to the same factors as the production of joint supply goods can be seen by interpreting the relationships shown in Figure 1 in terms of public goods production.[93] For simplicity, consider a two consumer world in which the demand for television programs depends on just two parties seeking these programs, Mr. H and Ms. M. The transmission over a cable channel of a program can satisfy the demand of both H and M, just as the slaughtering of a single steer can satisfy the demand for both meat and hides. By taking the different demand functions of H and M into account and combining them to form a single demand function, the analysis of public goods production dynamics tracks the production of joint supply goods as previously discussed in this article.

Reinterpreted in terms of public goods production, Figure 1 presents a summary of the supply and demand relationships influencing public goods production. The curve d represents the demand of H for more of the public good, in this case more access to television programs. The curve D represents the aggregate demand of H and M (that is, D represents the sum of the individual demands of H and M). The curve S represents the willingness of a maker of programs to supply more programs at various per program access prices. Overall, this represents a system in which program viewers (the consumers in this system) are willing to pay for access at any point on line D, where as providers of access are willing to supply

92. This example is one proposed by Demsetz, who focused on the demand for two parties for access to televised showings of taped programs. I have updated this example to specify the means of controlling access to the programs as being a cable television system that is capable of limiting access to particular shows and charging to provide access (on either a subscription or per-viewing basis). *See* Demsetz, *supra* note 74, at 294-95.

93. *See id.* at 294.

access at any point on line S.

The production of more programs will tend to reach an equilibrium at the point P-Q. At this price/volume point, the production of additional programs is efficient in the sense that net costs (taking into account opportunity costs of the program viewers) are minimized. Any less production would leave the viewers (H or M or both) with unsatisfied demand for which they would have paid more to satisfy than the supplier would have charged to satisfy it. Provided that the supplier was aware of this demand, sales of access to more programs would increase in number to meet the demand until the volume represented by P-Q was reached. Thus, underproduction (in relation to viewers' desires and willingness to pay) is corrected by increases in production and access up to the point where prices charged and willingness to pay are matched at point P-Q. In a similar fashion, efforts to produce too many goods and to charge too much for access will be discouraged. A supplier that tries this will be met with an unwillingness of viewers to pay and will need to move its offerings to a lower point along the S curve. The result will again tend to be a matching of supply and demand at point P-Q.

Demsetz described the general similarity of the factors governing the production of joint supply goods and public goods as follows:

> [J]ust as the number of goods which are supplied jointly (meat, hides, bones, etc.) is not limited to two except for expositional convenience, so the number of persons demanding a public good is not limited to two. With joint supply if the number of goods is increased or with a public good if the number of persons is increased, we will merely have complicated the geometry [of graphs such as Figure 1] without changing the analytical similarity of the two cases at all.
>
> What then is the difference between the two cases insofar as rate of output is concerned? There is no difference . . . , *provided that [parties] can be excluded from consuming the public good if they fail to pay for it*, which, of course, is implicitly assumed to be true in the joint supply problem.[94]

3. Advantages of Restricting Public Goods Under Access Controls

Controlling access to public goods and allowing private

94. *Id.* at 295.

originators of the goods to charge for access has two important advantages in mediating the private production of public goods.

First, it encourages the private production of additional public goods up to, but not beyond, the levels that meet the public's demand for more access. An access control system can thereby mediate the production of public goods at efficient output levels.

Second, an access pricing system can also regulate inputs to the private production of public goods in desirable ways. Such a system effectively puts a price on additional units of public goods, which can in turn influence the allocation of scarce resources towards the private production of additional public goods. This pricing mechanism tends to attract resources to the production of additional public goods and additional access opportunities where this is the best use of the resources (as measured by the public's value and willingness to pay for more access). An access control and payment system for public goods can therefore mediate the allocation of scarce resources to the private production of public goods in socially valuable ways.

This section analyzes both these aspects of access controls for public goods, explaining how access controls and access pricing practices governing the public's use of public goods can realize the types of societal benefits just described. The next section of this article describes how these same benefits can be realized through patent-based access controls and access pricing mechanisms for inventions.

a. Regulating and Targeting Private Production of Public Goods

A system controlling access to public goods and permitting the controlling party to charge a price for access achieves production incentivizing and regulating benefits by attaching a value to the production of additional quantities of public goods. By signaling to potential producers the willingness of public goods users to pay for more access to public goods, such a pricing system creates incentives for potential producers to create the types of public goods the public values and to expend production resources up to (but not beyond) the amounts that the public will pay for access to additional amounts of public goods. In this manner, the production of public goods is directed towards goods that the public values and is regulated to levels of production that match the limits of the public's willingness to pay for more units of public goods.

This type of system also signals the priorities of the public in its

demand for additional public goods. The willingness to pay for access to additional public goods indicates priorities in several dimensions, including the importance placed by the public on one version of a public good versus alternative versions of the same type of public good, on one type of public good versus another type of public good, and on public goods of various sorts relative to other types of goods and services. The willingness to pay for access to additional public goods signals the public's aggregate level of demand for additional access to the goods. These signals can be used by potential producers of public goods to target and regulate the production of additional public goods and to create additional access opportunities.

A system of pricing of public goods access in this manner is efficient in that it economizes on opportunity costs related to access to public goods. Where there is someone willing to pay for additional access to a public good at a particular price and there is some party willing to provide the additional access at a marginal cost less than that price, this type of pricing system will encourage the potential provider of the access to act and meet the demand of the other party seeking access.

Equilibrium will eventually be reached where the marginal costs of production of additional access to the public goods matches the marginal price that parties are willing to pay for that access. Such equilibrium will minimize the joint opportunity costs of the producers and users of public goods. At lower levels of production of public goods, the potential users will suffer opportunity costs. These will stem from their missed chances to have greater access to the public goods at prices that are less than the benefits they expect to realize from greater access. The missed opportunities to realize net gains of this sort are opportunity costs to the unsatisfied users of the public good. In a reciprocal fashion, the overproduction of public goods will create opportunity costs to the producers of the goods as some amounts of available access are not utilized and the amounts expended to make those available are not fully compensated. The opportunity costs here are related to the chances to use these wasted production resources elsewhere to greater advantage. A public goods access system of the type described here will tend to drive pricing and production of public goods access to levels where the opportunity costs of both producers and users of the public goods are minimized, thereby achieving a socially optimal production level.

b. Attracting Scarce Resources to the Production of Public Goods

Beyond regulating the private production of public goods to encourage the production of types and amounts of public goods that match public interests, a public goods access control system like the one described here can realize additional benefits by regulating the allocation of resources as inputs to the production of public goods. The pricing of opportunities to access public goods tends to ensure that scarce resources are allocated to the production of public goods where this is the most valued use of these resources, taking into account both their value in producing additional public goods and their value in other uses.

This type of "input management" advantage of access controls and access pricing for public goods has benefits in both encouraging and discouraging the use of resources for the production of public goods. Where there is a strong need for more public goods—as evidenced by a willingness of numerous members of the public to pay substantial amounts for additional access opportunities—potential producers of the goods will be encouraged to bring scarce resources to the production of public goods and to forgo the other products and benefits that different uses of the same resources might have produced. However, where the value placed on additional access to public goods is small, then scarce resources will not be attracted to the production of more access opportunities and scarce resources will tend to be allocated to other, more attractive ends.

The resource allocation decisions under consideration here are similar to those faced by a builder deciding whether to build a house on speculation (that is, not for himself, but rather to be sold to buyers upon completion). His projections of the ultimate sale price of the house will determine whether he initiates the house building project, as well as what type of house he will try to build and what materials he will use in constructing the house. Consider just one of the many resource allocations decisions he will make in assembling inputs to the building project. In evaluating the purchase of lumber for construction of the house, he will confront a market price for standard lumber (such as 2x4 boards and other standard components) that will be a key factor in determining whether and how he goes forward. If the costs of these and other construction components (plus the costs of the labor needed to complete the project) are less than the amount he projects he will be able to sell the house for, he will consider going

forward with the construction. Actually, a mere nominal profit (that is, a profit from a projected sale price that is just slightly greater than the projected costs of production) will probably not be enough to convince the builder to go forward. The projected net profits will need to be compelling in context—that is, the projected profits from building the house will need to be greater than the projected profits that the builder expects from the next best use of his time, efforts, and money. If the construction of the house (based on its projected price as influenced by future property controls over the house, if built) has a projected profit that trumps the other alternatives available to the builder, the builder will probably go forward with the construction and, by this action, divert and allocate a quantity of lumber to the purpose of building the house.

This allocation of lumber is a reflection of its projected value in the construction project. If this builder values it sufficiently, the lumber in question will go to the construction of the house. If the potential builder values the lumber less, then the lumber might go to the construction of other houses or other commercial properties. If lumber is generally valued at low levels at a given time, then the production of lumber might diminish and some potential lumber might be kept in reserve as trees until a later time when lumber was needed again in greater quantity and more trees were processed to produce more lumber. In these ways, the decisions of those who produce useful items and who value production inputs (like lumber in a construction project) allocate scarce resources to serve as inputs to production where this is the most valued use of the resources in question.

In so allocating scarce resources to the production of items, a product pricing system carries back upstream in the production process to affect decisions to produce scarce resources, the prices that are paid for them, and the degree to which these resources are devoted to various types of production in which they can play a role. By establishing access controls and access pricing for public goods, the production of public goods is brought into this resource allocation system. Projects for the private production of public goods compete with other projects (such as the private production of private goods or the construction of real estate properties) for the allocation of scarce resources that might be put to these other purposes.

Absent a scheme for pricing access to public goods and the establishment of an associated reward and incentive system for producers of those goods, the private production of public goods is

disadvantaged in the battle for scarce resources with other types of profitable production. Resources that are scarce and consequently costly to obtain will tend to be allocated by market systems to uses with clear payoffs. The production of public goods will proceed at less than publicly optimal rates and the allocation of scarce resources will be diverted away from public goods desired from the public. The result will be a misallocation of those scarce resources to ends less valued by the public, all as the consequence of the failure to translate the public's demand for additional public goods into access controls and access prices that would influence production decisions and attract more scarce resources to the production of additional public goods.

IV. PATENT RIGHTS: WHY WE NEED THEM

This section explores the benefits of applying Demsetz's analysis of public goods production to the private production of patent-eligible inventions. It argues that patent rights and rewards are best interpreted as means to implement invention access controls and pricing mechanisms that achieve an approximation of Demsetz's ideal system for incentivizing and regulating the production of private goods, operating in the case of patents in the particular realm of the production of patent-eligible inventions.

A. Patent Rights and Rewards as Means to Regulate Invention Production and Allocate Scarce Invention Inputs

The view of patent rights and rewards described in this article explains how patent rights establish prices for access to new inventions and thereby encourage the production of new advances and the allocation of scarce resources to such production. This view is developed here in a theory of patent rewards that is both descriptive (in helping us understand how patent rights affect invention production and decisions about how to allocate scarce resources to such production) and normative (in specifying why the invention incentives and resource shifts resulting from the enforcement of patent rights achieve socially beneficial levels of invention production and thereby serve the public interest).

The next portion of this section examines how Demsetz's framework for evaluating the private production of public goods provides valuable insights when applied to patent rights and the production of patent-eligible advances. Demsetz's views are used to

define a theory of patent rewards in which these rewards serve as efficient means to regulate the production of patentable advances. This theory is used descriptively to explain how strong controls over the making, use, and sale of patented advances (as enforced through patent rights) influence the types and amounts of new inventions targeted by potential inventors and further how impacts of patent rights on the valuation of new inventions has a derivative effect in attracting scarce resources to work on developing highly valuable new inventions.

In the concluding portion of this section, the analysis shifts to a normative mode in which the theory developed here is used to evaluate the public benefits from a strong, predictable patent system in which the nature of probable patent rights and rewards can be anticipated by potential innovators and meaningfully influence their decisions about how to conduct inventive projects. The analysis will argue that strongly enforced patent rights are beneficial to the public in ensuring that invention production is encouraged and targeted in relation to the nature and strength of public desires for various types of new inventions.

B. An Improved Patent Reward Theory for Regulating Invention Production

Demsetz's framework for analyzing the private production of public goods suggests an improved patent reward theory in which patents play central roles in pricing access to patented inventions. Patent rights serve as access controls regarding patented inventions.[95] These controls are imposed as means to implement a charging mechanism regarding invention access, with the resulting access fees serving as rewards for the production of patent-eligible inventions. This fee and reward system is imposed to encourage and regulate the private production of additional patent-eligible advances as public goods.

In order for patent-influenced access fees to reflect the full scope of user demands for new inventions, the fees for access paid by each party gaining access to a new invention design should equal (as closely as possible) the incremental utility seen by this party in gaining access. This incremental utility will reflect the new levels of utility achieved by the user in using the new invention over the utility

95. *See* 35 U.S.C. § 271(a) (2006).

experienced by the same party in using earlier (and usually non-patented) substitutes that perform the same tasks as the patented item. In a system of costless information and transactions, full price discrimination between users would be possible and individuals would be required to pay for access to new inventions at individually determined prices. Each potential user of a new product embodying a new design would be willing to pay an amount for access to the product up to the amount of their new value gained from access to the product. Potential inventors of new products would consider expending resources up to the aggregate demand of these users in developing new inventions.[96]

Seen this way, the full invention-pricing functions of patents and patent enforcement are apparent. Patent rights are not simply means to gain inventors some rewards for producing inventions. Patent rights are rather means to establish the value of inventions and to put the production of inventions into the larger picture of potentially valuable asset production. With a value attached to the production of additional inventions, efforts to produce new patent-eligible inventions can compete effectively for the attention of talented individuals and resource rich organizations amidst other potentially profitable and rewarding actions that might be chosen. At the same time, invention valuations enhanced by patent rights and invention access controls serve to attract scarce resources to the production of patent-eligible advances in efficient ways. Thus, interpreting patents as the linchpins of an invention access pricing system—the improved patent reward theory described in this article—provides new insights into the functioning of the patent system and the types of patent law reforms that may increase the public benefits from the patent system.

96. Ideally, again in a world of costless information gathering and pricing administration, price differentiation by product users would be possible and each would be charged for access to a new design in light of his or her personal willingness to pay for new utility. This type of price differentiation, were it possible, would maximize the total demand for new inventions, which would in turn increase the amounts that potential producers of the inventions would feel it wise to expend in producing the new inventions. A perfectly administrable system of access controls would make this type of particularly efficient price differentiation possible; as we shall see in later discussions, what the real patent system uses as a substitute is more in the nature of categorical price differentiation, where users of different embodiments of patent inventions are charged different patent-influenced prices that approximate their different demand and perceived invention utility characteristics, but do not achieve individualized price differentiation that would be even better if the costs of administering it were not prohibitive.

C. Descriptive Implications of the Improved Theory: How Patent Rights Regulate Invention Production

The potential impacts of patents on both invention inputs and outputs can be seen by applying Demsetz's framework for analyzing the impacts of property laws on the production of public goods to the particular case of the production of patent-eligible advances. This involves recasting his analysis into a patent context. The remainder of this subsection examines Demsetz's conclusions as they apply to the particular case of restricting access to patented inventions and thereby influencing and beneficially regulating the private productions of patentable inventions.

The aims of a patent-based invention reward or "pricing" system in this context are to summarize consumer demand and valuation assessments for inventions through patent-influenced prices for goods embodying inventions or through royalty payments to inventors for access to inventions. This type of patent-influenced reward system defines a set of demand and pricing characteristic for the production of additional inventions. The planning of potential inventors (and their backers, such as large corporations) will tend to respond to this demand with attempts to supply highly valued inventions. The result will be that choices about how many inventions to pursue, of what sort, and at what invention development cost will be resolved in research planning decisions in relation to the desires and demand characteristics of potential invention users.[97]

An example will make this type of operation of the patent system and its regulation of invention production clearer. For simplicity sake, consider a two consumer world in which the demand for mousetraps

97. Of course, the ability of parties pursuing inventions to plan their efforts in light of these factors depends on the availability to them of information on these aspects of invention demands and rewards. Limitations on the availability of this type of information may make an invention reward system useless in mediating the production of potentially useful inventions. Where the needed information is lacking, this may justify withholding the application of the patent system entirely and saving the associated administrative costs that administering the patent system would otherwise have produced. The possibility that information gaps and transaction costs may justify a more limited system of patent rewards and rights than the ideal reward system discussed at this point is considered at a later point in this article. The aim here is to profile the ideal system assuming few information gathering costs and low transactional costs. Later discussions will address how we may wish to back away from this ideal to take into account the practical deficiencies of our information gathering and transactional abilities and to tailor our real patent system in light of these real world limitations.

depends on just two parties seeking these traps, Ms. H and Mr. M.[98] Ms. H runs a restaurant and thus values the elimination of mice (and a better mousetrap that will do this effectively) very highly. Mr. M is a homeowner and is interested in a more effective mousetrap, but less strongly interested than Ms. H.

Returning to Figure 1, assume now that the demand curve d reflects the prices that Ms. H will pay for access to additional types of new mousetrap designs with increasing utility to her (with access charges implemented, for example, through patent-elevated prices imposed for buying traps with the new designs). The features of mousetrap designs that correspond to increased utility and value in the eyes of Ms. H are personal choices that do not matter in our analysis. She will, if she is a rational consumer, tend to be willing to pay more for access to designs that serve her better (in her own estimation of what is better), but the features of a design that make it better to her and how much she will pay for access to a new design are consumer-specific choices assumed here to be summarized in her demand curve.

Demand curve D in Figure 1 reflects in our current analysis the aggregate demand of Ms. H and Mr. M for access to new mousetrap designs (with the individual demand of Mr. M not shown, but mathematically ascertainable from the D and d curves shown and equal to D-d). Curve S reflects the willingness of a supplier who can control access to mousetraps with new designs to develop and supply new designs at various price levels, with volume reflecting the supply of more designs not more units of a single design. The socially optimal amount of production of the new mousetrap designs is encouraged by the combined demands of individual mousetrap users for new designs and the creation of market forces that encourage the development of new designs, but only up to the level of aggregate consumer demand. The resulting incentives should encourage the production of just enough new inventions as public goods—not too many (as to waste resources on inventions with less increased utility to invention users than the costs of developing the new inventions) and not too little (so as the leave invention users with unmet needs that could have been met through the cost-effective development of additional inventions at costs less than the utility gains the inventions

98. This example is derived from one proposed by Demsetz in which he focused on the demand for two parties for access to a televised showing of a taped program. I have translated his example from the world of copyrighted television programs into the world of patented inventions. *See* Demsetz, *supra* note 74, at 294-95.

would have produced).

Obviously, a two consumer marketplace for new designs is a trivial case. In the real world, these two parties might just link forces and seek to contract directly with the persons who could supply new mousetrap designs. Such a contract-based system could provide efficient incentives for invention by setting contract payments for successful inventions at levels that encouraged efficient efforts to develop and supply those inventions. The result would be a contract-implemented agency relationship in which the potential users of a new advance are the principals seeking efficient invention services from innovators serving as agents. Whether or not the work of the agent was efficiently encouraged or completed would depend on the accuracy with which the users could define their invention needs, the degree to which the agents could be incented to produce advances serving the users' needs without wasted efforts on other designs and tasks, and whether the amounts spent in contract-defined rewards exceeded the gains that the users realized from the new designs and mousetraps that resulted from the work of the innovators.[99] This sort of contract-based system could be implemented through direct contacts between innovators and invention users. It requires no patent system to link the interests of these parties.

