THE JOHN MARSHALL
REVIEW OF INTELLECTUAL PROPERTY LAW

# IDEAS INTO PRACTICE: HOW WELL DOES U.S. PATENT LAW IMPLEMENT MODERN INNOVATION THEORY?

STEPHEN M. MAURER

ABSTRACT

The U.S. Supreme Court's decision in *Graham v. John Deere* (1966) placed neoclassical economic insights at the heart of modern patent law. But economic theory has moved on. Since the 1990s, legal scholars have repeatedly mined the discipline to propose *ad hoc* rules for individual industries, such as biotech and software. So far, however, they have almost always ignored the literature's broader lessons for doctrine. This article asks how well today's patent doctrine follows and occasionally departs from modern economic principles. The analysis begins by reviewing what neoclassical economists have learned about innovation since the 1970s. Legal scholars usually divide this literature into a half-dozen competing and distinct "theories." Naively, this seems to suggest that any patent doctrines based on these theories must be similarly fragmented. This article offers a way out: far from being in conflict, the putatively separate "theories" share so many common assumptions and mathematical methods that they can usefully be analyzed as special cases of a single underlying theory. Furthermore, much of this theory is known. In particular, it predicts that any economically efficient patent system must accomplish three tasks: (1) limiting reward to non-obvious inventions; (2) choosing patent breadth to balance the benefits of innovation against the costs of monopoly; and (3) prescribing rules for allocating patent rewards where multiple inventors contribute to a shared technology. Remarkably, patent doctrine uses *Graham's* PHOSITA concept to address all three principles. This means that doctrinal solutions for one principle can have unintended impacts on the others. This article shows that any doctrinal architecture built on *Graham*'s PHOSITA test automatically allocates reward among successive inventors. Though reasonable, these default outcomes fall short of the economic ideal. This article analyzes how changes in the Utility, Blocking Patents, Reverse Doctrine of Equivalents, and the Written Description doctrines can mitigate this problem. However, other gaps are inherent and cannot be eliminated without abandoning *Graham* itself. This radically revised architecture would probably cause more problems than it solves.

Copyright © 2013 The John Marshall Law School

*Cite as* Stephen M. Maurer, *Idea into Practice: How Well Does U.S. Patent Law Implement Modern Innovation Theory?*, 12 J. MARSHALL REV. INTELL. PROP. L. 644 (2013).

# IDEA INTO PRACTICE: HOW WELL DOES U.S. PATENT LAW IMPLEMENT MODERN INNOVATION THEORY?

STEPHEN M. MAURER

# IDEAS INTO PRACTICE: HOW WELL DOES U.S. PATENT LAW IMPLEMENT MODERN INNOVATION THEORY?

STEPHEN M. MAURER*

## INTRODUCTION

Economists' understanding of patent incentives has grown immensely in the four decades since Professor Edmund Kitch wrote his seminal article.[1] This is clearly a good thing. At the same time, legal scholars have had a hard time mapping insights about incentives onto patent doctrine. Since the early 1990s, the usual response has been to divide the economic literature into a half-dozen or so "theories"—the precise number differs from author to author—each of which is said to be valid for some industries but not others.[2] But this means asking judges to start every case by selecting whichever theory or combination of theories best fits the facts at hand. This "many theories" approach vests broad discretion in judges and fragments patent doctrine into a potentially endless series of mini-statutes. While economists have argued that this kind of approach might be efficient,[3] it is hard to believe that Congress intended such an outcome.

This article begins by arguing that the "many theories" premise is false. Far from being in conflict, most of the economics literature shares common neoclassical[4] assumptions and methods. For this reason, the putatively separate theories *cannot* disagree in any fundamental way. But in that case, we can hope for a general economic theory leading to universal principles and global legal rules.

This article asks how well current patent law implements the lessons of neoclassical theory. In particular, we argue that any sufficiently unified neoclassical theory of innovation incentives must endorse three specific policy goals: (1) limiting reward for inventions that would be discovered in any case; (2) setting rewards that

* Berkeley Law School and Goldman School of Public Policy, University of California at Berkeley. Address: Goldman School of Public Policy, 2607 Hearst Ave., Berkeley, CA 94720-9657; smaurer@berkeley.edu.

[1] Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 265 (1977). As with most milestones, assertions that modern innovation economics dates from the 1970s are inherently arbitrary. One might equally argue that the main economic arguments for and against patent law were qualitatively understood in the 1860s. Mark D. Janis, *Patent Abolitionism*, 899 BERKELEY TECH. L.J. 899, 907–08 (2002) (documenting Victorian dissatisfaction with the patent system that parallel modern concerns); Fritz Machlup & Edith Penrose, *The Patent Controversy in the Nineteenth Century*, 10 J. ECON. HIST. 1, 1 (1950).

[2] See *infra* note 16 for an extended list.

[3] *See* SUZANNE SCOTCHMER, INNOVATION AND INCENTIVES 117–18 (2004) [hereinafter SCOTCHMER, INNOVATION AND INCENTIVES] (arguing that different industries have dramatically different R&D costs and/or ability to extract revenue from consumers, which make it "almost inevitable" that a one-size-fits-all patent reward "underreward[s]" some inventors and "overreward[s]" others).

[4] The term "neoclassical economics" conventionally refers to an approach to economics that relates supply and demand in a marketplace to an individual or firm's hypothesized maximization of profits and utility. *See* Antonietta Campus, "Marginal Economics," The New Palgrave: A Dictionary of Economics, v. 3, p. 323 (1987).

balance the benefits of innovation against the burden of monopoly prices and wasteful duplication of effort; and (3) allocating rewards to cover each innovator's costs when successive advances contribute to a common technology. Remarkably, we will see that current patent doctrine uses a single concept—*Graham's* "Person Having Ordinary Skill in the Art" ("PHOSITA")[5]—to address all three goals. Inevitably, this creates problems. We show that any version of patent law that uses a *Graham* non-obviousness threshold to address our first goal *automatically* sets "default choice" rules for reaching our third goal. Though reasonable, these rules fall well short of the incentives endorsed by neoclassical theory. In what follows, we show that the fit can be improved by adding defenses (e.g. the Reverse Doctrine of Equivalents) that require jurors to perform additional fact inquiries. That said, no amount of tinkering can eliminate the discrepancy entirely. This implies that *any* doctrine built on *Graham* non-obviousness is inherently flawed and that truly fundamental reforms would—assuming that they are possible at all—require fundamental changes to current law.

We proceed as follows. Part II discusses how innovation economics research has entered the legal literature. It argues that the "many theories" viewpoint is overstated and that most of the literature can be readily grouped within a coherent "neoclassical core." Part III uses this neoclassical core to identify three policy goals that any economically efficient patent statute must try to accomplish. Part IV reviews the Utility Doctrine and compares it against modern economic insights. Part V explores a puzzle: R&D costs are ubiquitous in economic theory but seldom if ever mentioned in doctrine. We argue that patent law overcomes this difficulty by using the PHOSITA concept to measure inventive cleverness as a proxy for effort and cost. We also discuss how litigation uncertainty improves the match between doctrine and economically efficient incentives. Part VI analyzes the allocation rules generated by *Graham*'s PHOSITA test in a simple Literal Infringement scenario. Strikingly, we find that the test generates reward allocations that are facially reasonable but highly imperfect. Part VII extends the analysis to the Doctrine of Equivalents. Part VIII examines how defenses based on the Reverse Doctrine of Equivalents, Pioneer Patents, and Written Description defenses can improve the alignment between doctrine and economically efficient incentives. Part IX assesses the gaps between patent doctrine and our three policy goals and asks whether a fundamentally different architecture could do better. Part X presents a brief conclusion.

## I. THE SEARCH FOR A GENERAL THEORY

No one doubts that a general theory of patent law is desirable.[6] How could they? We live in an age where general theories—sometimes even "Theories of Everything"[7]

---

[5] Graham v. John Deere Co., 383 U.S. 1, 19 (1966).

[6] John F. Duffy, *Patent System Reform: Harmony and Diversity in Global Patent Law*, 17 BERKELEY TECH. L. J. 685, 693 (2002) ("The most compelling justification for harmonization in patent law mirrors the justification for creating a patent system in the first place, for both are efforts to account for the positive externalities associated with the creation of technical information.") (emphasis omitted).