The patent system serves as a substitute for the type of contract-based incentive system just described in situations where there are many potential users of a new invention and contractual mechanisms are too unwieldy to be used by a group of interested parties to band together and create innovation incentives on behalf of their group. Demsetz's analysis describes a different means to link the interests of innovators and invention users through market processes, a mechanism which will play a much more central role than contracting in situations where potential invention users are numerous and dispersed widely such that effective contracting processes are not possible. The implications for the patent system of Demsetz's insights

99. In addition to agency problems that might interfere with this type of innovation through private contract incentives, in the absence of patent rights limiting use of the designs, the parties might also be worried about free rider effects—that is, benefits to competitors of the contracting parties who did not contribute to the costs of development. These effects would be present if the advances that were the targets of the contracting scheme were freely available to others not paying a share of the development costs. This might occur because physical limits on access to the resulting designs (or trade secret protections limiting access) were not sufficient to prevent knowledge of the designs by competitors once products reflecting the designs were publicly sold.

about the desirability of access controls for public goods are apparent if Ms. H and Mr. M in the prior example are thought of as very different individuals, each with a different set of demand characteristics and a different demand curve. Demsetz recognized that a producer of a public good faces the whole market demand for that good,[100] meaning that the public good can be supplied to all parties that seek it (and are willing to pay the access price demanded for use of the public good). The aggregate demand for new inventions of a particular type (such as a new mousetrap design) will reflect the sum of the incremental values that different types of invention users place on inventions and carry forward in their willingness to pay for access to the new advances. Hence, in the prior example, Ms. H and Mr. M can each be thought of as a representative of a group of similar invention users that will pay, as a group (aggregating the invention access payments of the similarly-situated group members), a certain amount for access to more inventions with a particular quality (with the nature of the quality of interest somewhat different by user group, just as the point of interest of consumers of hides and meat is somewhat different in the case of steer production).

Seen this way, the combined demand function graphed in Figure 1 summarizes the aggregate demand of different types of potential users for products based on new inventions. This demand reflects the interest and value to users of the new products in the functionality provided by the new products. The details of the functionality of interest (and the value that each category of invention user places on the functionality) will vary for different invention users, but an overall demand for new inventions can be projected from these component demands. From this overall demand, potential innovators can determine (at least roughly) what types of inventive projects will likely be met with favorable user reactions and at what sorts of payment levels.

The role of patent law in this system is to enable the exclusion of parties from access to patented technologies without paying for that access. Patent infringement suits leading to injunctive relief help to ensure that users of patented advances (as well as makers and sellers) do not have access to the advances without paying for the access. These suits (and the threat of them) are means to direct consumers to gain access to the patented inventions through channels authorized by

100. Demsetz, *supra* note 74, at 296-99.

the patent owners and with payments to those owners (in either patent-elevated sales prices or license fees paid to patent owners). Suits leading to damage recoveries supply payments for past unpaid access to patented inventions, payments that substitute for privately negotiated access fees charged by the patent owner. The threat of these suits encourages the payment of access fees without litigation, thereby creating a feedback effect promoting private processes for pricing and paying for invention access. Overall, the patent system and its remedies tend to preclude unpaid access to patented advances and to provide a basis for the type of invention "pricing" system envisioned by Demsetz. As such, patent law plays a central role in regulating the production of excludable public goods such as patent-eligible inventions.

The enhanced patent reward theory described here has many potential uses in helping us to understand the impacts of patent law features in diverse areas of the law.[101] Evaluating the range of inventions for which patent-based production incentives and regulation are particularly valuable can help us to define the proper scope of patentable subject matters and the corresponding scope of patent system influence. The desirability of limiting patent-based access controls to situations where the present utility of an invention is manifest and potential users can evaluate and price access to the inventions in terms of the incremental utility of the advances suggests that presently identifiable utility should be a minimum feature of advances constituting patentable subject matter. Furthermore, in order to aid in the operation of a system of access controls and pricing of the sort described here, a clear description of the present utility of an advance should be a minimum requirement of enabling invention disclosures in patent application documents.

The types of circumstances where we are most concerned about misallocation of scarce resources and most in need of the type of resource allocation mechanism provided by the type of reward system described here may inform our laws on non-obviousness. The range of parties who should be expected to pay rewards to invention developers and thereby beneficially influence the decisions of potential inventors about the scope and targets of inventive projects may inform our laws about tests for patent infringement triggering infringement remedies. Finally, because they are clearly central to a

101. Many of these implications of patent reward theory for patent law standards are explored more completely in Part IV of this article.

system of invention access pricing and rewards, standards for the relief available for patent infringement may be directly shaped by consideration of the enhanced reward theory described here. Implementation of the type of invention production system described here will depend on patent infringement damage payments that are large enough to properly compensate patent holders for the increased utility provided to past users of patented inventions. Patent remedy standards must also provide for a range of injunctive relief that will ensure inventors have the needed confidence over invention access controls to project inventive project payoffs and to shape their projects in terms of payoffs from meeting the full range of user needs that they can satisfy through efficient private production of new patentable inventions as public goods.

These, and other implications for patent law standards of the reward theory described here, are discussed at a later point in this article.

D. Normative Implications of the New Theory: The Case for Strong Patent Rights

An invention access pricing and reward system implemented through strong patent rights and access controls potentially can produce two key public benefits. Such a patent-based system of invention controls and access pricing can encourage the production of inventions in socially desirable forms and numbers. Such a system can also allocate scarce resources to the production of inventions in an efficient manner. This subsection considers these beneficial patent system impacts, thereby summarizing the normative case for strong patent rights.

1. Features of a Strong Patent System

A strong patent system is one where patent rights apply to particular inventions in a predictable manner and, if applicable, prevent unauthorized (and uncompensated) access to the inventions, also in a predictable manner. A patent system that is predictable in these two senses will provide inventors with means to protect rewards or "prices" for successful inventions and to make informed decisions about innovation projects and resource allocations.

The aim of patent rights to exclude parties from unauthorized access to patented inventions is to activate bargaining and pricing processes that aid patent holders in charging for access to their

inventions. The resulting access fees, charged through patent-influenced sales prices or through licensing fees paid by producers or users of patented advances, establish prices and values for various types of inventions. Exclusive rights and the injunctions that back them up are the starting points for privately-implemented invention pricing schemes. Patent holders are able to carry out these invention pricing schemes by selectively enforcing patent rights (where an access fee is not paid) and contracting for invention use (where an acceptable access price is agreed to be paid for use).

In this framework, damage remedies for patent infringement are backstops to clean up after failed efforts to charge for access, with dollar payments to a patent holder serving as substitutes for the access fees the holder would probably have demanded had an access transaction been constructed through private means. Additional punitive damage awards for willful infringement tend to encourage attention to possible patent rights conflicts before parties act, thereby promoting contracting to gain access to inventions and the payment of access fees.[102] In this sense, damage remedies regarding past infringement not only provide a substitute for privately negotiated invention access arrangements and pricing, these damage recoveries encourage care by potential infringers about the possibility of further infringement, thereby tending to bring these parties into the privately-conducted access pricing system.

These aspects of a strong patent system encouraging parties to establish private processes for implementing invention access pricing tend to ensure that means for charging for invention access are implemented in efficient ways. Patent holders have a direct stake in efficient charging mechanisms because inefficiencies will reduce what they take away from charging for invention access. These same incentives will encourage efficient frameworks for product differentiation and price discrimination in charging for invention access among different types of invention users. If patent holders have means for charging different prices for access to patented inventions by different subsets of invention users, the patent holders can charge different access fees for different users that reflect their different invention valuations and their different willingness to pay for access to the inventions. The costs of administering such a system of price discrimination will fall on the patent holders (in the sense that

102. Punitive damages of up to three times a patent holder's compensatory damages may be awarded in cases involving willful patent infringement. *See* 35 U.S.C. § 284 (2006).

the costs will come out of the potential profits to patent holders from enforcing their patent rights). This means that choices about whether to charge different prices for access to inventions (perhaps by supplying different types of products incorporating the invention), as well as decisions about how to implement different prices and how to group similar users for purposes of charging different access prices, will tend to be made efficiently to minimize administrative costs that would reduce the gains to patent holders.

2. Public Gains from a Strong Patent System

Two important types of public policy benefits can flow from a strong patent system like the one just described. An invention access control system like this can 1) encourage publicly optimal levels of invention production and 2) attract scarce resources to this type of production at publicly optimal levels. However, even if a patent system operating in an ideal form would produce these public advantages, transaction costs or other detrimental aspects of administering the system may swamp the advantages of a strong patent system and produce no net public benefits from having such a system. These transaction and administrative costs place practical limits on the public benefits from the patent system. Both the potential benefits of the patent system and the limiting impacts of patent system costs are discussed briefly in this subsection.

a. Encouraging Optimal Levels of Patent-Eligible Invention Production

A system of strong patent rights and invention access pricing mediates invention production in desirable ways. It encourages production of patent-eligible inventions at socially optimal levels. It can help to ensure that there is no underproduction of such patent-eligible inventions relative to the public's willingness to pay for more. By controlling invention access through patent rights, patent law assists inventors in implementing an invention reward and pricing system that targets and funds invention production levels so as to match the supply of additional patent-eligible inventions with public interest and value associated with such additional inventions.

In assessing the role of patents in regulating invention production, it is important to distinguish between the production of patent-eligible inventions and the production of other, more commonplace technical innovations. Patent-eligible advances are outliers among inventions, capable of production only by parties who

have and apply design insights that are beyond the obvious reasoning of the bulk of scientists and engineers in the same field.[103] These are simply a different inventive animal than the more straightforward innovations of the sort that are produced on a day-to-day basis by the great bulk of scientists and engineers working with standard design principles and obvious insights to extend earlier designs.

The patent system is premised on the view that we have plenty of day-to-day innovations (through the workings of normal competitive processes), but that we lack optimal levels of the outlier innovations comprising patent-eligible advances and that we need special incentives to encourage the creation of more of these outlier inventions.[104] Patent rewards are designed to give a special boost to the latter, on the ground that these are often important advances that shift technology progress in new directions and avoid stagnation in areas of well known technology and products.[105] Not every patented advance is practically significant in that it is a major commercial success or put into widespread use; yet, each is intellectually significant in that it diverges significantly from prior engineering directions and has the potential to lead designers in its field in important new directions.

The existence of the patent system is evidence that Congress views these sorts of patent-eligible inventions as special and the producers of these inventions as somewhat privileged. The privileged status is indicated by the monopoly control over invention use and implementation given by Congress to patent holders.[106] Such monopoly control is a rare exception to the principles and laws

103. *See* 35 U.S.C. § 103(a) (2006). *See also* Graham v. John Deere Co., 383 U.S. 1, 17 (1966).

104. The distinction between day-to-day innovations (which can be made by a wide variety of parties having average or greater skills in a particular engineering or scientific discipline) and patentable advances (which are not obvious to persons with average skills and can only be made by persons with exceptional insights) is maintained by non-obviousness tests. These tests place day to day advances outside the patent system (and make them unrestricted by patent rights) and advances beyond the average practitioner inside the system (and potentially subject to patent-influenced rewards if other patent law requirements are met). *See* 35 U.S.C. § 103(a) (2006).

105. Patents tend to increase technological diversity by increasing the incentives for parties to pursue, perfect, and disclose technological solutions to practical problems that are not already obvious to the bulk of engineers and scientists dealing with similar practical problems, thereby increasing the types of technological approaches that are known to engineers and scientists. *See* 35 U.S.C. § 103 (2006).

106. *See* 35 U.S.C. § 271(a) (2006) (giving the relevant patent holder control over parties who can make, use, or sell a patented invention during the term of a patent).

(primarily antitrust laws) encouraging active competition in most other spheres.[107] This special status of patent-eligible advances indicated by Congress should be remembered when we consider the merit of various versions of the patent system. The goal of the patent system is not to encourage inventions in general.[108] It is to encourage the special type of outlier inventions constituting patent-eligible advances.[109] Whether or not the patent system is a success will depend primarily on its effects on the production of these special outlier inventions, not on the impacts of the patent system on all types of inventions.

It may be argued that excessive patent rights regarding patent-eligible advances will have a negative side effect on the production of additional non-patentable inventions. This might be the case where non-patentable inventions are made using patented advances as bases for improvements or as components of broader new products. If use of patented designs is either unavailable or curtailed due to prices charged for such use, the production of new advances derived from older patented advances might be limited. In this sense, it is argued, patent rights should be interpreted and restricted to encourage the best possible overall levels of innovations in total, with strong enough rights to encourage some exceptional, patent-eligible advances, but not granted to such a strong degree that downstream innovation is curtailed.[110]

This is a false framework for analysis. Whether or not strong patent rights will encourage heightened levels of production of

---

107. *See* Thomas B. Leary, *The Patent-Antitrust Interface*, FED. TRADE COMM'N (May 3, 2001), http://www.ftc.gov/speeches/leary/ipspeech.shtm (noting the view of some in the antitrust community that patent law was an "alien intrusion, which undercut much more exalted competition values and which should be given the narrowest possible application").

108. Innovations that involve design methods or approaches that are obvious to persons of average skill in the relevant technology field are not patentable and, therefore, not encouraged by the patent system. *See* 35 U.S.C. § 103.

109. Patentable inventions—that is, inventions which are potentially encouraged by the promise of patent-influenced rewards—are outliers in innovation processes in that these advances are not ones that persons of average skill in the related fields could produce through obvious extensions of prior designs, but rather non-obvious advances over prior knowledge about designs. *See id.*

110. *See* FED. TRADE COMM'N, TO PROMOTE INNOVATION: THE PROPER BALANCE OF COMPETITION AND PATENT LAW AND POLICY 22 (2003), *available at* http://www.ftc.gov/os/2003/10/innovationrpt.pdf (explaining that "broad initial patents may raise significant problems for follow-on innovation. From this perspective, vesting the initial innovator with broad patent rights reduces the incentives of follow-on innovators and potentially impedes their access to upstream innovation inputs." (citation omitted)).

inventions as a whole depends on whether they are granted exclusively (or at least mostly) for inventions that would otherwise not have been made. The strength of the patent rights granted does not matter, so long as these rights are limited to advances that would not have been produced absent the promise of patent rewards. What does matter is the specificity of the rights, not the magnitude of the remedies. If patent rights and rewards are granted only for advances that would not have been made but for the lure of these rewards, this reward mechanism will tend to produce more patent-eligible advances than would otherwise be the case.[111] This will, in turn, provide more inventions that are potential platforms or components for later production of additional patented and non-patented advances. Even if patented advances are only available for use in later inventions after payment of patent-influenced access charges (or only available after the expiration of patent rights), the impact of the patent system in producing more patented advances will be to achieve a net social gain. The availability of some patent-restricted inventions is better for society than not having these innovations at all.

By providing more technology designs as starting points for further innovation, patent rights adding to our store of patent-eligible inventions will tend to enhance later invention opportunities and, over time, will probably increase total numbers of subsequent inventions of all types (both patent-eligible advances and other more obvious advances). Conversely, if patent rights and associated access charges are added to inventions that would have been produced by normal competitive processes, these charges become just deadweight losses that impose costs on the later production of inventions which depend on the patent-eligible advances. The production of some advances derived from the patented advances will be unnecessarily curtailed or foregone if these costs must be paid.

b. Encouraging Optimal Diversion of Scarce Invention Production Inputs

A system of invention access controls and access pricing can serve a second valuable public function in allocating scarce resources

111. *See* David S. Olson, *Taking the Utilitarian Basis for Patent Law Seriously: The Case for Restricting Patentable Subject Matter*, 82 TEMP. L. REV. 181, 191-204 (2009) (arguing that patents and related use restrictions due to patent rights should be limited to inventions that would otherwise not have been produced without the added incentives of patent-related rewards).

as invention inputs. Invention rewards defined by such a system tend to attract and allocate scarce resources to the production of patent-eligible advances in circumstances where such allocation is the highest, best use of these resources as measured by the interests and values of the public.[112] The invention pricing process becomes an intermediary system whereby the use of scarce resources is prioritized and directed to projects involving the largest public needs and greatest public impacts. By tying public values of new inventions to the private values of potential inventors and the organizations that back them, the private interests and agendas of the inventors and organizations are made at least roughly congruent with the aggregate public interest in new advances. As the inventors go forward with resource planning concerning potential invention projects, their resource acquisition and allocation decisions are made with the public value of potential inventions in mind. Hence, these inventors and their organizations become agents of the public in acquiring resources for the private production of public goods in new inventions and in prioritizing and economizing on the use of scarce resources for this purpose.

In this regard, the patent system supports market mechanisms for valuing potential inputs to the production of inventions, enabling markets to take this use of resources into account in additional to other potential uses for the same inventions. Where inventions are expected to be of high value by those who are capable of developing the inventions, resources (including the time of highly capable inventors and other resources needed to support the innovators' efforts) will be drawn to innovation projects. Where inventions appear less valuable, the same resources will tend to be devoted to other purposes where the resources appear to have greater value. By mediating the allocation of scarce resources potentially serving innovation but also potentially serving other valuable ends, patent rights and the invention valuation and reward processes these rights support serve valuable public functions in promoting the efficient use of scare resources.

112. For a more complete analysis of the mechanisms whereby potential patent rewards attract scarce resources to invention projects, see Richard S. Gruner, *Dispelling the Myth of Patents as Non-Rivalrous Property: Patents as Tools for Allocating Scarce Labor and Resources*, 13 COLUM. SCI. & TECH. L. REV. 1 (2011).

3. The Offsetting Effects of Administrative Costs

Public gains from the patent system of the sort described to this point may be offset and even eclipsed by the negative effects of various costs of administering the patent system and the enforcement of patent rights. Several types of costs may have this effect, including costs associated with reviewing patent applications, enforcing patent rights, and determining whether projected activities will conflict with patent rights. Additional costs include the lost value of activities not undertaken because they were mistakenly interpreted to be infringing, but which were actually legitimate in the face of patent rights. When costs such as these outweigh societal gains from the improved regulation of invention production under the influence of patent rights, the patent system produces a net loss to society rather than a net gain. Under these circumstances, it will be desirable to withhold the impacts of the patent system and exclude a corresponding realm of subject matters or activities from patent coverage.

A complete exclusion of a particular technology or type of infringement from patent rights and rewards on grounds of administrative costs will rarely be justified, however. This will be warranted only where the costs of patent enforcement will always (or almost always, to an extent that it is not worth assessing the paramount size of system costs on a case by case basis) trump the advantages of patent influence on invention production in the area of technology or useful advance under consideration.

In most technology settings, it is probably desirable to recognize the potential for projected patent rewards to draw out socially valuable innovation for the reasons described to this point and to leave it to particular inventors to determine whether the projected patent rewards for their anticipated work and projected inventions still provide net incentives for action after the costs of patent enforcement and net rewards these imply are taken into account. In short, an inclusive view on patent system scope and patentable subject matter will be mitigated by the case-specific factual assessments of particular potential inventors; if the costs of gaining patent rewards, when added to the costs of producing a successful invention, seem unlikely to produce net returns that trump the returns on similar investments of time and resources in other avenues of action, then the costs of administration in that setting (at least those felt by potential inventors who are contemplating the advantages of being patent holders) will put a practical limitation on the scope of the patent system in

particular cases.

V. PATENT RIGHTS: WHEN WE NEED THEM

Looking at the patent system as a means to promote the production of patent-eligible inventions as public goods, when do we need such a system? As a system that has many potential costs—in system administration expense, reductions in productive activities by persons who must pay elevated fees for access to patented advances, and in reduced competition to produce and sell patented advances—the patent system must have considerable advantages before it will overcome these costs and create net public benefits. In short, the preliminary public case is slanted against the patent system because of its manifest costs. When will the advantages of the patent system be great enough to overcome these costs and to establish a positive public policy case supporting a strong patent system?

An answer to this question is possible through extending the analysis of the patent system just described. The patent system is an artificial tool created by Congress to solve a problem: the potential underproduction of technical advances that are outliers in their technical fields. The patent system should be implemented where this problem is most acute. That is, the strongest public policy case for the patent system favors applying the system where the underproduction of technical advances has the most severe public consequences. If the patent system is a tool for matching the production of patent-eligible advances to public needs for such advances, then the patent system is most important where a mismatch between inventions and public needs is likely to have the greatest consequences.

In considering the type of mismatch between invention production and public needs that is the province of the patent system, it is important to remember that patents do not promote all inventions, but rather just those outlier inventions constituting patent-eligible advances. Hence, the case for a patent system (if one exists) must turn on identifying when there will likely be an under-production of these special patent-eligible advances absent patent rewards and where such under-production will have serious societal consequences.

An emphasis on the need for the patent system to overcome an invention production mismatch provides a guiding principle for assessing when we need the patent system. This principle leads us to three subsidiary questions: 1) When is the underproduction (in terms of unmet societal desires for inventions providing functional

improvements) of patent-eligible advances socially important? 2) When are normal competitive processes likely to encourage production of inventions at socially optimal levels such that the special boost of the patent system is not needed? 3) When are the costs of administering the patent system so great as to overwhelm the invention production advantages of the system, thereby justifying limits on the scope of the patent system?

To fully justify taking on the manifest costs of the patent system, we should probably reserve such a system for circumstances where the answers to all three of these questions support the existence of patent rights. This indicates that we should limit the patent system to settings where the underproduction of inventions has serious social consequences, where there are reasons to believe that scarce resources will be under-allocated to the production of socially important inventions, and where there are not predictably high transaction costs of administering patent rights that will exceed and cancel out the societal benefits of producing more inventions.

The remainder of this subsection addresses these questions in turn by way of providing an overall answer to the question: When do we need a strong patent system?

A. When Invention Underproduction Matters

If the patent system is primarily a means to encourage and regulate the private production of patent-eligible advances as public goods, then we most need the patent system where we are concerned about the underproduction of patent-eligible inventions. Put in functional terms, given that the patent system is a means toward the end of enhancing invention production, the patent system is most strongly justified where that end matters most strongly to us. Where invention underproduction is particularly important to society, using the patent system as a means to offset such underproduction may be correspondingly important.

The underproduction of patent-eligible inventions may be a problem for several different reasons (with more than one of these problems potentially present in a given technical area). At least three types of reasons for underproduction of patent-eligible inventions may extend over a wide range of technologies: 1) the failure of contracting processes to produce proper levels of invention because of the large number of potential users of a product or service and the burdens of reaching these parties to charge for invention access through contracts; 2) the likelihood of strategic behaviors limiting

innovation because potential inventors are unwilling to benefit competitors who may gain as free riders on realized inventions; and 3) the hesitancy of parties with specialized research capabilities to commit these to innovation with no practical means of realizing returns from product producers or users without encountering free rider problems. These potential reasons for invention underproduction (and the circumstances in which they may apply) are described in this subsection.

1. Where Contractual Incentives Have Limited Effectiveness

Where there are few potential users for particular types of innovations, contracting processes (in which rewards for successful inventions are set up as "bounties" for providing the inventions) can be structured so as to encourage parties to pursue inventions at optimal levels. A party (or manageable group of parties working together) can seek a particular type of invention by simply approaching a party capable of producing the needed advance and agreeing to contract terms providing for a specified payments to the potential producer upon delivery of an invention design meeting the user's functional specifications.[113] This sort of contract—essentially a specialized design services contract—would have the potential to create optimal incentives for the pursuit of cost-justified invention development efforts by the party agreeing to produce a design and for the devotion of resources to these efforts in an efficient manner. Where this sort of contracting process is likely to be effective, there seems to be little reason to also provide for patent rights. Indeed, given that the administration of the patent system would add costs to this situation, it would seem desirable to withhold patent rights in circumstances where contractually-implemented incentives for innovation seem capable of producing innovation efforts at socially

113. A similar set of contract based incentives might be framed through a bidding process in which the functional needs of one or more users were stated and multiple offerors were invited to submit bids to provide a completed device or device design. The resulting contract, now involving a self-identified design creator who identified a desirable design challenge in the bid specifications and who responded to the bidding, would have the same incentive characteristics as the contract described in the text. The main difference in the contracting process derived from bidding is that the potential design supplier will not have to be identified by the party seeking a new design, but will instead come forward to present their capabilities regarding the development of the desired new design and to convince the acquiring parties of the sufficiency of those capabilities to produce a successful result.

desirable levels.[114]

However, contractual mechanisms will typically break down (and the patent system will have a greater potential role as a substitute) where contracting processes are ineffective or particularly costly to implement. In these settings, the patent system and the invention incentives that it creates can serve as a contract substitute to cause potential innovators to see the merit of various inventions (and focus on efficient resource allocations to create the indicated inventions) with the perspectives and value judgments of potential invention users in mind. The potential rewards of the patent system will provide potential inventors with this sense of perspective of invention users. These incentives can make potential inventors behavioral agents of invention users for purposes of pursuing inventive projects. As such, patent rights can serve as a substitute for equivalent agency contracts in settings where the latter are not possible or not cost-effective to implement.

Why might innovation-inducing contracts be difficult or impossible to implement and the patent system be needed to provide substitutes? There may be many reasons, but at least two seem likely to be widespread and to provide broad justifications for reliance on the patent system across diverse technologies and fields.

First, the number of potential users of an invention may be so large as to make it unwieldy to implement a contractual system for charging for access to inventions and creating corresponding rewards for invention creation. In these circumstances, only a subset of the potential user group could be banded together to form an administratively workable contracting group and could only be expected to pay a reward for successful invention designs that corresponded to their advantage from gaining access to a particular invention. This would create some incentives for invention, but not the same scope of incentives as if all the relevant and interested potential users of a successful invention could be expected to provide portions of the rewards for successfully producing an invention. Since contracting processes cannot be expected to include and reflect the user preferences and value placed on a particular invention by large groups of invention users, these processes will create suboptimal

114. I have previously argued that this is a desirable limit on the patent system and should be included as a component of standards for patentable subject matter. *See* Richard S. Gruner, *Intangible Inventions: Patentable Subject Matter for an Information Age*, 35 LOY. L.A. L. REV. 355, 448-50 (2002).

levels of incentives (and suboptimal levels of invention) where the size of potential invention user groups is large.

Second, where the identity of potential producers of a useful invention are unclear, it may be difficult or impossible for interested potential users of a particular type of advance to contact the relevant, capable innovators and to establish the needed invention incentives for the range of organizations or individuals who are capable of responding with a successful invention. Here the source of a possible gap in innovation incentives is not the number of potential invention users, but rather the lack of knowledge of the users about the capabilities and identities of potential innovators. Absent some targeting information about who might be a successful innovator, parties seeking to create contractual incentives for innovation will be impaired in efforts to contact one or more viable innovators and to implement invention-focused incentive contracts.[115] For the sorts of non-obvious advances that patents can reward and potentially influence, only a small number of exceptional innovators in a field may have the domain-specific knowledge or analytic abilities needed to achieve unusual insights leading to non-obvious advances. The ability of contracting parties to find and conclude incentive contracts with these scarce innovators may be limited at best. Due to lack of knowledge about who is a viable innovator in the realm of non-obvious advances, contracting processes aimed at promoting these sorts of advances at optimal levels may be difficult or impossible to target effectively.

For these reasons (and perhaps others), contracting processes aimed at creating optimal incentives for the production of non-obvious advances may be difficult or impossible to implement in many fields where there are numerous potential invention users or uncertain sources of innovation. Given that many fields (particularly those of widespread social importance) probably have one or both of these characteristics, the need for patent incentives as a substitute for contract-based incentives may be great across many innovation settings.

115. Even an open ended reward contract (in which the relevant parties agreed to pay the first party to submit a design with certain characteristics a particular reward) would have to be publicized to the potential innovators and administered to determine who had made a successful submission.

2. Where Free-Rider Effects Among Competitors are Substantial

Invention underproduction may also be a problem because of strategic behaviors between competitors who might otherwise engage in innovation to stay ahead of each other in their competitive activities. In serving a potential set of customers, a particular type of business will tend to keep the future interests of the customer set in mind, including the interests of the customers in outlier innovations of the sort that the patent system is aimed at influencing. This type of motivation would tend to cause a business to act as an agent of its customer for purposes of innovation. It would, as part of this mission, tend to include some mix of work on outlier innovations (produced by itself or by contractors working for the business) to the extent that the projected success of work on such outlier innovations initially appeared cost effective in light of the probable commercial value of the projected results and the projected likelihood of project success. In short, businesses might serve as surrogates for consumers in producing inventions at optimal levels, in which case the special incentives of patents will not be necessary. Innovation risk taking on behalf of the public (as well as on behalf of the businesses themselves) may be undertaken as part of business risk taking more generally.

Where there are two or more businesses competing for the same sales to customers, however, strategic behaviors among the businesses may cause each to hold back from efficient innovation, leading to a lack of optimal invention production. The problem that tends to limit optimal invention production in these circumstances stems from fear on the part of each potential innovator of free rider effects. Potential invention sources (both organizational and individual) who lack means to control inventions after they are produced will be loath to invest large sums and scarce personal resources in the discovery and development of new innovations, only to see competitors reap the main business benefits from the new innovations. This would be the case if a competitor could copy a new design, get an equivalent product into the marketplace, and gain the business benefits of that design nearly as well as the originator of the design, without paying any of the design costs. Fear that this will occur will cause each potential innovator to hold back from innovation, resulting in a less than optimal total production of new advances.

This type of free rider problem will be most serious where the

ease of free riding is greatest and the likelihood that potential innovators will appreciate this and hold back from innovation accordingly is also greatest. This will be the case where the costs of innovation are high and the costs of copying and gaining from an advance as a free rider are low. This combination of circumstances will produce the greatest gains to free riders and the greatest injuries to the innovators who are victimized by the free riders. This will also create the greatest apparent threat from free riders and, correspondingly, the greatest deterrents to innovation that patents can offset to public benefit.

3. Where Free-Rider Effects Limiting Specialized Research Activities are Substantial

Free rider impacts may also significantly limit innovation where this type of work would otherwise best be undertaken by research specialists. Given the complexity of many types of modern scientific and engineering research, the research function is now often quite different from the function of producing and marketing goods and services to consumers.[116] Effective and efficient research is conducted by parties with skills and resources that are divorced from those of production and marketing.[117] In some large corporations, these specialized research functions are simply conducted in a different department or unit of the overall corporate enterprise.[118] However, in an increasing number of contexts, the research functions leading to socially significant products and services are conducted in organizations that are primarily dedicated to research as their primary tasks. The two primary examples of these specialized research organizations are startup companies developing new technologies and university research labs producing new technological advances in conjunction with new scientific and engineering knowledge.

These specialized research enterprises are conducted in sole or in

116. Research activities, like product production and marketing, have been targets of efforts to gain and apply specialized knowledge and techniques. As specialization has increased, many firms have either organized research departments to focus on specialized research tasks or have sought research results from outside concerns that have specialized in the research steps needed to produce new products. *See* Barnett, *supra* note 51, at 55-64 (describing disaggregation of research, design, and production functions in semiconductor industry to realize specialization benefits).

117. *See id.*

118. *See* Maximilian von Zedtwitz & Oliver Gassmann, *Market Versus Technology Drive in R&D Internationalization: Four Different Patterns of Managing Research and Development*, 31 RES. POL'Y 569, 577 (2002).

part for commercial gains, although for some (primarily university organizations) this may be a secondary mission.[119] In order to initiate large projects, the managers of these enterprises must be able to present a convincing case to parties controlling the necessary resources (which are often staggeringly large) that there will be means to produce commercial gains from successful completion of the research that is being considered.[120] A plausible business objective is necessary, even though there is never a guarantee that a successful scientific or engineering result will be produced or that such a result, if achieved, will translate into a commercial success. Research results must have a foreseeable commercial value (even if this is contingent) if projects are to be funded (or at least aided) in light of their commercial potential.

The difficulty of free rider effects in this context is that, if successful research results are not capable of being controlled by research organizations, they will have little if any means to gain a return on their research efforts and investments. Once successful results are made public, any party will be able to use them in the absence of patents and the research organization will have little if any means to gain a return on its valuable discovery. Concern about free riders will destroy any vision of future profit potential in commercially significant lines of research. The research organization will bear the costs; the free rider will reap the gains. There will be no means to link commercialization to research costs in a way that will produce research payoffs and cost recoveries. Absent such a manifest link, research enterprises will be loath to start out on costly projects and the production of innovations at optimal levels will be curtailed.

As with the counterpart problem between competitors, this

119. In university settings, the primary research objective may be scientific knowledge enhancement, with the identification, dissemination, and commercialization of related practical advances as incidental secondary goals. *See, e.g.*, HARVARD UNIV. OFFICE OF TECH. DEV., INVENTOR'S HANDBOOK: A FACULTY GUIDE TO INTELLECTUAL PROPERTY AND TECHNOLOGY DEVELOPMENT 4 (2009) (citing as goals underlying technology transfer practices at Harvard University as including "advanc[ing] the development of the Harvard faculty's groundbreaking discoveries for the good of society," as well as helping to "enhance intellectual exchange, foster the widest possible recognition for [faculty] efforts and attract financial sponsorship of [faculty] research").

120. Parties seeking funding for support of technology research must typically compete with other parties seeking allocations of the same scarce resources for other projects (both other types of technology development and projects unrelated to technology development). In order to prevail in this competition for resources, the projected returns from technology development efforts must often be high. *See* Gruner, *supra* note 112, at 23-24.

version of free riding will be most serious–and the need for patents as counteracting forces will be greatest–where the costs of research and innovation production are largest and the costs of acting as a free rider are lowest. The nature of a new technology may suggest that it will be relatively easy to implement and replicate in products once uncertain design features of the new technology are worked out in research activities. This would suggest to research specialists who are potential innovators that their successful results will be easily reusable by free riders providing commercially valuable products to users and thereby precluding any effective means for charging for access to the research results.

The presence of multiple companies that are already large and highly active in the relevant product and sales market—and that are, therefore, well-positioned to serve as effective free riders if there are no intellectual property controls over the use of commercially valuable research results—will signal to potential research specialists that there is a high likelihood of a free rider use of research results absent intellectual property controls.

In these circumstances, research organizations will see that they have much to lose from the initiation of what will probably be uncompensated research and, correspondingly, potential free riders will be likely to capitalize on most or all opportunities to take up unprotected new advances and commercialize them in lieu of parties who have formal relationships (and compensation arrangements) with the research organizations that discovered the advances. In circumstances of high research costs and highly likely free riding which will cut off research compensation and return, research organizations are highly likely to be deterred from initiating costly research and the public will suffer from resulting suboptimal production of new innovations.

B. When Resource Under-Allocation Matters

Even where we are concerned about the underproduction of innovations (perhaps due to the probable weakness of contractual incentives or the disincentives of free-rider effects), competitive processes may be sufficient to alter resource allocations and bring competitive forces into play to increase invention production. Normal economic forces should, in the presence of sufficient resources for reallocation to new innovative tasks and opportunities, incent businesses to pursue innovative projects and adopt business practices that respond to unmet consumer demands for innovation and that

overcome problems like free-rider effects. Relying on these sorts of competitive processes to produce more innovation will probably be more effective than going to the more extreme and costly solution of invoking the patent system for increased incentives. The patent system will be needed (and its application through corresponding patentable subject matter standards most clearly justified) where there are reasons to believe that normal resource allocation processes will not bring adequate resources to support socially desired innovation and that resource under-allocation in this respect will be a problem.

Under what circumstances, then, are normal competitive processes in commercial settings unlikely to be adequate and the special spur of patent rights likely to be needed to bring sufficient scarce resources to the production of patent-eligible advances? To answer this question, we need to focus on how normal competitive processes allocate resources to the production of various items in socially desirable quantities, and the reasons why these normal processes may break down in connection with the production of patent-eligible advances. Both of these topics are considered in this subsection.

1. Allocating Resources to Invention Inputs Through Competitive Processes: The Disruptive Influence of Resource Scarcity

Normal competitive processes—including labor markets allocating the employment of talented individuals and securities markets allocating investment resources to various companies with research agendas—will tend to bring key innovation resources to bear on projects of interest to the public where the sources of these resources are numerous and the means to reallocate the resources are active and well-informed.[121] Where resources are scarcer, however, and the impacts of a few resource allocation decisions are critical

121. An active, well-informed market is generally an efficient and effective allocator of resources to the production of goods and services desired by the public. As noted by Robert Kuttner:

> At the very core of the market system is the price mechanism. Prices indicate what millions of individual goods and services are "worth" to willing sellers and willing buyers. Prices thereby function to apportion economic resources efficiently: they signal sellers what to produce; consumers how to buy; capitalists where to invest. . . . Markets, therefore, can claim to embody and express freedom of choice, as well as efficient allocation of scarce resources.

ROBERT KUTTNER, EVERYTHING FOR SALE: THE VIRTUES AND LIMITS OF MARKETS 11 (1996).

(because there are only a few parties who control the relevant resources and their decisions fully control the relevant public impacts of those resources), resources may not be allocated in accordance with the public's interest in useful advances and the consequences of resource misallocation in this respect may be severe.[122] Patent incentives more clearly tying the advantages of certain resource allocation decisions concerning scarce resources to the public's perception of innovation value will be advantageous in these settings to overcome otherwise present misallocations of scarce innovation inputs.

Normal competitive business processes are likely to encourage the production of inventions at socially optimal levels, provided that there are numerous persons with the analytic talents and research resources needed to produce a particular type of invention. Where this is the case, the needed inputs for invention will be available from multiple sources, meaning that they should be both easy to find and easy (that is, minimally costly) to attract to innovation projects. One or more inventors will tend to put together the resources needed for effective innovation and to produce new inventions capable of meeting public demand for such inventions.