[7] *See, e.g.*, STEPHEN W. HAWKING, THE THEORY OF EVERYTHING (2005).

—are praised to the skies.[8] The only real question is whether we know how to build such a thing.

This Section begins by asking what we mean by a general theory. For the sake of definiteness, we illustrate our argument with paradigmatic examples from physics. We then review various "theories" that legal scholars have identified in the economics literature. Significantly, most of this work shares the same neoclassical assumptions (rational actors, profit maximization) and mathematical methods (optimization calculus). Within this "neoclassical core," any disagreements are superficial and can usually be traced to special industry-specific factual assumptions. From this perspective, the putatively separate "theories" are more accurately seen as special cases of a common underlying theory. This should encourage patent law reformers to search for broad principles and rules of general application.

### *A. The Gold Standard: Classical Physics*

We begin by recalling what a really powerful theory looks like. Here the paradigmatic example is Newtonian physics. Consider the path of a moving object. Plainly, one can imagine many different versions of this problem. For example we can analyze the object's path when it is allowed to move freely in three dimensions like a planet orbiting the sun (Problem A). Alternatively, we can analyze the much simpler problem of how fast it moves when forced to follow a track or inclined plane. (Problem B). On the one hand, both problems are clearly related. But on the other, the solutions look very different on the page.[9] The question is, should we emphasize the commonalities or points of difference? One could say that Problem A is a "different theory" than Problem B, but that distinction is misleading. In practice, physicists almost always say that B is a "special case" of the "general theory" A.[10]

The difference is more than semantic. Problem B, being less general, is almost always easier to solve than Problem A. Indeed, this is usually why Problem B was chosen in the first place. On the other hand, each special case identifies features that must also appear in the general theory.[11] For a physicist, finding this general solution could be an end in itself. General solutions also have practical power. For example, engineers might use a general solution to plot lowest-energy trajectories for spacecraft traveling between any two points in the solar system. In both cases, there is an obvious incentive to unify the special cases as much as possible.

---

[8] For a recent example, see *In Praise of Particle Physics: Higgs Ahoy!*, THE ECONOMIST (Dec. 17, 2011), http://www.economist.com/node/21541825 (describing a triumphant account of how physicists found evidence for the Higgs boson, a long-hypothesized particle needed to reconcile the Standard Model of particle physics with observed subatomic particles and masses).

[9] *Compare* 1 RICHARD P. FEYNMAN, ROBERT B. LEIGHTON & MATTHEW SANDS, *Motion*, *in* THE FEYNMAN LECTURES ON PHYSICS 8–10 (Addison-Wesley 1964) (analyzing a ball rolling down an inclined plane), *with* 1 FEYNMAN ET AL., *Newton's Laws of Dynamics*, *in* THE FEYNMAN LECTURES ON PHYSICS 9–7 (calculating planetary orbits).

[10] *See, e.g.*, RICHARD WOLFSON, SIMPLY EINSTEIN 174 (2003) ("Special [in Einstein's special theory of relativity] means specialized, limited, restricted—to the special case of uniform motion.").

[11] *Id.*

Not surprisingly, this can be hard. One obvious problem is that physicists may not know the general theory.[12] Fortunately, it is often enough to suspect that a general theory exists and guess at principles.[13] On the other hand, some general theories are known but lack general solutions.[14] Even if a general theory exists, engineers may still be restricted to special (and solvable) cases.[15]

The corresponding agenda for patent law is clear. Economists, like physicists, have been groping toward a general theory of innovation. And lawyers, like engineers, would like to exploit this knowledge to design doctrines that accelerate innovation at the lowest possible cost. As in our physics example, we cannot know whether this agenda will be successful. At the same time, we can be sure of one thing—we will never know unless we try. This article begins the experiment.

### *B. How Lawyers See the Economics Literature*

It has become conventional for legal scholars to stylize the economics literature in terms of competing theories.[16] The practice seems to have originated in a 1990

---

[12] This was the case for electromagnetism between 1831 and 1865, and relativity between 1905 and 1917. See, e.g., ROY PORTER, THE BIOGRAPHICAL DICTIONARY OF SCIENTISTS (Oxford Univ. Press, 2d ed. 1994) (entries for Faraday, Maxwell, and Einstein).

[13] For example, Maxwell's final theory of electromagnetism includes one law previously known to Faraday and a second law that Ampere discovered and Maxwell corrected. Graham Hall, *Maxwell's Electromagnetic Theory and Special Relativity*, 366 PHIL. TRANSACTIONS ROYAL SOC'Y 1849, 1855–56 (2008), *available at* http://rsta.royalsocietypublishing.org/content/366/1871/1849.full.pdf+html.

[14] The legend of Mohammed's Coffin provides an entertaining example. According to medieval folklore, the Prophet Mohammed was buried in such a way that magnets kept his coffin perpetually suspended between heaven and earth. Samuel Earnshaw proved in 1842 that no general solution to the Mohammed's coffin problem exists. There are, however, solutions for special cases including the one-dimensional problem and time-varying electric fields. Philip Gibbs & Andre Geim, *Levitation Possible*, RADBOUD UNIV. NIJMEGEN (Mar. 18, 1997), http://www.ru.nl/hfml/about-hfml/levitation/diamagnetic/levitation-possible/. A lawyer would call these solutions "*ad hoc* rules."

[15] *See, e.g.*, Günter Rothe, *Two Solvable Cases of the Traveling Salesman Problem*, TECHNISCHEN UNIVERSITÄT GRAZ 1 (1988), http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.8.2643. "The traveling salesman problem is known to be NP-hard . . . which implies that no algorithm is known currently which finds an optimal tour in polynomial time." *Id.* One approach around this challenging aspect is by "identif[ying] restricted classes of the problem for which efficient solutions are possible." *Id.*

[16] *See* Amelia S. Rinehart, *Patents As Escalators*, 14 VAND. J. ENT. & TECH. L. 81, 95–96, 108–10 (2011) (discussing prospect and option pricing theories); Robert P. Greenspoon & Catherine M. Cottle, *Don't Assume a Can Opener: Confronting Patent Economic Theories With Licensing and Enforcement Reality*, 12 COLUM. SCI. & TECH. L. REV. 194, 197 (2011) (discussing reward, prospect, and commercialization theories); Roberto Mazzoleni & Richard R. Nelson, *Economic Theories About the Benefits and Costs of Patents*, 32 J. ECON. ISSUES 1031, 1038–44 (1998) (discussing reward, disclosure, commercialization, and broad prospect theories); Ted Sichelman, *Commercializing Patents*, 62 STAN. L. REV. 341, 341, 377–78 (2010) (discussing reward, prospect, commercialization, signaling, and transaction cost theories); Andrew Blair Stanek, *Increased Market Power as a New Secondary Consideration In Patent Law*, 58 AM. U. L. REV. 707, 731 n.136 (2009) (identifying five economic theories posited over time for the patent system); Wolrad Prinz zu Waldeck und Pyrmont, *Research Tool Patents After* Integra v. Merck--*Have They Reached a Safe Harbor?*, 14 MICH. TELECOMM. & TECH. L. REV. 367, 371-73, 386 (2008) (discussing inducement-to-invent, incentive to

article by Professors Robert Merges and Richard Nelson, who criticized earlier scholars for "assum[ing] that invention is the same in all technologies."[17] They divided the economics literature into three categories.[18] First, there were traditional models that examined incentives for a single inventor working in isolation.[19] These almost always focused on "a simple tradeoff" between "incentives to the inventor and underuse of the invention due to patent monopolies."[20] Second, there were more recent theories in which innovation proceeded on the basis of "many actors" operating