Normal competitive processes should prove adequate in these circumstances to support socially optimal innovation production—that is, innovation production at levels that meet the demand of the public for the type of advance in question and minimize opportunity costs for desired but unavailable innovations. Errors in resource allocation by a few innovators will not be significant in under-producing socially desired advances because the large number of potential innovators (supported by the ease of assembling related invention inputs) will ensure that a substantial number of potential innovators will still make the needed resource allocation decisions and pursue the types of innovations that the public wants. The special added incentives of patent rewards implemented through controls on invention access will not be needed to boost invention production where there are many potential innovators and the resources to support their work.

122. Where there are few resource allocators—and few allocation decisions—the consequences of even one misallocation of resources to a project that is not the most highly valued by the public among the alternatives available may be particularly significant as there will not be other resource allocators who will support different allocations and better serve the public's needs.

Where the inputs to the production of a particular type of advance are scarce—either because few persons have the relevant analytic abilities and research skills or because there are few organizations or other parties with sufficient resources to support effective innovation—regular competitive processes may break down in bringing needed resources to bear on the types of innovation valued by the public. Even a few misallocations of scarce resources may preclude the types of innovation projects desired by the public because there are no additional quantities of the resources in question to allocate to innovation. A tighter coupling of invention value and rewards to resource allocations is needed in these settings. Each resource allocation decision should be made with a strong influence of the public's perception of invention value.

Patent rewards are a means to achieve this. By permitting invention originators to charge the public for access to their advances, these originators are encouraged to more closely equate invention value (as seen by the public) with the outputs of a successful innovation project. Decisions about the allocation of scarce resources to projects aimed at producing patentable inventions are thereby more closely influenced by the public value of the inventions at stake. Patent rewards for innovations above and beyond normal commercial rewards may be needed to attract these sorts of scarce inputs to the production of inventions and to ensure that the needed resources are not put to other less socially valuable uses.

2. Sources of Scarcity in Invention Inputs

Outlier inventions (of the sort patents cover) have features that often depend on scarce resources. These are inventions that are, by definition, not within the design capabilities of persons having average skill and knowledge within the relevant design fields.[123] This means that the outlier inventions that are the subjects of most patents are produced by only a few highly capable scientists and engineers (rather than by the great bulk of persons in the same fields who have average or less than average skills). The persons with the sorts of skills and knowledge needed to produce patented inventions are scarce resources. The special lure of patent rewards may be needed to attract these scarce innovators to the production of patentable advances, drawing them away from the other ends that such parties

123. *See* 35 U.S.C. § 103 (2006).

with scarce knowledge or skills might pursue.

As I have analyzed in more depth elsewhere, inventors capable of producing patentable advances (and the resources needed to support these inventors) may be scarce for a variety of reasons.[124] Knowledge in many science and engineering disciplines is now very complex and specialized, meaning that only a few persons may have mastered and be able to apply any particular body of knowledge to produce a related type of invention.[125] Advanced training in universities tends to track this same pattern of specialization, producing only a few parties with advanced training in particular disciplines.[126] Supporting resources needed for innovation in particular fields—including laboratory facilities, research staffs, administrative staffs, equipment, and supplies—are extensive and only possessed by a few corporations or universities.[127] The scarcity of the innovators or supporting resources in these various types of situations imply that the lure of patent rewards may be needed to attract these resources to the production of patent-eligible advances rather than to other productive activities that the resources might support.

3. Indicators That Invention Inputs are Not Scarce and That Normal Competition Will Probably Suffice to Produce Inventions

Conversely, some features of invention production processes or contexts tend to ensure that there are many potential providers of particular types of inventions and that, accordingly, there are not the sorts of conditions of inventor scarcity that need the special incentives and resource allocations achieved by the patent system. Where the skills necessary to produce new advances are widespread and costs of producing inventions are small, neither the relevant researchers nor the supporting resources are scarce resources. The likelihood of many new advances—including, as part of the overall mix, a substantial component of outlier advances—is high. Under these conditions, there will not be a case for application of the patent system.

Conditions of this sort are present, for example, in many business environments, where numerous parties may be capable of

124. *See* Gruner, *supra* note 112, at 17-23.

125. *Id.* at 19-21.

126. *Id.*

127. *Id.* at 23-24.

identifying and pursuing certain types of new business practices innovations and where the resources needed to try out new methods may be small. This combination of many potential innovators and sufficient supporting resources to produce and test advances suggests that business method patents may not be needed to spur advances in most business settings.

However, even though business method patents are not needed in all business settings, they may still serve a valuable resource allocation purpose in narrow business domains. Business method patents may have strong justifications in specific business settings where the generally prevailing conditions of abundant potential innovators and low innovation costs are not met.

For example, in some areas of highly rare business skills such as computer-enhanced financial methods, business innovators having both financial training and computer-based analysis skills and who are capable of working at the most advanced levels may be rare. The elite computer-savvy financial specialists capable of producing elaborate computer-based financial system advances are a scarce resource. Patent incentives may be valuable to encourage these rare parties to apply their skills and talents to the development of new financial services techniques. Assuming that the development of new financial methods is of substantial societal importance, then patent incentives can ensure that the rare skills (and associated computer resources) of high-level financial specialists are applied to the development of new financial methods rather than to other tasks of significant value to financial concerns and users of the services of those concerns.

Thus, even in a domain such as business methods where there are sometimes many possible innovators, there may be subdomains where innovators are scarce and patents are more justified. This will be the case where innovation depends on very specialized knowledge or unusual resources and there are very few parties with both the needed knowledge and the necessary supporting resources to produce outlier innovations.

This suggests the need for careful analysis of the merit of patents in subdomains of fields and that a categorical exclusion of all business methods from the patent system as non-patentable subject matters may not be wise. Rather, the appropriateness of patents for specific types of business methods should be screened through the context-specific knowledge and innovation technique standards of non-obviousness tests. By applying the latter to determine if a given

advance was one that only rare innovators could have produced (because the advance was beyond the knowledge and skill of persons with average or less than average abilities in the same field), business method patents and other patents in fields where patents are not generally needed to spur innovation can still be a force in exceptional subdomains where key analytic skills or innovation resources are rare. In these settings, the need for patent incentives to attract scarce innovators and scarce innovation resources to important tasks is particularly strong and the case for patent system inclusion is compelling.

C. When Administrative Costs Trump Patent System Advantages

Even where socially important inventions (with serious consequences if under produced) are present and there are reasons to believe that special incentives such as patent rewards are needed to attract scarce resources to the production of these important inventions, application of the patent system may still not be desirable where the costs of administering the system are large and these costs trump the apparent advantages of patent impacts under ideal conditions. If the administrative costs of granting and enforcing patents will overwhelm the social benefits of increasing the production of the important inventions in question, the patent system will not produce any net social gains and no patents should issue. Patent rewards should be recognized only where society realizes a net benefit in the increased production of patent-eligible advances even after patent system administrative costs are taken into account.

This subsection considers some of the circumstances where the costs of administering the patent system are likely to be large and when patent rights should be withheld in light of these probable costs.

1. Limits Based on Difficulties in Identifying Infringement and Triggering Rewards

One potentially important source of transactional costs in administering the patent system lies in accurately identifying unpaid access to patented advances as triggers for providing patent holders with remedies for that unpaid access. In traditional patent law terminology, the costs in this category include costs in accurately

determining patent infringement.[128] These may include costs of determining patent-specific standards for measuring infringement (including costs of claim construction as a preliminary to assessing infringement), costs in gathering and analyzing facts related to infringement determinations, costs incurred by defendants due to over-inclusive findings of infringement that force defendants to pay for access to patented technology where no access and advantage to the defendant has actually occurred, and costs related to under-inclusive findings of patent infringement that are paid by plaintiffs who are forced to accept unpaid access to patented technology (and to bear uncompensated research costs) where such access has incurred, but defendants are not legally required to pay for the access.

The costs of accurate determinations of infringement depend in part on the similarity of infringing and non-infringing items and the difficulty of telling them apart. Where distinguishing these (and providing remedies only for infringing items) is difficult, the tendency to mistarget and misapply patent remedies will be great. Hence, types of inventions that are little different from non-patented predecessors (or which are different only in ways that are difficult to detect) might be excluded from the patent system on the ground that the costs of accurate infringement findings and closely tailored patent rewards are too great to justify the invention production advantages of these findings.

Where, for example, a patented technology incorporates only small, difficult to detect changes in a prior technology, such a technology might be excluded from patent protections on the grounds that accurately targeting access controls and patent rewards for such a technology will regularly take more resources than the societal gains that are likely to result from more production of the inventions. Indeed, no greater production of such inventions will be likely if potential innovators perceive that infringement claims will be hard to press (because infringement will be hard to prove) and that infringement remedies will be curtailed accordingly. The same sort of logic might justify excluding advances that are used in circumstances that are frequently difficult for patent holders (or parties working on their behalf) to scrutinize in indentifying when infringement is present. Certain business methods, for example, are used so

128. Harry Surden, *Efficient Uncertainty in Patent Interpretation*, 68 WASH. & LEE L. REV. 1737, 1754 (2011) (noting the potentially significant transaction costs involved in estimating and assessing whether particular actions entail patent infringement).

frequently in private, concealed business environments that infringement will rarely be possible without major expenditures. The difficulty of identifying infringing activities and seeking corresponding remedies and patent-influenced rewards will undercut the incentives that patents can create for such advances.

Another source of difficulties and transaction costs in making infringement findings relates to the ability of potential defendants to disguise their products and activities so as to conceal infringing products and practices. The more defendants can misrepresent the nature of their products and practices in ways that utilize patented technologies while avoiding patent infringement remedies, the weaker that patent incentives become. Where defendants can disguise or conceal infringement, the costs of making accurate infringement findings go up (or the costs resulting from misapplied patent liability go up). At some point, the ease of disguising or otherwise concealing infringing items will be so great that infringement inquiries are not worthwhile and patent rights will also be a waste. This will be the case where the costs of accurate infringement findings overshadow the advantages of producing more patent-eligible advances in a particular area of technology.

2. Limits Based on Difficulties in Determining Proper Reward Size (Damages)

A patent-implemented invention access control system of the sort described in this article will only tie invention production incentives to invention access payments representing consumers' full measures of invention value if patent remedies for infringing conduct provide meaningful substitutes for voluntary access payments. This means that, where a party gains the benefit of invention use without paying access fees (in the form of licensing fees or patent-elevated product sales prices), the relevant patent holder should be able to claim damages for the party's infringing conduct that are roughly equal in amount to the gains the patent holder would have received from a voluntarily negotiated access fee. In this way, patent damages will provide a roughly accurate substitute for market-determined access fees and the rewards and incentives for invention production that those fees normally provide.[129]

129. Current standards for compensatory damages based on patent infringement provide for damages that are substitutes for privately determined invention access fees, thereby serving the function of patent damages described here. These damages are determined by reference to

Further, heightened damage awards—that is, punitive damages—may be needed to encourage parties with notice that they are about to engage in infringing conduct and a clear opportunity to gain access to an invention through payment of access fees to be penalized when they make the opposite choice and knowingly engage in infringement.[130] Awards of punitive damages in these circumstances will, like awards of compensatory damages, also necessitate determinations of the value that invention users place on access to patented inventions since the aim of the punitive damages will be to not just equal and offset these gains to intention users, but to exceed these gains and thereby create a net deterrent to knowing choices to infringe. This sort of deterrent will discourage parties from ignoring the production system for pricing patented inventions that is implemented through the patent system. It will discourage parties from adopting a business strategy of ignoring normal commercial channels for accessing patented inventions, channels that have been authorized by the patent holder and that play a role in pricing and rewarding the production of patented inventions. The likelihood of these heightened remedies and penalties will discourage and deter the next round of potential intentional infringers from ignoring market processes for invention pricing and the invention production system that these pricing mechanisms support.

A key problem and limitation in the administration of patent damages may be that the transaction costs of measuring the likely value and willingness to pay off infringers may be so large as to overwhelm the incremental advantages of more accurate estimates of user value and willingness to pay. Where greater accuracy in determining the willingness of defendants to pay consumes extensive resources in discovery or court processes, little, if any, advantage may

private transactions for determining the amounts that parties would pay for access to patented inventions, thereby making patent damages roughly equal to and a substitute for market determined fees for access to patented inventions. Under prevailing patent damage standards, compensatory patent damages are to equal no less than the licensing fees that a patent holder and infringer would have been likely to have determined in an arms length negotiation to be reasonable compensation for the infringer's use of a patented invention. *See, e.g.*, ResQNet.com, Inc. v. Lansa, Inc., 594 F.3d 860, 868-69 (Fed. Cir. 2010).

130. Punitive damages are available under the Patent Act. *See* 35 U.S.C. § 284 (2006). These sorts of enhanced damages are awarded to respond to (and deter) willful infringement and other culpable conduct. *See* Beatrice Foods Co. v. New England Printing & Lithographing Co., 923 F.2d 1576, 1579 (Fed. Cir. 1991) ("Under our cases, enhanced damages may be awarded only as a penalty for an infringer's increased culpability, namely willful infringement or bad faith.").

be achieved in the better targeting of incentives for invention. Better invention production incentives will not be realized in these circumstances because the net rewards to patent holders will not be significantly different than they would have been without efforts to measure the damages more accurately. The greater awards that the plaintiffs might ideally have received with more information on invention users' valuation of inventions and willingness to pay for access to the inventions will have been eaten up by discovery and court costs. The result will be that plaintiffs will be no better off than if less accurate determinations of defendant valuations of intention access were used in determining infringement damage awards.

This type of transaction cost analysis suggests two possible limits on the patent system. First, litigants could simply be left to take these sorts of net recovery dynamics into account in decisions about when to terminate litigation by settlements. Where the costs of further discovery and court proceedings do not seem likely to enhance the recovery that will be obtained (taking the costs of further litigation steps into account), the litigants can simply reach a settlement and maximize the plaintiff's net recovery as determined at that point in the development of the legal proceedings. There are at least two advantages to this approach. First, it allows the litigants to take into account diverse factors in assessing the strengths of a case on both sides and the likely litigation costs to come if a case is continued rather than settled. Second, it encourages the litigants themselves to formulate new valuation frameworks in considering case valuations and settlement terms.

A second possible solution to the problem of transactional costs in particularizing damage measurements lies in assessing damages for broad categories of defendants without worrying about the further factual findings necessary to particularize the findings to individual defendants within the categories. Categorical damage assessments will avoid the costs of achieving more particularized assessments of how much each defendant in a class would be willing to pay for access to a given technology. So long as the damages assessed for a class of defendants approximates the average for the class including the defendant, the amounts recoverable by plaintiffs will approximate the rewards that they would achieve through more individualized remedies (although some defendants will end up paying more or less than they would be required to pay as damages in a system that made more particularized assessments of damages and enforced patent remedies and rewards accordingly).

3. Limits Based on Difficulties in Administering Remedies

A further source of transaction costs that may justify some curtailment of the patent system relates to the costs of administering remedies when imposed. While damage remedies may be relatively easy to impose, injunctive relief may be particularly difficult to administer for several reasons.

First, it may be difficult to ascertain the actual nature of a defendant's past or future infringement and to tailor an appropriate injunction to prevent that infringement in the future. Where a given defendant has disguised the nature of its actual use of a patented technology, an injunction realized in litigation may restrict the apparent activity, but not the full scope of the defendant's uncompensated use of the technology involved. Where defendants are particularly good at disguising or counterfeiting activities, many instances of infringement may be overlooked and left out of injunctions. If this is the case, the mistargeting of injunctions (and the related under-compensation of patent holders as the defendants make uncompensated use of patented technologies outside of the mistargeted injunctions) will be particularly severe.

Second, even where an injunction is properly targeted to prevent improper use of a patented technology without compensation to the patent holder, policing adherence to the terms of the injunction may be difficult. Errors in such policing will either over or under compensate patent holders, with each type of error leading to social costs in over or under incentivizing inventions relative to socially optimal levels.[131]

Over-enforcement of injunctions (like overly broad findings of liability) will tend to advance the production of patent-eligible advances over the level that is socially optimal. This will, in effect, divert the use of some resources to the production of these advances where the resources would be better applied to other tasks to achieve better overall satisfaction of societal demands for both inventions and other products potentially produced with the same resources.

---

131. Over payments of rewards to inventors will, of course, not affect the incentives related to the remedies as these inventions will already be complete when the remedies are granted. However, such over payments will be witnessed by other inventors and similar overpayments will be anticipated by these inventors in the future. The result will be a tendency towards over production of the types of inventions associated with overpayments. In a similar fashion, underpayments will tend to lessen invention production below socially optimal levels.

The under-compensation of patent holders through the under-enforcement of injunctions may also produce several adverse consequences. If deviations from injunction terms have no practical consequences, the injunctions will have little or no effect in stopping uncompensated access to patented technologies and will not encourage users to pay for access to patented technologies. The greater the "leakage" of patented technologies into the hands of users in violation of injunction terms, the greater the uncompensated weaknesses in the invention production system because the incentives for creating similar inventions in the future will be perceived as being likely to be affected by similar leakage. As a result, incentives for innovation will be less substantial than would be the case with full rights enforcement. In addition, the ability of some invention users to gain access to patented technologies without paying for such access will encourage others to do the same (either because they no longer perceive much deterrence in the threat of patent litigation leading to impotent remedies or because they are in competition with the users gaining free use of the patented technology and feel the need to act like those users to keep up in the competition).

At some level of injunction enforcement costs, the costs of policing and seeking more accurate and defendant-specific judicial enforcement of injunctions will be greater than the incremental gains that the patent holders involved will realize. In these settings, withholding additional policing and enforcement will increase the net gains and incentives realized by patent holders and encourage invention production at greater levels than would be the case with more costly injunction enforcement.

VI. GENERAL PRINCIPLES FOR RECOGNIZING PATENTABLE SUBJECT MATTER: RECONSTRUCTING STANDARDS IN THE POST-*BILSKI* ERA

Patentable subject matter standards specify the range of practical discoveries that fall within the patent system and that are potentially subject to the types of production incentives and resource attractions that patent rights can provide.[132] As described in this article, the patent system is a tool for encouraging and regulating the production of patent-eligible advances at production rates that match public

132. *See* 1 DONALD S. CHISUM, CHISUM ON PATENTS: A TREATISE ON THE LAW OF PATENTABILITY, VALIDITY AND INFRINGEMENT § 1.01, at 1-7 (2011).

demands.[133] Where patents can serve this end, patentable subject matter should be recognized. Where achieving this end is impossible—because public demands for advances cannot be translated effectively into production incentives or because the costs of doing so overwhelm the public advantages of trying—the patent system should not be applied and advances should be found to be non-patentable subject matter. In this way, the purpose of the patent system in regulating invention production provides a guiding principle for the construction and interpretation of patentable subject matter standards.

Patent law is peculiarly in need of such a guiding principle. Patentable subject matter standards are at once ill-defined in current case law, yet needed in a clear form by numerous parties with stakes in the patent system and its operation. Persons who are strongly interested in clear patentable subject matter standards include potential patentees, patent holders and others concerned about the validity of present patents, and potential infringers who are worried about the scope of patent infringement liability. All of these parties are greatly disserved by the uncertainty of present patentable subject matter standards; all would benefit from greater clarity in these standards to inform future behavior.

This section describes the present lack of meaningful standards for patentable subject matter, assesses why past standards have been inadequate in light of the purposes of the patent system described in this article, and proposes a new patentable subject matter standard that will serve the purposes of the patent system more directly and completely.