---

invest, reward, disclosure, and anticommons theories); Carol M. Nielsen & Michael R. Samdzija, *Compulsory Patent Licensing: Is It a Viable Solution in the United States?*, 13 MICH. TELECOMM. & TECH. L. REV. 509, 517–19 (2007) (discussing prospect, cumulative innovation, anticommons, and patent thicket, theory); David B. Conrad, *Mining the Patent Thicket: The Supreme Court's Rejection of the Automatic Injunction Rule in* eBay v. MercExchange, 26 REV. LITIG. 119, 132–137 (2007) (discussing reward, disclosure, commercialization, and prospect theories); F. Scott Kieff, *Coordination, Property and Intellectual Property: An Unconventional Approach to Anticompetitive Effects and Downstream Access*, 56 EMORY L.J. 327, 398–435 (2006) (discussing reward, rent dissipation, anticommons, prospect, and commercialization theories); Richard S. Gruner, *Corporate Patents: Optimizing Organizational Responses To Innovation Opportunities and Invention Discoveries*, 10 MARQ. INTELL. PROP. L. REV. 1, 39, 47, 50, 52 (2006) (discussing reward, prospect, disclosure, and rent dissipation theories); Peter Fox, *It's Not Over the for the Product of Nature Doctrine Until the Synthetic Super-Heavy Element ('SHE') Sings*, 79 TEMP. L. REV. 1005, 1011 n.47 (2006) (discussing reward, prospect, race-to-invent, and rent dissipation theories); Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L. REV. 123, 132–33, 147, 149 (2006) (discussing reward, portfolio, prospect, incentive, and signaling theories); Christopher A. Cotropia, *Patent Claim Interpretation Methodologies and Their Claim Scope Paradigms*, 47 WM. & MARY L. REV. 49, 117–23, 120 n.374 (2005) (discussing competitive innovation, prospect theory, and cumulative innovation); Shubha Ghosh, *Patents and the Regulatory State: Rethinking the Patent Bargain Metaphor After Eldred*, 19 BERKELEY TECH. L.J. 1315, 1353–69 (2004) (discussing prospecting, information costs, business asset, and industry regulation theories of patent law); Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1600–14 (2003) [hereinafter Burk & Lemley, *Policy Levers*] (discussing prospect, competitive innovation, cumulative innovation, anticommons, and patent thicket theories); Julian D. Forman, *A Timing Perspective on the Utility Requirement in Biotechnology Patent Applications*, 12 ALB. L.J. SCI. & TECH. 647, 665–80 (2002) (discussing "prospect," "reward," and Anticommons theory); Mark. D. Janis, *Second Tier Patent Protection*, 40 HARV. INT'L L.J. 151, 209–11, 218 (1999) (discussing prospect, reward, and anticommons theories); Phanesh Koneru, *To Promote the Progress of Useful Articles? An Analysis of the Current Utility Standards of Pharmaceutical Products and Biotechnology Research Tools*, 38 IDEA 625, 629–53 (1998) (discussing reward, patent-induced, prospect, race-to-invent, and rent dissipation theories); Craig A. Nard, *Deference, Defiance, and the Useful Arts*, 56 OHIO ST. L.J. 1415, 1419, 1419 n.13 (1995) (discussing reward, disclosure, prospect, and rent dissipation theories).

Probably the most systematic "many theories" analysis is found in A. Samuel Oddi, *Un-Unified Economic Theories of Patents—the Not-Quite-Holy Grail*, 71 NOTRE DAME L. REV. 267, 268 (1996) (asking if any one of five economic theories of patent incentives supplies a "unifying economic theory of patents"). Unlike the present article, however, Professor Oddi is not looking for a unified theory of innovation that might (or might not) lead to ideal patent rules. Rather, he asks whether existing law follows from one of his economic theories. *Id.* at 270–71. Indeed, he even rejects theories on the ground that they are not "consistent with the [current] patent statute, its substance and procedure." *Id.* at 271. Not surprisingly, the quest fails. Indeed, Oddi concludes that some decisions are "better explained by doctrine than by any economic theory." *Id.* at 290.

[17] Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 843 (1990).

[18] *Id.* at 868–78.

[19] *Id.* at 868.

[20] *Id.*

over "many rounds."[21] Following Kitch, these theories often produced arguments in favor of broad patents that deliberately promoted "single-firm domination of a technological prospect."[22] Third, and finally, Merges and Nelson argued that the usual neoclassical assumption that firms are rational and maximize profits could fail.[23] In this case, Kitch-type arguments disappear so that narrow patents are once again preferred.[24]

Scholars have updated Merges and Nelson's categories over the years. In their influential 2003 article, Professors Dan Burk and Mark Lemley subdivide Merges and Nelson's second (multi-inventor) category into "Prospect Theory," "Competitive" innovation, and "Cumulative Innovation."[25] Similarly, they divide the third category into "Anticommons Theory" and "Patent Thickets."[26] Despite small variations from author to author, this expanded list continues to dominate contemporary academic discussion.[27]

Inevitably, this "many theories" viewpoint implies a choice. In the easy cases, one published theory will indeed fit all of the available facts. This, however, can only happen by accident. As Merges and Nelson acknowledge, individual industries will more commonly display "some attributes" of the published theory while "not . . . shar[ing] all the features."[28] They may also have "particular characteristics of [their] own" that do not fit any theory at all.[29] Logically, judges can address such cases in just two ways: pick an existing category knowing that the fit is imperfect, or else let the number of recognized "theories" proliferate indefinitely. While Merges and Nelson do not address the issue in detail, they are plainly open to the latter choice.[30] Despite arguing that four categories cover much of the U.S. economy,[31] they admit that still more categories could be needed in the future.[32] Worse, they remark

---

[21] *Id.* at 869.

[22] *Id.* at 870–72.

[23] *Id.* at 872.

[24] *Id.* at 873–75. Merges and Nelson argue, in particular, that real firms may not be satisfied with sub-optimal performance, have organizational biases that limit their rationality, and/or be unable to agree on probable outcomes even in principle. *Id.* at 872–73.

[25] Burk & Lemley, *Policy Levers*, *supra* note 16, at 1600–10.

[26] *Id.* at 1624–30.

[27] *See supra* note 16.

[28] Merges & Nelson, *supra* note 17, at 882–83; *cf.* Dan L. Burk & Mark A. Lemley, *Biotechnology's Uncertainty Principle*, 54 CASE W. RES. L. REV. 691, 738 (2004) [hereinafter Burk & Lemley, *Biotechnology's Uncertainty Principle*] (arguing that the biotechnology industry is "properly described in part by the anti-commons theory . . . and in part by prospect theory").

[29] Merges & Nelson, *supra* note 17, at 882.

[30] *Id.* at 842–43.

[31] *Id.* at 880–84. The four categories are "discrete invention[s]" (for example, the King Gillette safety razor and the ball point pen), "cumulative technologies" (such as electric lighting, automobiles, aircraft, radio, semiconductors, and computers), "chemical technologies" (bulk chemicals and pharmaceuticals), and "science-based industries" (such as biotechnology). *Id.*

[32] *Id.* at 880 (arguing that the U.S. economy is adequately described by "at least" four categories).

that industries may host multiple theories so that the correct choice can differ from firm-to-firm and even product-to-product.[33]

This modern trend marks a fundamental departure from patent law's historic emphasis on rules of general application.[34] Despite this, modern scholars continue to treat policy questions as an exercise in picking and choosing between economic theories.[35] Judges have been similarly fond of inventing industry-specific rules.[36]

### *C. Searching for Unity: The Neoclassical Core*

Once legal scholars decided on a "many theories" approach, the rise of industry- and sometimes even product-specific rules became inevitable.[37] But was the decision necessary? In what follows, we review the economic literature's main strands and find evidence for their underlying unity.

We begin by examining the literature on individual inventors. This genre is variously called **reward**,[38] **incentive to invent**,[39] and **inventive**[40] theory.

---

[33] *See id.* at 882–83 (arguing that chemical products with a single application are "discrete entit[ies]," but chemical products with multiple applications are not). Additionally, processes for making chemicals are "cumulative" technologies. *Id.* at 883.

[34] Samantha A. Jameson, *The Problems of the Utility Analysis in* Fisher *and Its Associated Policy Implications and Flaws*, 56 DUKE L.J. 311, 339 (2006) (noting that the patent system "[t]heoretically . . . 'provides technology-neutral protection to all kinds of technologies'") (citing Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific*?, 17 BERKELEY TECH. L.J. 1155, 1156 (2002) [hereinafter Burk & Lemley, *Patent Law*]); Burk & Lemley, *Patent Law*, at 1160 (noting that patent law traditionally stressed "unified rules" for a "homogeneous world").

[35] *See* Alan Devlin, *Patent Law's Parsimony Principle*, 25 BERKELEY TECH. L.J. 1693, 1721 (2010) (arguing that patent rules should be "discriminately" adjusted to correct "over-rewarding" of inventors in the business methods, computer software, and medical diagnostics industries); Michael A. Carrier, *Cabining Intellectual Property Through a Property Paradigm*, 54 DUKE L.J. 1, 38–42 (2004) (identifying eleven industries where "patents are less critical for innovation than the traditional theory would posit," and implying the need for narrower patents and/or a "greater role for defenses to claims of patent infringement").