A. Past Standards Reevaluated

This subsection provides a brief overview of past patentable subject matter standards that have each enjoyed brief acceptance, only to be rejected by reviewing courts that have found weaknesses in the standards upon full analysis and consideration. My object in this critique is to point out the deficiencies of these past standards in focusing the patent system and its incentives on the types of access controls and innovation production incentives described in this article. Following this critique, the next subsection proposes a new patentable subject matter standard tailoring the sweep of the patent system to

133. *See supra* Part IV.

advances that can be encouraged by the access controls and payment systems the patent system facilitates.

1. Supreme Court Standards

a. Benson Test

Much of the current confusion in patentable subject matter standards stems from the Supreme Court's cryptic analysis of such standards in the case of *Gottschalk v. Benson*.[134] In *Benson*, the Court held that a computer-based information processing method—divorced from any particular implementation in a specific computer or other piece of hardware—did not constitute patentable subject matter.[135] The invention at issue was a method for converting information from one computer-readable format to another computer-readable format.[136] The Court interpreted the claims of the disputed patent as covering all implementations of the method described in the patent, including implementations "not limited to any particular art or technology, to any particular apparatus or machinery, or to any particular end use."[137]

The Court's reasons for finding this advance unpatentable were described in very convoluted terms. Different portions of the Court's opinion in *Benson* seemed to focus on different reasons why the invention at stake was not patentable subject matter.[138] However, the discussions were very unclear, both individually and in their relationships to each other.[139]

One stated concern of the Court was that the information processing method at stake was no more than an unapplied idea.[140] The Court noted that just as "one may not patent an idea," one could not patent the "formula for converting [binary coded decimal (BCD)] numerals to pure binary numerals . . . ."[141] This discussion suggests

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

135. *Id.* at 71-72.

136. The patent at issue described the method sought to be patented as "a method for converting binary-coded decimal (BCD) numerals into pure binary numerals." *Id.* at 64.

137. *Id.*

138. For a complete critique of the *Benson* decision (including extensive criticisms of the Court's analyses), see Donald S. Chisum, *The Patentability of Algorithms*, 47 U. PITT. L. REV. 959, 971-91 (1986).

139. *See id.*

140. *See Benson*, 409 U.S. at 71.

141. *Id.*

that the Court was concerned with the lack of practical details in the patent claims regarding the implementation of the claimed invention and that a more concretely described and implemented computer-processing sequence might have passed muster.

The need for a complete description of a patentable advance remains an important consideration under the view of the patent system advocated in this article. To serve as the focus of access controls and user valuations of the sort described here, a patented invention must include sufficient practical details to distinguish the advance from other, similarly functioning items or processes. An advance lacking distinctive implementation characteristics would be hard to single out and control in an access system and hard for users to evaluate in order to determine how much they would pay for access. Ideally, to avoid limiting access to the wrong items or practices through the enforcement of patent rights, the description of an advance in a patent should be complete enough to distinguish the advance from other items and practices found in the same operative environments and potentially confused with the new advance. Hence, the Court's analysis in the *Benson* case provides some support for the type of patent system advocated in this article.

The problem with *Benson* is that the Court did not explain the basis for its concern over the lack of practical details in the invention at issue there. The Court's opinion failed to address why the lack of practical details in the advance under scrutiny precluded patenting of the advance. It also failed to discuss the types of practical design features that would have transformed the unpatentable idea present in *Benson* into a patentable invention. It is unclear from the Court's analysis whether a somewhat more concretely defined computer processing method (perhaps with additional concrete details on the computer software code needed to implement the method or the computer hardware equipment needed to implement the method) would have been sufficient to make the advance patentable subject matter. In short, the *Benson* decision indicates that inventors must go beyond mere ideas to create patentable subject matter, but provides little guidance in its discussion of unpatentable ideas about what distinguishes an unpatentable idea from a patentable implementation of the idea.

Unfortunately, the *Benson* Court did not limit its criticism of the invention at issue to concerns over patenting an idea. The Court confused matters further by mentioning several additional but ill-described grounds for finding the invention to be nonpatentable

subject matter.[142] The Court shifted its perspective several times in its analyses and identified several other reasons why the advance under analysis was not patentable subject matter, without indicating which if any of these rationales were independent grounds for rejecting the patentability of the invention in question. [143]

Some portions of the Court's opinion suggested that it was concerned the patent at issue might be interpreted to protect computer software *per se*. The Court noted that the invention at hand was described in terms of information processing steps (easily implemented in corresponding computer software code) without any further physical device details and that the Court was seeking to exclude such computer-implemented information processing methods from the patent system.[144] These portions of the opinion were taken by some observers as an indication that the Court was trying to announce (somewhat cryptically) that computer software was unpatentable.[145]

Yet another portion of the Court's opinion indicated that the information processing algorithm underlying the advance at stake was unpatentable because the algorithm held great promise for application within the computer field and should remain available for use by later designers of computer processes or devices rather than being restricted under the control of one party through patent rights.[146] The Court gave no indication of how to identify algorithms with this potential future importance, nor any explanation of why the development of such important algorithms should not be encouraged through the recognition of associated patent rights and incentives. The Court seemed to paradoxically note the importance of the method at hand, yet deny the application of patent rewards and incentives that would encourage the production of similarly important methods in the

142. See *Benson*, 409 U.S. at 72.

143. Chisum, *supra* note 138, at 971-92.

144. *Benson*, 409 U.S. at 71-72.

145. *See, e.g.*, Pamela Samuelson, Benson *Revisited: The Case Against Patent Protection for Algorithms and Other Computer Program-Related Inventions*, 39 EMORY L.J. 1025, 1030 (1990).

146. The Court noted that:

> The mathematical formula involved here has no substantial practical application except in connection with a digital computer, which means that if the judgment below [finding the patent covering the advance to be enforceable] is affirmed, the patent would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself.

Benson, 409 U.S. at 71-72.

future.

The confusion created by this multiplicity of rationales given by the Court to explain its decision in *Benson* (without any indication as to which were particularly important), as well as the lack of underlying explanations for why the Court was emphasizing many of these factors, left the *Benson* decision as a source of confusion rather than clarity regarding patentable subject matter standards. Lower federal courts (and to a degree the Supreme Court itself) have struggled to interpret and apply the holding in *Benson* for many years.[147] Not surprisingly, given the several rationales mentioned in *Benson* for rejecting the patentability of the method analyzed there, different parties have found support in *Benson* for many different patentable subject matter tests. Unfortunately, the rationales underlying the Court's analyses are so limited and unexplained that the *Benson* decision provides little help in determining the scope of patentable subject matter in new technological contexts. The concern expressed in *Benson* over the need for a practical advance with defined features is laudable as far as it goes, but the case does not make clear why we are concerned about these features. Nor does it specify how we should determine if a given invention has sufficiently defined features to make the invention patentable subject matter.

b. Chakrabarty Test

Diamond v. Chakrabarty[148] is generally understood and remembered as a case concerning biotechnology patenting and the importance of including living innovations within the patent system. The case is not limited to consideration of the patenting of items derived from nature, but rather offers some guidance regarding the minimum features of patentable subject matter generally. The discussion in *Chakrabarty* of the minimum features of patentable subject matter was embedded in the Court's evaluation of the invention at issue—a bio-engineered bacteria for use in cleaning up oil spills.[149] In its discussion of standards for identifying a new manufacture and how those standards would apply to the invention in dispute, the Court indicated that distinctiveness of an advance and identifiable new usefulness of an advance were the keys to

147. *See* discussion *infra* Part 2(a).

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

149. *Id.* at 308-10.

determining if the advance was patentable subject matter.[150] Regarding the patentability of the contested bacteria, the Court reasoned that the: "respondent's micro-organism plainly qualifies as patentable subject matter. His claim is not to a hitherto unknown natural phenomenon, but to a nonnaturally occurring manufacture or composition of matter—a product of human ingenuity 'having a distinctive name, character [and] use.'"[151]

In the last portion of this passage, the Court seems to be applying a general standard for identifying patentable subject matter. The Court indicates that an item is patentable subject matter if the item is 1) man-made (that is, a nonnaturally occurring product of human ingenuity) with 2) a distinctive name, 3) a distinctive character, and 4) a distinctive use. Such a man-made, distinctive item would qualify for a patent if other patent law standards (such as novelty, utility and non-obviousness of the item) were satisfied.

This standard tracks well with the tests that would limit the patent system to an access control and invention reward system of the sort described in this article. In order to be distinctly controlled (to limit access) as well as to be effectively evaluated by users for value and possible access payments, an advance must have a distinctive character and description (for both identification and segregation in access controls) and a distinct use (for evaluation of the incremental value of the distinctive use as a precursor to the willingness to make access payments). An advance with these features will be a good candidate for the sort of access controls that the patent system can implement.

There may still be concerns that this test will include too many technologies in the patent system because there may be few public benefits of extending patents to all distinctive, non-naturally occurring advances. If there are few public benefits from encouraging the production of more inventions of a particular distinctive type (or in encouraging more careful use of resources in the production of such inventions), these public benefits may not be worth the administrative costs and other transaction costs of implementing patent rights and rewards. These possible costs of broad patent system scope—and some of the means for reducing these costs—are addressed in the next subsection of this article describing a proposed patentable subject matter standard.

---

150. *See id.* at 310.

151. *Id.* at 309-10 (quoting Hartranft v. Wiegmann, 121 U.S. 609, 615 (1887)).

c. Bilski Test

In its 2010 analysis of patentable subject matter standards in *Bilski v. Kappos*,[152] the Supreme Court once again emphasized that an advance that is at heart no more than an information processing scheme should be considered an abstract idea and unpatentable.[153] In that case, the Court assessed the patentability of an invention involving a method of matching or "hedging" risks in investing in commodities.[154] This was seen as merely an abstract idea because the risk matching arrangements at the heart of the advance involved information analysis and processing.[155] The Court also emphasized that the addition of post-solution activity (that is, post-information processing actions) to record a result will not transform an information processing advance into patentable subject matter, nor will the limitation of use of an information processing advance to a specific industry or field of use produce patentable subject matter.[156]

The *Bilski* Court rejected the notion (adopted previously in the case by the Federal Circuit) that the presence of a specific machine or physical transformation in a process was needed in order to make the process patentable subject matter.[157] The Court found that the presence in a process of a machine or transformation was "an important and useful clue" as to whether the process constitutes patentable subject matter, but is not the sole test for such subject matter.[158] Unfortunately, the Court offered no indication about why the presence of a machine or physical transformation was important, even as a clue. It also left unstated how this clue should be interpreted. Is the presence of a machine or transformation a sufficient indicator of patentable subject matter or could some advances having one or both of these features still fail to qualify for patentable subject matter? Where the clue is missing—that is, where an invention fails to include either a machine or transformation—what other factors might still indicate that patentable subject matter is present? What are the necessary and sufficient features of patentable subject matter as seen by the Supreme Court? All of these key considerations were left

152. Bilski v. Kappos, 130 S. Ct. 3218 (2010).
153. *Id.* at 3228-31.
154. *Id.* at 3222.
155. *Id.* at 3231.
156. *Id.*
157. *Id.* at 3225-28.
158. *Id.* at 3226.

unresolved by the Court's glib reference to the (somewhat) rejected machine and transformation test as a "clue" to patentability.

The Court did adopt (without really defining) a general standard for patentability in *Bilski*. After rejecting the machine or transformation test as a sole test for patentable subject matter, the Court found the invention at issue in the case to not constitute patentable subject matter because claims to this invention "are attempts to patent abstract ideas."[159] Thus, the test the Court applied was to determine if an advance was a mere "abstract idea." The Court's earlier case law—most notably the *Benson* decision—had repeatedly held that abstract ideas, as opposed to applied ones, were not patentable.[160] The Court in *Bilski* provided this approving description of its earlier analyses in *Benson*:

> In *Benson*, the Court considered whether a patent application for an algorithm to convert binary-coded decimal numerals into pure binary code was a "process" under § 101. The Court first explained that "'[a] principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right.'" The Court then held the application at issue was not a "process," but an unpatentable abstract idea. "It is conceded that one may not patent an idea. But in practical effect that would be the result if the formula for converting . . . numerals to pure binary numerals were patented in this case." A contrary holding "would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself."[161]

The Court went on to explain its holding in *Parker v. Flook*[162] as another example of a case where the Court rejected the patenting of an advance because the advance was, at bottom, a mere abstract idea.[163] The invention at issue in *Flook* was an alarm system based on a novel use of a mathematical formula to analyze physical features of a chemical process and to activate an alarm when the process reached a particular state.[164] The *Bilski* Court saw this as an abstract idea (the formula) supplemented by "post-solution activity" (i.e., the steps needed to activate the alarm) which did not figure in the patentability

159. *Id.* at 3230.
160. *See* Gottschalk v. Benson, 409 U.S. 63, 64-67 (1972).
161. *Bilski*, 130 S. Ct. at 3230 (quoting *Benson*, 409 U.S. at 64-67, 71, 72).
162. Parker v. Flook, 437 U.S. 584 (1978).
163. *Id.* at 585-86.
164. *Id.* at 594.

analysis.[165] Hence, the invention, as pared down by the Court, was no more than a new formula applied in a particular context. Since a formula is an abstract idea, the advance put forth in *Flook* was seen as an abstract idea as well and was found to not constitute patentable subject matter.

As with the advances at issue in *Benson* and *Flook*, the invention in dispute in *Bilski* was found not to constitute patentable subject matter because the invention was no more than an abstract idea. The Court's brief analysis of the invention as possible patentable subject matter was as follows:

> The concept of hedging, described in claim 1 and reduced to a mathematical formula in claim 4, is an unpatentable abstract idea, just like the algorithms at issue in *Benson* and *Flook*. Allowing petitioners to patent risk hedging would pre-empt use of this approach in all fields, and would effectively grant a monopoly over an abstract idea.[166]

The Court also noted that the portions of the patent at issue in *Bilski* which specified the context where hedging was to occur and which claimed control over hedging practices only in the commodities and energy markets did not change the outcome of the patentable subject matter analysis.[167] These field-of-use limitations were not sufficient to change the abstract character of the claimed invention and even with these limitations the invention at issue was an unpatentable abstract idea.[168]

The *Bilski* test, like that in *Benson*, reflected a valuable concern that inventions be applied and not purely abstract in order to be patentable. Under the view of the patent system described in this article, the limitation of patentable inventions to applied designs serves to ensure that access controls related to patent rights will be limited to applied items and contexts and that potential invention users will have practical features of patented inventions to look to in evaluating the value of the inventions and the amounts that they are willing to pay for access to the inventions. Unfortunately, the Court's analysis in *Bilski* tells us little about what makes an invention an applied advance rather than an unpatentable abstract idea. The Court would have been more helpful and precise—and would probably have

165. *Bilski*, 130 S. Ct. at 3230.
166. *Id.* at 3231.
167. *Id.*
168. *Id.*

come to the same result—had it indicated that an abstract, unpatentable advance is one with so few specified features and results that patent holders will be unable to control access to the advance as claimed by the inventor and potential users will be unable to evaluate its value and determine an access price they are willing to pay.

Efforts to interpret and apply the Supreme Court's *Bilski* test in lower courts are still in their early stages. The Federal Circuit provided more details on how the *Bilski* test might be applied to distinguish a patentable advance from an abstract idea in *Research Corporation Technologies, Inc. v. Microsoft Corporation*.[169] The Federal Circuit's analysis in *Research Corporation* represents a significant effort to refine and give workable meaning to the *Bilski* test.

The inventions under consideration in *Research Corporation* were computer-based processes for rendering a halftone image on a computer screen.[170] The court found that these advances were patentable subject matter because they fell within the Patent Act's definition of a patentable process and because the advances did not fall within the judicially-recognized exception to patentability for abstract ideas.[171]

The court in *Research Corporation* saw the *Bilski* test as requiring a two step analysis of patentable subject matter.[172] First, an invention should be analyzed to determine if it falls within one of the statutory categories of generally patentable inventions—that is, to determine if the advance is a process, machine, item of manufacture, composition of matter, or an improvement of one of these.[173] Second, if an advance falls within one of these categories, the advance should be assessed to determine if it is within any of the special exclusions from patentability recognized by the Supreme Court, including the exclusion for abstract ideas.[174]

The Federal Circuit felt that an invention that falls within one of the statutory categories will normally constitute patentable subject matter unless there is clear evidence that the advance warrants special

169. Research Corp. Technologies, Inc. v. Microsoft Corp., 627 F.3d 859 (Fed. Cir. 2010).
170. *Id.* at 862-66.
171. *Id.* at 868-69.
172. *Id.* at 867-68.
173. *Id.* at 867.
174. *Id.* at 867-68.

exclusion from patentability.[175] With respect to the exclusion of an advance from patentability on the ground of that the advance is no more than an abstract idea, the Federal Circuit noted in *Research Corporation* that:

> this disqualifying characteristic should exhibit itself so manifestly as to override the broad statutory categories of eligible subject matter and the statutory context that directs primary attention on the patentability criteria of the rest of the Patent Act.[176]

In short, the court felt that most inventions that fall within the statutory invention categories addressed in the Patent Act should be deemed patentable subject matter and submitted for review under the other criteria for patenting (such as novelty, non-obviousness, and enablement tests) specified in the Patent Act.

In determining whether the advance at issue in *Research Corporation* was patentable subject matter, the Federal Circuit considered it important that the advance presented "functional and palpable applications in the field of computer technology."[177] The advance addressed a distinct functional need in the field of the advance—that is, a need in the field of computer display controls for a method of rendering halftone images on computer displays.[178] The advance included specifically described physical components, thereby distinguishing the advance from a mere abstract idea.[179] Furthermore, the court noted that the advance was an improvement over previous methods already used commercially to achieve the same end and that "inventions with specific applications or improvements to technologies in the marketplace are not likely to be so abstract that they override the statutory language and framework of the Patent Act."[180]

Refined as described by the Federal Circuit in *Research Corporation,* the *Bilski* test confirms the status of an advance as patentable subject matter if the advance meets the criteria to fall within one of the statutory categories of patentable inventions and the advance is not abstract as indicated by:

1) the functional ends served by the advance in its practical field;

---

175. *Id.* at 868.
176. *Id.*
177. *Id.*
178. *Id.* at 868-69.
179. *Id.* at 869.
180. *Id.*

> 2) the palpable components comprising the invention;
>
> 3) the need in the relevant field for the functionality provided by the advance; and
>
> 4) the likely commercial significance of the advance, as indicated by the existence in the marketplace of commercially successful items for which the advance is either a method of operation or an improvement.[181]

This version (or extension) of the *Bilski* test offers great promise as a means to implement and contain a system of invention access controls of the sort that the patent system was intended to advance. The limitation of patent controls to advances with palpable components and identifiable functional ends will tend to restrict patent rights to inventions that are capable of access limitations (which can be implemented by restricting access to the distinct functional results achieved by the invention or to the palpable components of the invention or to both).

The preexisting need in the field of the advance for a process with the functionality of the advance indicates that users are likely to value the advance and to be willing to pay an access charge to gain the use of the advance. The preexisting need suggests a user demand for inventions like the advance and a corresponding likelihood that users will pay rewards to inventors for satisfying the demand.