[36] *See generally* Burk & Lemley, *Policy Levers*, *supra* note 16, at 1593 (claiming that the Federal Circuit goes to "inordinate lengths to find biotechnological inventions nonobvious" while simultaneously "impos[ing] stringent enablement and written description requirements . . . "). The situation in software is exactly reversed. *Id.*

[37] *See generally id.* at 1577 (suggesting patent laws, as "actually applied to different industries[,] increasingly diverge"); Merges & Nelson, *supra* note 17, at 843 (suggesting several models of technical advance in industry and that the scope of patent depends on the nature of technology in an industry).

[38] For recent examples, see Lisa L. Ouellette, *Do Patents Disclose Useful Information?* 25 HARV. J. LAW & TECH. 531, 541 (2012); Greenspoon & Cottle, *supra* note 16, at 196; Rinehart, *supra* note 16, at 87, 89 n.42.

[39] Kieff, *supra* note 16, at 399–400; Yusing Ko, *An Economic Analysis of Biotechnology Patent Protection*, 102 YALE L.J. 777, 791–92 (1992). The usage has the advantage of emphasizing economic incentives over natural law arguments based on justice to the inventor. Oddi, *supra* note 16, at 273–74.

[40] Laurie L. Hill, *The Race to Patent the Genome: Free Riders, Hold Ups, and the Future of Medical Breakthroughs*, 11 TEX. INTELL. PROP. L.J. 221, 238 (2003). Scholars sometimes note a fourth **incentive to disclose** theory which holds that patents are needed to persuade inventors to publish information that would otherwise be kept secret. *See id.*; Ko, *supra* note 39, at 795–98; Liza

Regardless, the central focus is the same: adjusting the size of the patent reward for maximum benefit to society.[41] On the one hand, the reward must usually be high enough to cover inventors' R&D costs[42] and may require a bonus to accelerate R&D or overcome the probability of failure or unexpected obstacles.[43] On the other, new products are pointless unless consumers can afford them. Making the patent monopoly too broad interferes with this goal.[44] While these points were already well-known to Thomas Jefferson,[45] modern neoclassical theory makes them precise.[46] Most legal scholars have absorbed this more careful and sophisticated language.[47]

So far, we have concentrated on theories that address inventors in isolation. However, real inventors often coexist and compete with one another. Professor Kitch's **prospect theory**[48] was the earliest installment in this multi-inventor literature.[49] It argued that multiple, uncoordinated inventors were bound to duplicate each other's work and that this necessarily causes duplication and waste. Broad patents overcome this problem by letting a single owner coordinate development.[50] Variants of Kitch's argument add that patent owners can also use their power to suppress rent-seeking[51] and promote information-sharing.[52] Many scholars argue that these ideas are most relevant for R&D programs that "commercialize" primitive prototypes into marketable products.[53] For this reason, Kitch's basic idea is often referred to as "development,"[54] "incentive-to-innovate,"[55] or "innovation theory."[56]

---

Vertinsky, *An Organizational Approach to the Design of Patent Law*, 13 MINN. J.L. SCI. & TECH. 211, 221 (2012).

[41] Kieff, *supra* note 16, at 435.

[42] Burk & Lemley, *Policy Levers*, *supra* note 16, at 1585–86. Unless otherwise stated, we define the "cost" of an invention as the minimum amount of money needed to conduct a particular R&D project.

[43] *See generally id.* at 1585–87.

[44] *See, e.g.*, Merges & Nelson, *supra* note 17, at 870, 884; Ko, *supra* note 39, at 782.

[45] *See* Graham v. John Deere, 383 U.S. 1, 7–9 (1966).

[46] *See generally* Daniel M. Hausman, *Economic Methodology in a Nutshell*, 3 J. ECON. PERSP. 115, 117 (1989) (discussing the transition from "classical to neoclassical economics").