Finally, the presence in the marketplace of a predecessor to the advance indicates that potential users of the advance will probably assign a value to the advance in the marketplace, just as they already assign a value to the predecessor that is already found in the marketplace. This provides further evidence that a willingness to pay access fees and to create invention rewards is likely to follow the entry of the new advance into the marketplace even if access to the advance is somewhat limited by patent rights.

2. Federal Circuit Standards

a. Freeman-Walter Test

In a series of opinions in the late 1970s and early 1980s, the Court of Claims and Patent Appeals (as a predecessor to the Federal Circuit court) developed a test for patentable subject matter that

181. *Id.* at 868-69.

focused on whether an invention submitted for patenting involved an algorithm. The resulting standard was called the *Freeman-Walter* test in light of the two primary cases in which the test was set out.[182] This standard was intended to provide a more useful clarification of the test used (but poorly described) by the Supreme Court in *Benson*.[183] The test evaluated the presence of patentable subject matter in light of the following:

> First, the claim is analyzed to determine whether a mathematical algorithm is directly or indirectly recited. Next, if a mathematical algorithm is found, the claim as a whole is further analyzed to determine whether the algorithm is "applied in any manner to physical elements or process steps," and, if it is, it "passes muster [as patentable subject matter]."[184]

This two part test sought to ensure that inventions which were little more than algorithms for processing disembodied information were excluded from patenting. Where information processing algorithms or relationships were used in situations divorced from a particular practical application, the combined effect of the first and second parts of the test was to exclude the unapplied advances from patentable subject matter.

This type of exclusion was appropriate under the access control view of the patent system discussed in this article since an unapplied information processing algorithm lacks the sort of user utility that is the starting point for user evaluations of advances and a willingness to pay for access to the advances. Absent identifiable utility, users of an advance will not be interested in access and will not determine how much they are willing to pay for such access. An advance that lacked such a practical result and associated likelihood of user valuation is not amenable to the type of incentive system underlying the patent system.

If unapplied advances were accepted for patenting, incentives for the production of such advances might be set either too high or too low. A party developing an information processing method might over or under estimate the range of ultimate uses and the value that users would place on the advance since the range of applications of the advance (and the value of the advances) would not yet be

182. *In re* Freeman, 573 F.2d 1237 (C.C.P.A. 1978), *amended by In re* Walter, 618 F.2d 758, 766-68 (C.C.P.A. 1980).

183. *See* Gottschalk v. Benson, 409 U.S. 63, 64-67 (1972).

184. *In re* Pardo, 684 F.2d 912, 915 (C.C.P.A. 1982) (quoting *Walter*, 618 F.2d at 767).

determined. The developer of the advance might over or under estimate the practical ends to which the advance would be put. Absent an anchor for valuation in particular practical ends, the value of an unapplied advance (and the amount of resources that it would be cost effective to put into producing the advance) could not be determined with any accuracy. Absent at least some roughly accurate means to project the value of a targeted invention, the well-targeted yet regulated production of new advances through the type of patent-implemented access pricing scheme described here would not be possible.

Only once something more than just a pure, unapplied information processing method is developed (through linking the method to one or more practical results) can the value of an information processing advance be translated into consumer demand and patent rewards based on that demand captured through access controls and resulting reward payments to innovators. Hence, unapplied information processing algorithms or relationships lacking identified practical implementations were appropriately excluded from the patent system.

The *Freeman-Walter* standard looked to the physical application of an information processing algorithm as an indicator that an advance was an applied design, not merely an abstract idea in the form of an unapplied algorithm.[185] However, this test was both over- and under-inclusive of the advances that should have been included in the patent system. Some items involving physical applications of algorithms would not be susceptible to the type of access controls and valuation systems administered through the patent system. Conversely, some applications of algorithms would be amenable to access controls (through controls over the results of using the algorithms), valuations, and rewards to inventors of the sort that the patent system can implement. Hence, as noted by the Supreme Court in *Bilski*, physical elements serve as partial, but incomplete criteria for distinguishing between patentable subject matters (for which the patent system is a useful reward and production mediation tool) and unpatentable subject matters other useful advances (which are best produced without the access limitations and administrative costs of the patent system).

185. *Id.*

b. State Street Test

Recognizing the weaknesses and incompleteness of the *Freeman-Walter* test, the Federal Circuit eventually tried again and reformulated its patentable subject matter standard. In a standard articulated by one of the great masters of patent law, Judge Giles Rich, the Federal Circuit announced and refined a new patentable subject matter standard in a series of decisions. The standard was first mentioned in passing in *In re Alappat*,[186] discussed further in the context of a computer-based business method in *State Street Bank & Trust Co. v. Signature Finance Group, Inc.*,[187] and discussed and explained more completely in connection with a pure information processing advance in *AT&T Corp. v. Excel Communications, Inc.*[188]

While it was applied by the Federal Circuit across several types of inventions in these cases, the standard that emerged was most commonly associated with the *State Street* case because the standard was used there to uphold the patenting of a business method and this result received considerable notoriety.[189] This standard—referred to here as the *State Street* test for convenience—was aimed at distinguishing abstract ideas from patentable applications. The standard specified that an advance was patentable subject matter if the

186. *In re* Alappat, 33 F.3d 1526 (Fed. Cir. 1994) (en banc). *Alappat* involved a computer system that controlled the display on a cathode ray tube screen. The system evaluated electronic signals and determined how to best control the screen to display a graphic image corresponding to the signals. The only new feature of the system relative to earlier cathode ray tube systems were the information processing steps implemented by the computer system. The Federal Circuit found that this invention constituted patentable subject matter because the invention was "a specific machine [that produces] a useful, concrete, and tangible result." *Id.* at 1544.

187. State St. Bank & Trust Co. v. Signature Fin. Grp., Inc., 149 F.3d 1368 (Fed. Cir. 1998), *cert. denied*, 525 U.S. 1093 (1999). *State Street* involved a computer-based system implementing a particular financial investment management method. The invention in this case helped to administer a "hub and spoke" method for the central investment of funds from multiple financial institutions. The system provided for the collection of funds from multiple institutions (the "spokes") for investment by a single central fund (the "hub"), with frequent status reports made to the contributing institutions. The Federal Circuit found this system to be patentable subject matter because the system met the standard announced in *Alappat*—that is, because the computer-based system was a specific item that produced a useful, concrete, and tangible result. *Id.* at 1373.

188. AT&T Corp. v. Excel Commc'ns, Inc., 172 F.3d 1352, 1353 (Fed. Cir. 1999), *cert. denied*, 528 U.S. 946 (1999). The innovation in AT&T was a new electronic record keeping method for recording information on long distance calls. The method was held to be patentable subject matter because the method was a specific method producing a useful, concrete, and tangible result. *Id.* at 1357-58.

189. *See e.g.*, Kevin Schubert, *Should* State Street *be Overruled? Continuing Controversy over Business Method Patents*, 90 J. PAT. & TRADEMARK OFF. SOC'Y 461 (2008).

advance was 1) a specific machine, item of manufacture, composition, or process that, when used, produced a 2) useful, 3) concrete, and 4) tangible result.[190]

The Federal Circuit explained its objectives in applying the *State Street* test as follows:

> In *State Street*, this court, following the Supreme Court's guidance in *Diehr*, concluded that "[u]npatentable mathematical algorithms are identifiable by showing they are merely abstract ideas constituting disembodied concepts or truths that are not 'useful.' . . . [T]o be patentable an algorithm must be applied in a 'useful' way." In that case, the claimed data processing system for implementing a financial management structure satisfied the § 101 inquiry because it constituted a "practical application of a mathematical algorithm, . . . [by] produc[ing] 'a useful, concrete and tangible result.'"
>
> The *State Street* formulation, that a mathematical algorithm may be an integral part of patentable subject matter such as a machine or process if the claimed invention as a whole is applied in a "useful" manner, follows the approach taken by this court en banc in *In re Alappat*. In *Alappat*, we set out our understanding of the Supreme Court's limitations on the patentability of mathematical subject matter and concluded that:
>
> > [The Court] never intended to create an overly broad, fourth category of [mathematical] subject matter excluded from § 101. Rather, at the core of the Court's analysis . . . lies an attempt by the Court to explain a rather straightforward concept, namely, that certain types of mathematical subject matter, *standing alone*, represent nothing more than *abstract ideas until reduced to some type of practical application*, and thus that subject matter is not, in and of itself, entitled to patent protection.
>
> Thus, the *Alappat* inquiry simply requires an examination of the contested claims to see if the claimed subject matter as a whole is a disembodied mathematical concept representing nothing more than a "law of nature" or an "abstract idea," or if the mathematical concept has been reduced to some practical application rendering it "useful." In *Alappat*, we held that more than an abstract idea was claimed because the claimed invention as a whole was directed toward forming a specific machine that produced the useful,

190. *See AT&T*, 172 F.3d at 1357-58.

concrete, and tangible result of a smooth waveform display.[191]

While the Federal Circuit related its *State Street* standard in this way to concerns over not patenting abstract ideas, the court still did not explain why patenting abstract ideas was a problem or how its test for identifying a non-abstract and patentable idea was effective in achieving the goals of patent law. Without a sense of the objectives of the *State Street* standard, there was little way for subsequent courts or other analysts to determine the meaning of the components of this standard and to apply to standard in an effective, consistent way.

Ultimately, the Federal Circuit rejected the *State Street* standard in its opinion in *Bilski*, in part because the court felt that the *State Street* standard was too inclusive and brought too many types of advances within the patent system.[192] Yet the Federal Circuit speaking in *Bilski* did not articulate criteria for determining what constituted an excessively broad sweep of the patent system and why the *State Street* standard produced such a consequence. Indeed, while the Federal Circuit had purported to apply it for some years, the court rejected the *State Street* standard in its *Bilski* opinion with little analysis and explanation of the reasons for rejecting a well-established standard that had been frequently applied in lower courts and in patent analyses by private parties.[193]

The only reasoning the Federal Circuit offered in its *Bilski* opinion to explain its rejection of the *State Street* test was to note that this test focused too much on the results achieved by an advance and not on the features of the advance itself.[194] However, this may have been a criticism resulting from the Federal Circuit's consideration of only part of the *State Street* standard. The Court mistakenly referred to the standard as requiring only that an advance produce a useful, concrete and tangible result in order to constitute patentable subject matter.[195] The full *State Street* standard required such results plus that an advance constitute a specific machine, item of manufacture, composition, or process.[196] This further required element—if interpreted to exclude advances lacking specific practical details from

191. *Id.* at 1357-58 (citations omitted) (alterations in original).

192. *In re* Bilski, 545 F.3d 943, 959-60 (Fed. Cir. 2008), *aff'd*, 130 S. Ct. 3218 (2010).

193. For an overview of the history of the application of the *State Street* standard and some of its implications, see Schubert, *supra* note 189.

194. *Bilski*, 545 F.3d at 959-60.

195. *Id.* at 959-60, 960 n.19.

196. *AT&T*, 172 F.3d at 1357-58.

patenting—might have allayed the concerns of the Federal Circuit that the *State Street* test was inadequate to prevent abstract ideas from receiving patents.

While it was technically not before the Supreme Court in *Bilski* (having been rejected below by the Federal Circuit),[197] the *State Street* standard was nonetheless disavowed (if not outright rejected) by all members of the Supreme Court in various opinions issued in the *Bilski* case.[198] However, the precise reasons for the Court's doubts regarding the *State Street* standard were not articulated, perhaps because the Court's goals for a patentable subject matter standard were ill defined and, absent these goals, there was no clear reasoning that the Court could present as to why the *State Street* test did not serve patent law goals.

Viewed from the perspective of this article, the *State Street* standard captured some of the criteria needed in a well-focused standard for patentable subject matter. The *State Street* standard limited patentable advances to specific inventions that were capable of producing useful, concrete and tangible results. Inventions meeting this test would tend to be ones that could be incorporated in workable access controls and valuation methods.

The requirement in *State Street* of a specific device or process as a threshold feature of patentable subject matter suggests that patentable items must be sufficiently defined to be distinguished from other items, which is a necessary characteristic if access controls (based on patent rights) are to be imposed on the patented items but not on other like items. A sufficiently specific advance in this regard is one where the characteristics of the device or process are stated in patent claims with sufficient detail that courts, litigants, and potential actors can be sure when a patented item is present rather than just another item which is based on the same unpatentable abstract idea or information relationship. Had the somewhat ambiguously stated first prong of the *State Street* test been interpreted this way, it would have been an important feature of a test appropriately limiting patentable subject matter to advances where access controls and the types of rewards furthered by the patent system are administratively possible.

Likewise, the requirements under *State Street* that a patentable

197. *See Bilski*, 545 F.3d at 959-60.

198. *See* Bilski v. Kappos, 130 S. Ct. 3218, 3231 (2010); *id.* at 3232 n.1 (Stevens, Ginsberg, Breyer, and Sotomayor, J., concurring); *id*. at 3259 (Breyer and Scalia, J., concurring).

advance produce a useful, concrete, and tangible result also seem likely to further the invention production goals of the patent system if these requirements are properly interpreted. Clearly, a useful result is necessary if invention users are to see value in access to an advance and to be willing to pay a price for such access. A concrete result—in the sense that the result is not uncertain, but rather regularly produced by use of the advance and manifest in a well-defined, concrete way—is needed for invention users to be willing to assess value from initial uses of an advance, to project that they will gain similar advantages from similar uses in the future, and to be willing to pay for future access to gain the same advantages in the future. The requirement that the result of an advance be tangible—having a useful impact in interpreting or working with objects in the tangible world—ensures that patent rights and associated access controls are restricted to either physical devices and processes or results that are important in measuring or managing physical circumstances (with results tied to those circumstances, which will distinguish them from pure intellectual thoughts or information analyses). This last requirement of tangible results—that is, results affecting or relating to the tangible world—could, if interpreted to restrict patents to advances where the results are trackable and access to them controllable—serve to limit patent rights and controls to advances that are amenable to the types of access controls and production incentives that can be administered through the patent system. Hence, while it was very ambiguous in its initially articulated form, the *State Street* test had the rudiments of criteria for desirable bounds on the patent system. Some clarifying and narrowing interpretations, keeping in mind the goal of ensuring workable invention access controls and access payment mechanisms, might have saved the *State Street* test as an administrable and desirable patentable subject matter test.

c. Machine-or-Transformation Test

In the Federal Circuit's *Bilski* decision, the court sitting *en banc* rejected the *State Street* standard as too ill defined and too broadly inclusive of patentable subject matter.[199] Instead, the court substituted a "machine-or-transformation test" that it felt was dictated by earlier Supreme Court precedents.[200] Although the Supreme Court later held the machine or transformation test was not the sole criteria for

199. *See Bilski*, 545 F.3d at 959-60 nn.18 & 19.

200. *See id.* at 961-63.

determining the scope of patentable subject matter,[201] it did find some value in the test as a "clue" to patentability.[202] The Federal Circuit's analysis in *Bilski* was as follows:

> A claimed process is surely patent-eligible under § 101 if: (1) it is tied to a particular machine or apparatus, or (2) it transforms a particular article into a different state or thing. . . . [This "machine-or-transformation test" is] the governing test for determining patent eligibility of a process under §101.[203]

This machine-or-transformation test was rejected by the Supreme Court as the sole patentable subject matter test on the ground that it improperly barred patents for some advances that were intended by Congress to fall within the range of the patent system.[204] However, the Supreme Court gave little indication of how it saw the purposes of the patent system envisioned by Congress or why the machine-or-transformation test failed to serve these purposes.

As with the emphasis on physical elements or steps in the Federal Circuit's *Freeman-Walter* test for patentable subject matter, the emphasis on physical elements or impacts of a process in the machine-or-transformation test was a partial step towards a desirable patentable subject matter standard, but only if these sorts of physical features are interpreted and analyzed in the context of an invention access control system. Physical features of an advance may be important in assessing whether an advance should be seen as patentable subject matter if those elements (or their absence) indicate that uses of the advance will be easy to limit and that, therefore, an access limiting system for such an advance will be reasonable to administer through patent rights. Physical elements will frequently aid in ensuring that related invention embodiments or the results that they achieve can be controlled and access given to those embodiments only for persons who pay for such access. However, such physical elements may not necessarily be co-extensive with the presence of an excludable advance that can be subject to meaningful access controls. The ultimate standard should be workability of access controls.

Absent an overarching criteria such as this for assessing the character of useful advances and dividing them into patent-eligible and non-eligible subsets, focusing on relevant types of elements such

201. *See Bilski*, 130 S. Ct. at 3225-27.
202. *See id.* at 3226-27.
203. *Bilski*, 545 F.3d at 954, 956.
204. *See Bilski*, 130 S. Ct. at 3225-27.

as physical elements or effects will be meaningless. The presence or absence of these elements will lead to over- and under-inclusive decisions on patentability. Some inventions with physical elements may not lend themselves to easy exclusion from use and, therefore, should probably be omitted from the patent system by finding them to be non-patentable subject matter. In contrast, some intangible advances may be sufficiently tied to physical circumstances, physical implementations, or other physical use features to provide means for excluding users from the processes or from the results of those processes and to thereby provide viable access controls and access payment mechanisms. Such excludable advances (even without substantial physical instantiation) should be treated as incorporating patentable subject matter.

B. Summarizing Decades of Confusion in Patentable Subject Matter Tests

The Supreme Court's opinion in *Bilski* did little more than refer backwards to the Court's own prior case law on patentable subject matter—case law that that had spurred great confusion in lower courts for decades as to what criteria to use for analyzing patentable subject matter. The Court's analysis did not provide any meaningful new guidance and has left the analysis of patentable subject matter in an even more uncertain state than before the Supreme Court's *Bilski* decision. Widespread uncertainty and confusion in patentable subject matter analyses began for courts, the USPTO, and practitioners with the Supreme Court's decision in *Benson* in 1972.[205] Since then, we have seen the rise and later rejection of the *Freeman-Walters*, *State Street*, and machine or transformation tests, all viewed in their time by the Federal Circuit as needed clarifications of the uncertain standards left behind by the Supreme Court's case law.

With the Supreme Court's refusal to add any new guidance regarding standards in *Bilski*, we once again return to this uncertain Supreme Court case law for its very limited guidance on how to conduct patentable subject matter standards. Patentable subject matter analyses—one of the most important types of evaluations in patent law, governing the range of the patent system itself—rest on underpinnings that have been widely recognized as weak and ill defined for over thirty years. The incentive goals of the patent system

205. *See* Chisum, *supra* note 138, at 992.

and the interests of those who rely on this system for certainty about the system's probable rewards and restrictions deserve far more clarity in patent law standards than the efforts of the Supreme Court and the Federal Circuit have produced.

The primary reason why judicial efforts to define criteria for patentable subject matter in recent years have failed is that these efforts have not been grounded in the objectives of the patent system. Since patentable subject matter standards effectively set the boundaries of the patent system and its impacts, patentable subject matter scope should conform to the objectives of the patent system to set the proper boundaries for the system through the articulation of patentable subject matter standards. If we do not pay attention to the goals of the system, how can we assess what technologies are valuably included or excluded from the system?

The notion that Congress intended to include a wide range of useful advances within the incentives of the patent system suggests the proper framework for defining patentable subject matter standards. Congress implemented the patent system as a means to raise incentives for the creation of useful items and processes that are outliers in their fields, with the incentives achieved through the sorts of access controls and payment systems described in this article. Given this incentive purpose (and the lack of any explicit technical boundaries on the achievement of this purpose detailed in the Patent Act), Congress's apparent aim was that the patent system should be coextensive with the system's purpose—that is, that the patent system should serve a technology-enhancing function wherever the system is capable of doing so. Accordingly, Congress's intention was for the patent system to apply to all types of useful advances that can be encouraged through the sorts of access controls the patent system can implement (but also that no other technologies be included).

This purpose provides a guiding principle for the construction of patentable subject matter standards. Such a technology-neutral principle offers a means to construct patentable subject matter standards that focus on the practical features of advances which will fall within the patent system without knowing in advance what the advances will entail or how they will operate (e.g., whether they are tangible or not) and without arbitrarily tying the patent system only to types of technologies that resemble industrial or technical discoveries of the past. Such neutral criteria should—given the forward looking goals of the patent system to promote the production of new advances that are non-obvious departures from our current knowledge—help to

keep the patent system as a force for the encouragement of presently unforeseen technologies as "times change" and fundamentally new fields of useful advances are possible.[206]

C. Towards a Revised Patentable Subject Matter Standard

This subsection uses the neutral principle of the suitability of advances for access controls and access payments as a criterion for the construction of new patentable subject matter standards. It is hoped that these proposed standards will fill the void left behind by Supreme Court precedent to date and will be extendable as needed based on interpretations and modifications made in light of new technologies and social practices. With such forward-looking, yet well-defined features, patentable subject matter standards can tailor the scope of the patent system to advances for which the system can serve the incentive function envisioned by Congress and the drafters of the Constitution.

1. Goals of a Revised Patentable Subject Matter Standard

The standard proposed here is a reflection of the view explained in this article that the patent system exists to encourage and regulate the production of patent-eligible advances. The proposed standard includes practical advances within the range of patentable subject matter if the production of advances can be effectively included in access controls and access payment practices that will incentivize the production of patent-eligible advances as public goods at production rates that match public demands. Where an advance is capable of access controls that lend themselves to such a payment and production modulating system, the advance should be considered patentable subject matter. Conversely, where advances or their uses

206. The need to keep the patent system forward looking and to ensure that it will provide incentives for as yet undiscovered types of advances was recognized by a plurality of the Court in *Bilski* which notes:

> [T]imes change. Technology and other innovations progress in unexpected ways. For example, it was once forcefully argued that until recent times, "well-established principles of patent law probably would have prevented the issuance of a valid patent on almost any conceivable computer program." But this fact does not mean that unforeseen innovations such as computer programs are always unpatentable. Section 101 is a "dynamic provision designed to encompass new and unforeseen inventions." A categorical rule denying patent protection for "inventions in areas not contemplated by Congress . . . would frustrate the purposes of the patent law."

Bilski, 130 S. Ct. at 3227 (alteration in original) (citations omitted).

are such that this type of access limitation and production incentivizing cannot be realized effectively—because demands of the public for advances cannot be translated into production incentives or because the costs of doing so generally overwhelm the corresponding advantages—then the patent system should not be applied and the advances should be determined to constitute unpatentable subject matter. In short, the feasibility of access controls (and related access payment systems) should govern the overall scope of the patent system and the range of patentable subject matters.

A system of invention production incentives of the sort described in this article will need to include 1) a means for potential users to recognize an advance as a distinct item from other similar items performing similar functions, 2) sufficient present utility (and projectable future utility) for the users to evaluate the merit of the new advance (relative to the merit of other items that perform the same function) and to determine an amount that they are willing to pay for access to the new advance and its new incremental utility, 3) a mechanism for controlling access to the advance such that persons will seldom gain access to the advance without paying for such access, and 4) a mechanism for transferring access payments from users of an advance to the inventors of the advance, either directly or through intermediate payment steps. With these four components, a system of access controls and associated access payments will tend to encourage the production of new advances in accordance with the value judgments and demands of potential users of the advances. Absent any one of these features, the link between enhanced user value and rewards to inventors will be broken and the invention incentivizing and resource allocation goals of the patent system will not be served.

2. Minimum Features of Technologies Susceptible to Production Incentives Through Patent Rewards

Given that the main purpose of the patent system is to set prices for access to patented technologies and to thereby encourage the production of patent-eligible advances at levels that meet public demand for such inventions (and at levels that minimize opportunity costs suffered by the public when such inventions are under-produced), the patent system should apply to technologies and production settings where this access pricing is possible and likely to be effective in modulating invention production. Hence, a minimum feature of an invention constituting patentable subject matter should

be that the invention's production can be encouraged through the types of access controls and access pricing encompassed by patent laws. This implies further that patentable subject matters should be limited to advances for which there are sustainable methods of exclusion from use and further means of providing selective access only upon payment of relevant access charges.

There are several basic features that a technology must have in order to be susceptible to this type of production incentive system. Absent any one of these features, it seems unlikely that a system of invention access controls and related invention access pricing will produce rewards to innovators that will encourage invention production:

1) *Distinct Definition of Invention*: In order to be the focus of user valuation assessments and related access payment arrangements, an invention will need to be separately identifiable. That is, the invention will need to be defined with sufficient particularity to be understood in its practical operation and to be separated from other similar items or processes. Distinctive definition will be a minimum feature of patentable subject matter for several reasons.

   First, the definition of an invention should be sufficiently particularized and linked to practical results to give potential users means to analyze the results and their related utility. These utility assessments will be the basis for invention valuations by potential users and, therefore, the willingness of potential users to pay for access to the inventions.

   Second, the definition of an invention constituting patentable subject matter should include concrete implementation details, such as workable descriptions of the invention's structure (if any) and operation, so as to give potential users means to understand the steps and costs of making, using, and maintaining the advance. Often, the net utility to users of an advance will depend significantly on the costs of making the advance (or acquiring it) and using it, as well as the advantages of using it. The net advantages of use of a new advance can only be understood with information on all of the factors affecting these net advantages, including the means of use and the resources needed to support such use. Hence, concrete knowledge and description of the structure and operation of an advance are important minimum features of

patentable subject matter, needed to provide key informational inputs to user value determinations. These details must be both known (an issue of patentable subject matter) and described completely in a patent application (an issue of the sufficiency of invention disclosures and descriptions in the relevant patent specification).[207]

Third, the definition of an invention will need to be particular enough to permit viable access controls and access payment mechanisms to be implemented and administered consistently and without large fact-finding costs. Absent a detailed definition of an advance—including clear means to distinguish it from unrestricted items with similar structural features or usage results—attempts to control access to the advance as part of an access payment system may tend to be over- or under-inclusive, improperly sweeping up other non-patented items within access controls or failing to control all instances of access to the patented advance and thereby allowing some users free access to the advance.

Two dimensions of distinctiveness in defining a patentable advance may be relevant here: distinctiveness as to structural elements of the advance and distinctiveness as to functional results achieved by using the advance. A complete description of a patentable advance should include sufficient practical details to both distinguish the patented item from similarly functioning items with different structural features and to distinguish the patented item from similarly structured items with different functions. A description lacking these sorts of distinctive details regarding the structures or results of a new invention will provide little means for invention-specific access controls of the sort needed in applying the patent system effectively to incentivize and modulate invention production.

Complete invention definitions, fully comparing an invention

207. Even where an advance constitutes patentable subject matter (and might therefore qualify for a patent under a complete description of the invention and upon meeting other patent law tests), an inventor may fail to gain a patent (or may obtain a patent that can later be voided) if the description of the invention provided in the inventor's patent application (in the portion of the application commonly referred to as the invention "specification") is not sufficient to inform average practitioners in the field of the invention how to make and use the invention. *See* 35 U.S.C. § 112 (2006).

to prior designs and then clarifying the new features which distinguish the advance from its predecessors, may also aid potential users of the advance in assessing the incremental utility associated with the advance. The nature of items used previously for accomplishing the same ends—often unpatented items free for all to use without patent-related payments—should be apparent from the discussion in invention definitions of the predecessors. By limiting patents to advances that involve clearly articulated practical distinctions from prior item designs, potential users of the advances will be able to focus their attention on the new features of the advance that distinguish it from its predecessors and that potentially provide new utility to the users. This will, in turn, support the valuation of the new utility by potential users and the willingness of users to pay for access to the new advances.

2) *Regular Invention Operation*: An invention constituting patentable subject matter should operate in a regular manner to achieve predictable utility if users are to have confidence in the future utility of the advance and to be willing to pay for access to that future utility. Absent this type of regular operation and functional results, potential users of an invention will not be able to project the value of future access to the invention (because they will lack confidence that presently achieved utility from the advance will correspond to future utility and value) and will be unwilling to set and pay a price for access to the invention.

3) *Manifest Utility Coextensive with Invention Definition*: A patentable invention should produce identifiable results and manifest utility when used in accordance with its definition so that potential invention users can assess how much they wish to pay to gain access to the invention and the benefits it produces. Absent manifest utility (as evidenced through the results the invention presently produces or such further results as can be projected reasonably when the invention is brought to the public through patent disclosures), the type of access valuations and pricing that can be administered through the patent system will produce few if any rewards to inventors, making the costs of administering and enforcing patent rights simply wasted resources.

The need for manifest utility to support invention valuation processes explains why advances that are defined only in highly abstract terms without important implementing details should not be seen as patentable subject matter. These advances lack the types of implementing details that correspond to immediately available utility in using the advances. Sufficient practical details should be present to dictate identifiable results from use of advances, such that the utility of these results and value from using the advances can be assessed. This will allow potential users of the advances to make valuation determinations and to arrive at access payment decisions. Lacking such bases for evaluation, advances should not be admitted to the type of access pricing system implemented through patent rights.

4) *Means to Limit Access*: A patentable invention and the settings in which it is used should be such that there are means to prevent access to the invention (or at least use of the invention) by persons who do not pay for such access (through either license fees or patent-influenced purchase prices). Absent means to prevent access on this sort of conditional basis, the type of access pricing that the patent system can support will not be possible. If persons who do not pay are able to gain access on a similar basis to those who do, there will be little incentive to pay for such access. Indeed, those persons who pay for access and are then forced to compete with persons who gain access without paying will be disadvantaged since both groups will have access to the functionality of the invention in question, but only those who have paid will bear the costs of gaining access. Furthermore, absent clear means of preventing access to inventions without access payments, parties capable of producing new advances may see little way to gain returns on their innovation investments and may forego inventive efforts rather than effectively giving away their costly inventions in the absence of access controls.

5) *Payment Mechanisms*: The manner and settings in which an invention is used must be amenable to pricing of access and transfers of access payments to invention originators. The intermediate steps through which this will occur may be complex and may imperfectly transfer the prices users pay for

> access to the parties who produce inventions. Nonetheless, some mechanisms for transferring most of these access charges to inventors (or the organizations that fund inventive efforts) are needed to encourage optimal levels of invention production.

The more imperfectly these mechanisms work—that is, the more that the costs of administering payment systems filter out components of payments and cause only fractions of amounts that users pay for access to be passed on to invention originators—the less effectively the patent system will encourage and regulate invention production in accordance with user demands.

Furthermore, the more that payment mechanisms work on categorical bases—for example, specifying a single price for a patented item that is sold to a group of users where each user in the group has a somewhat different value for use of the item—the more likely it is that some invention users will have unmet demands for additional inventions. This will tend to be the case because no user will be likely to pay more for invention access than that party feels such access is worth, but some potential invention users might be willing to pay more for additional inventions that are not produced at the access price that the relevant group of users is willing to pay.[208] These additional inventions will not be produced if access pricing mechanisms are not particularized to the individual level and the additional demands of those users at higher prices are not translated into additional demands seen by invention originators.

This is another way of saying that the ideal form of the patent system would be one where there is perfect price discrimination of invention access across potential invention users—where each user pays what he or she actually thinks that use of the invention is worth. Such a system would present incentives to inventors that would encourage the creation of inventions filling the broadest range of user demands and minimizing opportunity costs due to the under-production of inventions which would have cost less to produce than their values to users.

However, such a system of perfect price discrimination for access pricing would be too costly to administer and we generally will make do with categorical assessments of access prices—that is, with a

208. This is a result of imperfect price discrimination of access pricing in a system where access to an invention is made available to a group of users at a single price.

system providing access to patented items across categories of users at prices that reflect the willingness of substantial number of users in each category to pay for access. Groupings of users in particular categories (and policing different pricing of invention access by category) will depend on the costs of identifying and separating different types of invention users and in charging each type a different price for use of an invention. Where these costs are high, efforts to distinguish between users will tend not to be made, meaning that the users involved will be lumped into one group and charged one price for invention access rather than being subject to price discrimination.

6) *Modest Transaction Costs in Administering Patent Rights*: The types of invention access controls and use payment mechanisms described in the above criteria must be capable of implementation without large transaction costs if demands for new inventions are to be translated accurately into incentives for invention production. If transaction costs of implementing access controls and access payment mechanisms are large for a particular category of advance, then these costs may soak up amounts that users will pay for access, leaving little or no net rewards for invention users. Such a result would negate the production incentive function to be served by patent laws and rewards. Indeed, transaction costs that soak up even a substantial fraction of the amounts users are willing to pay for inventions may substantially skew the production of such inventions away from optimal levels. Such costs—even if they leave behind some net rewards for invention producers—will tend to make these rewards so low as to cause many potential inventors to see other uses of their resources as producing greater potential returns and to divert their efforts away from potential inventive efforts.

Transaction costs in administering patent rights may arise from several features of innovations. High administrative costs due to any of the following factors may justify excluding advances from patentable subject matters.

Feasibility of Access Controls: Generally prevailing costs of preventing access to inventions for persons refusing to pay for such access might justify excluding some types of advances from patentable subject matter. Such costs may be high where advances are not susceptible to physical access limitations or other similarly effective means for preventing unpaid access.

Advances involving new devices or compositions of matter should generally be amenable to effective access controls since access to units of the devices or compositions can be restricted through physical constraints and access limitations. Similarly, low cost access controls for processes that achieve physical transformations of items will often be feasible. This will be the case because these sorts of physical transformations will typically be realized by physical equipment and access to the relevant patented process can be controlled by limiting physical access to the corresponding equipment.

Limiting access to method advances with no physical embodiments or transformations will sometimes raise more problems. Difficulties may arise for these sorts of inventions in implementing policeable systems for access controls (through which persons can be excluded from using an advance until they have paid for access) and payment remedies (through which persons gaining unpaid access can be detected and an appropriate remedy imposed that achieves the equivalent of the access payments that should have been made, plus a punitive payment where there is a need to deter knowing non-payment of access fees). Where these sorts of access controls and payment mechanisms will typically be hard or impossible to implement, the reward system of patent laws will break down and no attempt should be made to save it. Rather, the associated type of advance should be deemed non-patentable subject matter and left to other means for encouragement, if any.

Concern over ease of invention access and use and the inability of patent system processes to effectively stop such use suggests yet another reason (other than the lack of immediate utility) that pure ideas, as unapplied to practical tasks and used only in mental processes of individuals, should not fall within the patent system. The use of such ideas absent access payments would be hard, if not impossible to prevent. Both the means to stop the spread of ideas from one person to another, as well as the means for determining whether particular ideas have been used by additional parties would be highly difficult to implement. Application of the patent system to these sorts of abstract, unapplied ideas is properly withheld due in part to these difficulties. Of course, we also limit the application of the patent system regarding the spread and reuse of ideas in mental processes because the First Amendment protects the communication and intellectual understanding of ideas and the patent system must be interpreted and applied so as not to interfere with the dictates and

ends of this fundamental constitutional protection.

In sum, this analysis suggests two threshold inquiries that will bear on whether transactional costs or difficulties in administering access controls suggest that the patent system should not be applied at all. First, are physical controls (of the type normally exerted over personal property) sufficient to ensure access controls over either physical items embodying an invention or over physical equipment through which the invention is carried out such that physical controls will provide a low cost, reliable means to implement invention access controls and related payment mechanisms? Second, where an advance involves few if any physical features, will there be some other regularly successful and minimally expensive means of preventing access to the advance in order to ensure payment? If the answer to both of these questions is no, then a type of advance should probably be excluded from the patent system on transaction cost grounds.

Note that this type of analysis may call into question the propriety of patenting many business methods that do not involve devices or physical transformations of an item. The use of these methods in business contexts may be particularly hard to detect and prevent absent access payments. Hence, the analysis presented here supplies grounds for excluding some but not all business methods from patentable subject matter. Whether or not a business method should be deemed patentable subject matter should turn on whether the method is of a type (and used in a context) that use of the method is excludable—that is, that access to the business method can be controlled and constrained absent access payments.

Clarity of Use Determinations: Another type of transaction cost that may figure centrally in determining the outer boundaries of patentable subject matter relates to the distinctiveness of the boundaries between an advance and prior items or practices. This distinctiveness will influence the ease with which use or non-use of the new advance can be determined. If a new design is highly similar to an old one (or can easily be disguised as an old one), then identifying unauthorized and unpaid uses of the advance as part of an access control and payment system will be particularly difficult. In these settings, patentable subject matter might not be found because of the difficulties (and costs) of patent infringement determinations for the advances were they to be included in the patent system.

The practical implications of this type of transaction cost analysis suggest limitations on the patenting of advances where, for example, subtle types of discretion must be exercised or

determinations made as part of carrying out an advance and the necessary criteria for making these cannot be clearly defined in patent applications and claims. In these settings, patents should be withheld, in part because the patent infringement analyses that would be required to identify infringement and implement patent rewards would be time consuming and speculative.[209]

Another related ground on which a given advance might be seen as falling outside patentable subject matter is that the advance is defined only in highly abstract terms, lacking practical implementation details from which the presence and use of the same advance might be measured as part of an access control system. Here, the abstract nature of the definition of an advance suggests a use measurement problem. Where only an abstract definition is given and a set of results would be consistent with (or at least similar to the results from) both use and non-use of the abstract advance, measurements of infringement from actual activities of asserted infringers may be highly uncertain and costly. Hence, the definition of an advance only in abstract terms may be proper grounds for exclusion of the advance from patentable subject matter because of difficulties in determining whether the advance is present and in implementing a related access control and payment system through patent rights.

D. A Proposed New Standard for Recognizing Patentable Subject Matter

Based on the considerations described in this article, the following patentable subject matter standard would be a valuable advance over present uncertainty in such standards and corresponding weakness in patent-mediated invention incentives:

> Patentable subject matter is present if an advance includes 1) a distinct definition of structure, operation, and use consequences such that potential users can recognize and evaluate the methods, costs, and results of using the advance, 2) sufficient presently apparent or immediately projectable utility in its defined form to permit users to evaluate the merit of the new advance and to

209. Another reason to withhold patents in these circumstances is that the persons providing the advances have not given the public enough information to use the advances effectively until they have provided clear definitions of the analysis or decision criteria sufficient for other parties to replicate the advances and to achieve utility from using the advances.

> determine how much, if anything, they are willing to pay for access to it, 3) sufficient physical instantiation or other controllable inputs or results such that access to the advance can generally be limited to parties paying for such access, and 4) features and contexts for use such that uses of the advance can be identified and mechanisms for transferring access payments from users of the advance to the inventors of the advance, either directly or through intermediate payment steps, can be implemented without excessive costs.

The proposed standard is consistent with (although more detailed than) the standard used by the Federal Circuit in *Research Corporation*.[210] There, the court recognized patentable subject matter in an advance based on evidence of the functional ends served by the advance, the palpable components comprising the invention, the need in the relevant field for the functionality provided by the advance, and the likely commercial significance of the advance, as indicated by the existence in the marketplace of commercially successful items to which the advance was either applied or used as an improvement.[211] The proposed test focuses on similar factors, but relates the factors more clearly to invention access controls.

The proposed standard, if applied across various present and future technologies, will insure that the patent system and its restrictions and costs are limited to advances that are workable targets of the types of access controls, payments to inventors, incentivizing of invention, and prioritization of resource allocations to inventive activities that the patent system was intended to achieve.

The test proposed here is dictated by the nature of the patent system and its goals, not by the types of technologies and advances that the system is intended to foster. These standards will ensure that patent rights have their proper influence on the allocation of resources to inventive efforts in competition with other demands for the same resources. This will help to give patent rights their proper place as counterparts to personal and real property rights in the attraction and allocation of resources to the production of useful products of human ingenuity. Viewed as means to value and attract scarce resources to the production of patent-eligible advances, patent rights can be seen as one more component in the broader scheme of property rights.

---

210. *See* Research Corp. Technologies, Inc. v. Microsoft Corp., 627 F.3d 859 (Fed. Cir. 2010).

211. *Id.* at 868-69.

Patent rights can encourage invention production in the same ways that other property rights have influenced production activities concerning real and personal property production for many years.

The test for patentable subject matter proposed here represents a technology-neutral, historically unbound set of criteria for determining the scope and limits of the patent system. As such, the test is suitable to frame the promise of patent rewards for future inventors considering potentially difficult and expensive efforts to discover previously unknown types of technologies. To define the characteristics of patentable technologies in terms of technology criteria (or based on the technology-specific historical background of a technology, whether the technology emerged from industrial, liberal arts, or business sources, or whether the technology will be used in industrial, liberal arts, or business contexts) is probably fundamentally flawed. Our most useful discoveries of the greatest benefit to society in increased utility may come from and be applied in any of these areas. The patent system should back efforts to produce new, useful tools for societal use in whatever fields these tools can be produced and used.

Hence, a set of criteria like those proposed here, which are not technologically limited or historically constrained, but rather framed in terms of the operation of the patent systems and its capabilities, offer the best hope of realizing the full potential of the patent system. These standards will ensure the maximum benefit from the patent system, applying that system where the system has strengths and a positive social role, and withholding the system where it is either ineffective or too costly and cannot serve its intended social function.

VII. ADDITIONAL PATENT LAW IMPLICATIONS OF PATENT RIGHTS AS RESOURCE ALLOCATORS

While the primary focus in the doctrinal analyses in this article has been on patentable subject matter standards, the interpretation developed in this article of the patent system as an invention production incentivizing and resource attracting system has further implications for the formulation and application of other patent law standards. Indeed, given that the encouragement and regulation of invention production are among the fundamental purposes of the patent system, it is hardly surprising that these goals should be kept in mind in shaping many aspects of the patent system. The full implications of these patent system goals in shaping patent law standards are worthy of considerable additional analyses. This section

concludes with some brief thoughts on how the view of the patent system as a means for incentivizing and regulating invention production may warrant reevaluation and reformulation of patent law doctrines beyond patentable subject matter standards.

A. Utility Standards

To qualify for a patent, a particular advance must realize at least a small amount of positive utility when used.[212] Viewed as part of a system of access controls and rewards to invention producers, this utility requirement ensures that patent system controls and costs only apply to inventions with utility that users can assess and value. Absent some immediately apparent advantage in using the invention, the type of patent reward system outlined in this article will be useless and the costs of administering the patent system will be wasted.

Furthermore, by insisting that patent applicants pursue, produce, and identify in their patent applications some clear utility for their advances, we ensure that inventors are encouraged to complete their efforts to perfect their inventive projects to the point of producing some manifest positive utility, not just hoped for results. This encouragement to complete their efforts and to bring forth at least some practical examples of how their invention can be used advantageously helps to ensure that patent restrictions and costs are only applied to inventions that have some value and a corresponding likelihood of generating access payments back to invention originators. Since these payments are at the heart of the invention production incentivizing that the patent system can promote, only inventions that have some meaningful present value and corresponding likelihood of generating these sorts of payments should fall within the patent system.

B. Non-obviousness Standards

One of the key purposes of the patent system as envisioned and described here is to attract and allocate scarce resources to inventive projects. Among the key types of resources that the patent system allocates is the time of a few highly skilled inventors. Patent rights are not presented to every party who produces a new advance with practical implications, but rather only to those parties who produce advances that are both new and non-obvious relative to prior

212. 35 U.S.C. § 101 (2006).

knowledge in the same field of technology or endeavor.[213]

The interpretation of non-obviousness standards should be shaped to ensure that patent rights and the incentives that such rights create are focused on attracting the attention and talents of highly skilled (and rare) individuals to inventive efforts. This can be accomplished by tailoring non-obviousness findings to permit patenting of an advance only where the information, analytic techniques, or inventive skills needed to produce the advance were not widely held at the time the advance was made.

Where few, if any, parties other than the originator of an advance shared the same information, analytic abilities, and skills used in producing the advance, then the advance should be treated as non-obvious in the field where it was produced. This will ensure that patent incentives are focused on parties with rare knowledge and abilities (which are key inputs needed to promote certain lines of technical innovation). At the same time, by withholding patent rights in cases of widely shared information, analytic abilities, and research skills sufficient to support research projects, this type of non-obviousness standard will avoid incurring the costs and limitations of patent rights in settings where numerous parties are potentially successful innovators and the talents and knowledge of a few especially capable (and scarce) innovators need not be attracted through the special incentives of the patent system.

By limiting patents to situations where key researchers and supporting resources are scarce inputs to innovation, non-obviousness standards can tailor the impacts and costs of the patent system to the invention under-production and resource allocation problems that the patent system was designed to address.

C. Infringement Standards

1. Direct Infringement

A party is deemed to directly infringe a patent—and to be liable to the patent owner—where the party makes, uses, sells, or imports a patented invention without the consent of the patent owner.[214] Whether or not the item made, used, sold, or imported by an asserted

213. *See* 35 U.S.C. §§ 102(a), 103 (2006).

214. 35 U.S.C. § 271(a) (2006). Several additional less common types of patent infringement are recognized for conduct that aids or encourages the types of patent infringement mentioned in the text. *See* 35 U.S.C. § 271(b)-(g) (2006).

infringer (termed the "accused item" in patent analyses) was the same as the patented invention is determined by a one-to-one comparison between the elements of the patent claims and the features of the item of the asserted infringer.[215] Where there is a complete match, direct infringement is found.[216]

This sort of direct infringement analysis should be conducted as part of the policing of a system of patent-mediated invention access controls of the sort discussed in this article. What constitutes an element of a patented invention, and whether the same element is present in an accused item, should be determined and understood as persons analyzing the invention and considering the scope of related access controls and licensing would have approached the same questions. If an item asserted to be infringing would have been seen by most parties in the same field as one controlled by the patent under analysis and only properly produced or used with a license or under other consent of the patent holder, then the same item should be found to be directly infringing when made without such consent. In this way, direct infringement findings will be, in effect, findings that the infringers were abusers of the access controls and payment requirements imposed by the patent system. Direct infringers, under these standards, will be persons who have disregarded the access limitations imposed on inventions by the patent system, despite clear opportunities (from reading the relevant patents and the understanding of those patents that would have prevailed in their field) to appreciate that they were undertaking activities and gaining access to patented inventions without paying the access fees demanded by the patent system.

2. Doctrine of Equivalent Infringement

Access controls provided by the patent system can also be abused where parties gain unpaid access to the functionality of a new invention without implementing precisely the same design as the new invention. Under these circumstances, a party may infringe a patent on the new invention under the doctrine of equivalents.[217]

Viewing the patent system as a means to encourage and regulate

---

215. *See* Southwall Technologies, Inc. v. Cardinal IG Co., 54 F.3d 1570, 1575 (Fed. Cir. 1995).

216. *Id.*

217. *See* Graver Tank & Mfg. Co., Inc. v. Linde Air Products Co., 339 U.S. 605, 608-09 (1950).

invention production offers some guidance about the proper scope of the doctrine of equivalents. This doctrine, as interpreted by the Supreme Court, provides that an item which is not identical to a patented invention nonetheless infringes the patent on that invention if the item operates through the same means, serves the same function, and achieves the same result as the patented invention.[218] An item that has this trio of similarities to a patented invention—that is, the same means, function, and result—is viewed as the equivalent of the patented invention.[219] The Supreme Court has further indicated that the equivalency of an item to a patented invention should be assessed element by element, meaning that an item must have an identical or equivalent element to every element of a patented invention before infringement under the doctrine of equivalents can be found.[220]

The doctrine of equivalents can serve a valuable function in a patent-implemented invention access control system by preventing parties from gaining unpaid access to the functionality of patented advances by making cosmetic or functionally unimportant changes in those advances. By viewing the purpose of the doctrine of equivalents as being to prevent unpaid access to the functional advantages of a patented invention, the nature of equivalent advances that should be deemed infringing under the doctrine of equivalents becomes clearer. Any advance that achieves functional results that are similar to those of a patented advance and that incorporates components and operating sequences that are similar to those of a patented advance should be deemed infringing.[221] Persons using these types of similar advances

218. *Id.*

219. *Id.*

220. Warner-Jenkinson Co., Inc. v. Hilton Davis Chem. Co., 520 U.S. 17, 40 (1997).

221. Under this approach, whether or not a similar advance was directly derived from a patented advance by modifying the latter will not matter for purposes of assessing infringement under the doctrine of equivalents. The salient issues are similarity of the structure and operation of the two items and whether these similarities account for the functionality of the item that is similar to the patented advance. If so, the value of the patented advance is being captured in the similar item and a user of the latter should pay the same charges for access to this functionality as a user of the patented advance.

While the derivation of a modified advance from a patented item is not strictly indicative of infringement under the doctrine of equivalents, the process by which derivation occurred may have some bearing on whether infringement should be found under the doctrine. For example, it will often be relevant to consider whether, in making changes to the patented item, designers sought to incorporate meaningful functional differences in the modified item (perhaps because the modifications were part of a designing around process aimed at avoiding patent infringement). Where significant functional differences were targeted and pursued in good faith,

are gaining access to the practical advantages of the patented invention and should only be able to do so by paying for access.

A system that lets uncompensated use of these sorts of equivalent items slip by without infringement findings and remedies will not reward inventors for the full range of increased utility that their advances have brought to society. This will undervalue their inventive efforts and tend to encourage future inventors in similar circumstances to make choices to pursue other types of activities, resulting in the under-production of valuable inventions relative to public demand. Toleration of this type of uncompensated use will also encourage parties to create more and more functionally equivalent items that work like patented advances but that do not incur patent-related charges when used. These sorts of imitators (who are effectively free riders on the inventive success of the producers of patented advances) will tend to prevail in their own fields of competition over those parties who pay for access to patented advances. Imitators will tend to prevail in these settings since the imitators of patented advances will not suffer the costs of patent-related access payments that their competitors will bear. Over time, this sort of industry change under competitive pressures will advantage clever users of equivalent items that gain the advantages of patented items without incurring the costs of producing the patented

---

it makes it more probable that meaningfully different functional attributes were attained in the modified device. This, in turn, would provide some circumstantial evidence that the modified item is not similar to the patented one in structural, operational, and functional details and, accordingly, that the second advance should not be deemed an infringing item under the doctrine of equivalents.

Under this approach, persons seeking to modify patented designs to create new functionality are treated like designers generally, meaning that they are able to realize commercial gains from their designs free from patent enforcement threats from holders of patents on the unmodified items. Furthermore, if their new designs have features meeting patent standards, the designers of the modified items can seek patents of their own for the new features. Interpreting the doctrine of equivalents as allowing this sort of modification of patented items to achieve new functionality (where that new functionality does not depend on the protected functional mechanisms of the patented item) furthers the technology diversification goals of the patent system by bringing more technological alternatives into public access and use.

In contrast, attempts to make trivial (that is, functionally unimportant) changes in a patented advance in order to avoid findings of patent infringement are particularly injurious to the ends of the patent system. These efforts to "camouflage" infringement will just make instances of infringement harder to find and compensate, thereby undercutting the type of reward function that the patent system was intended to implement. Intentional efforts to conceal infringement in this way deserve strong condemnation and deterrence. One way to achieve this is to treat efforts aimed just at creating superficial differences between patented and similar items as instances of willful infringement, with the potential for punitive damages as a remedy.

items.

To avoid this result—and the resulting gutting of patent-related payments and rewards as innovation incentives—findings of infringement under the doctrine of equivalents should be robustly made and enforced to police the boundaries of use of the functional results of patented advances. Such functionally-driven findings of infringement are needed to ensure that the parties who gain access to those results—as evidenced by use of items with features that produce the same functional results as those of a patented advance with substantially similar components and operational characteristics—are forced to pay for access to this functionality in the same manner as users of the patented advance itself.

D. Damages

By treating the patent system—including the relief afforded in patent suits when the normal access controls imposed by the patent system are not observed and a party makes unauthorized use of a patented invention—as a scheme for matching user demand with invention incentives, the proper scope of patent damages can be clarified. Patent damages are a substitute for the amount that a user would be willing to pay for access to a patented advance over the amount that they would pay for the next best substitute for performing the same task. Where there are active markets for the patented item and the next best substitute, the difference between these market prices should generally determine the per item patent damages that can be recovered.[222] Where there is no such market, an estimate of the utility gains and value to the infringer of access to one unit of the patented item (less the cost of producing the unit and providing this access) should define the per unit damages that should be recoverable for infringement. This will equal the amount that the infringer would reasonably have been willing to pay for the type of access to the patented advance that the party utilized. The object of damage awards of this sort is to approximate the amounts that would have flowed to the patent holder were the patent system to have successfully controlled access to the patented invention and the infringer to have gained access to the invention via legitimate (that is, non-infringing)

222. Where the production costs of the patented item and its second best substitute are about equal, the patent damages (equal to lost profits) per item will be equal to the difference in the prices of the two items. Where one item is more or less costly to produce than the other, the per unit lost profit figure needs to be adjusted to reflect this difference.

channels and business processes.

Additional punitive damages may be needed to discourage choices by potential infringers not to seek patented advances from legitimate sources—that is, to discourage choices to engage in infringement knowing that access to an advance was constrained by patent rights. Without punitive damages for choices to forgo access controls and patent-influenced payments to patent holders, potential infringers, confronted with knowledge of patent controls and invention use limitations, will have little reason to respect and adhere to the access controls of the patent system. They will instead be encouraged to choose to infringe by accessing patented advances without payments of patent-influenced prices with the hope that they will not be detected or, in the alternative, with the hope that, upon detection, the worst that will happen is that they will have to pay patent damages roughly equal to the amounts that they would have had to pay for access through legitimate channels of access sanctioned by the patent owner.

To avoid this decision calculus favoring decisions to infringe, knowing infringers must face heightened damages that, even in light of the chance that their infringement will not be detected at all, create a perception of net projected costs in choosing to infringe and a net projected advantage in seeking access to a patented advance through means consented to by the patent holder.

E. Injunctions

Injunction standards crafted in accordance with the view of patent law described here would focus on the need to prevent uncompensated access to patented inventions. Where a patent holder's past practices have established a generally available access payment scheme for an invention—perhaps because the patent holder has a well-established and non-exclusive licensing scheme specifying a price for access to a patented technology—injunction relief for infringement may not be needed at all, so long as an infringer pays damages equal to the access fee (that is, licensing fee) the infringer would have paid under the patent holder's access pricing scheme. In this context, damages are a sufficient substitute for privately imposed access charges and further injunctive relief is not needed.[223]

223. This criterion for determining whether an injunction is appropriate in a patent case is equivalent to a key component of the standard adopted by the Supreme Court in the *eBay* case. *See* eBay Inc. v. MercExchange, LLC, 547 U.S. 388, 391-94 (2006).

Injunctions to counteract patent infringement will be needed primarily where damages are not easily available substitutes for privately-imposed access charges. Injunctions will help force potential users of patented advances to deal with access controls and to bargain with patent holders as the parties controlling access if the potential users are to have further access to the patented inventions at all. In this respect, injunctions serve as a means to police and preserve the access control system that sets up mechanisms of payments and rewards back to inventors, with the amount of the actual payments and rewards determined through bargaining between patent holders and infringers after injunctions have issued.

Where there is no established price for additional access (perhaps because the patent holder intends to be the sole producer of a technology or intends to grant an exclusive license to only one party), the reward function of the patent system will generally be best served by protecting the patent holder's control over access to a patented invention through the issuance of an injunction barring unauthorized use of the invention. This will ensure that any subsequent access of the infringer will be at a price that the patent holder has bargained for and consented to. The patent holder will have incentives to determine the incremental utility seen in the invention by the infringer and to set the price for access at but not above this point (since a higher price will not be paid by a rational licensee or user of the patented invention). A patent holder will tend to cap their access price at this point since a price—that is, licensing royalty—too high will just cause the potential user to turn away from a license and to use the second best item or practice to achieve the same functionality as the patented item would have provided.

VIII. CONCLUSION

Patent rights exist to serve societal ends. They encourage talented inventors and the resource managers who back these inventors to devote scarce time and resources to the search for new man-made tools for societal use. In encouraging more—but not too much more—attention to the development of these tools, the patent system serves a function like no other legal regime in encouraging and regulating socially desirable inventive efforts. Patent rights encourage and regulate the production of patent-eligible advances with characteristics that match public demands for new advances.

A lot is at stake in determining the proper scope and means for operation of the patent system. The parties and resources influenced

by this system are among our most talented and most valuable. The inventions that patents promote—ranging from life saving drugs to society transforming communications devices—work tremendous changes in our daily activities and offer the best chance for new efficiencies and effectiveness in our economies, our societies, and our lives.

We need a strong patent system because we need the useful tools it generates and the better world they enable. We need a patent system when—but only when—it can encourage the creation of socially valuable tools and can regulate the allocation of scarce resources to the important tasks of creating these tools. While it comes at a price, our patent system aims at promoting the production of our best innovations, propagating the use of these advances, and topping these advances with further innovation. The strength with which we back these innovative processes has much to do with the pace of progress in our society and with the strength of the United States economy as one of the primary sources of valuable innovation in the world.

Ultimately, the patent system is a case in point of "you get what you pay for." We need a patent system because we need what it produces. Valuing these ends, we should be willing, without carelessness and waste, to enforce patent rights strongly and to pay related fees for access to functionally valuable advances produced by talented innovators and risk-taking resource investors. This is the patent bargain—a good exchange enabled by patent rights and carried out through a costly, but worthwhile, patent-mediated system for inducing and regulating invention production.