[47] *See generally* Arti Kaur Rai, *Regulating Scientific Research: Intellectual Property Rights and the Norms of Science*, 94 NW. U. L. REV. 77, 86, 116–118 (1999) (arguing that patents internalize the positive externalities of innovation and discourage free-riding); Forman, *supra* note 16, at 663 (identifying "traditional view of the patent bargain" in terms of free ridership, public goods, and monopoly pricing); Hill, *supra* note 40, at 238 (noting that freeriding may delay invention).

[48] Kitch, *supra* note 1, at 266. In contrast to later commentators, Kitch himself did not adopt a "many theories" viewpoint, saying, "[t]he reward theory is not questioned on its own terms. Rather it is argued that the reward theory offers an incomplete view of the functions of the patent system." *Id.*

[49] Burk & Lemley, *Policy Levers*, *supra* note 16, at 1603–04.

[50] *Id.* at 1615–16.

[51] Forman, *supra* note 16, at 665; Kenneth W. Dam, *The Economic Underpinnings of Patent Law*, 23 J. LEGAL STUD. 247, 253 (1994).

[52] Ko, *supra* note 39, at 801.

[53] *See generally* Burk & Lemley, *Policy Levers*, *supra* note 16, at 1583.

[54] *See* Forman, *supra* note 16, at 665 (arguing that Kitch's theory is focused on "the further development of prospects"); Rai, *supra* note 47, at 120–21.

[55] Teresa M. Summers, *The Scope of Utility in the Twenty-First Century: New Guidance for Gene-Related Patents*, 91 GEO L.J. 475, 491–93 (2003); *see also* Ko, *supra* note 39, at 799 (incentive

Kitch's theory was based on accidental duplication, i.e. the idea that non-interacting inventors would inevitably duplicate each other's work.[57] **Racing** theory extends this idea by pointing out that the case for interacting inventors is even worse.[58] Patent law, after all, only rewards the *first* inventor while second-comers get nothing. The resulting patent races can be beneficial to the extent that they accelerate innovation and provide redundancy against failure.[59] But the drive to be first—even by an hour or two[60]—can also force competing R&D programs deep into diminishing returns. This, of course, is wasteful. As in the classical reward literature, racing theory teaches that policymakers should pick whichever reward level offers the greatest net benefit.[61]

Finally, **cumulative innovation** addresses how the patent reward should be divided when multiple inventors contribute to a common technology.[62] Here, the key insight is that a first generation ("1G") invention offers two distinct benefits to society. First and most obviously, it yields *present value* by telling society how to make products for immediate use.[63] Second, it delivers *option value* by providing a platform that second generation ("2G") inventors can build on to invent still more products.[64] Ideally, inventors should proceed whenever *all* benefits *including* option value exceed total costs. This cannot happen, however, so long as each inventor's reward is limited to the revenues earned by his or her products. In this case, inventors who produce intermediate (pre-commercial) inventions receive nothing for option value so that the technology is never developed at all. Correcting this problem means allocating part of the 2G (3G, 4G, 5G . . . ) inventor's revenues back to the invention's forebears. We return to this problem below.

---

to innovate); A. Samuel Oddi, *Beyond Obviousness: Invention Protection In the Twenty-First Century*, 38 AM. U. L. REV. 1097, 1110 (1989).

[56] Hill, *supra* note 40, at 238–39. Hill somewhat confusingly divides prospect theory (suppressing wasteful and duplicative investments) from innovation theory (covering commercialization costs). *Id.* Vertinsky argues that Kitch's prospect theory is more usefully seen as an aspect of commercialization. Vertinsky, *supra* note 40, at 221.

[57] Rai, *supra* note 47, at 819–20.

[58] Mark A. Lemley, *The Myth of the Sole Inventor*, 110 Mich. L. Rev. 709, 752 (2012).

[59] *Id.* at 753–54.

[60] *See* Peter Carlson, *The Bell Telephone: Patent Nonsense?*, WASH. POST (Feb. 20, 2008), http://articles.washingtonpost.com/2008-02-20/news/36916760_1_bell-great-grandson-seth-shulman-telephone-gambit (describing how Alexander Graham Bell filed his patent application for the telephone the same day as his nearest rival, Elisha Gray); *see also* Dolbear v. Am. Bell Tel. Co., 126 U.S. 1, 18 (1888).

[61] Greenspoon & Cottle, *supra* note 16, at 197; *see also* George W. Dent, Jr., *The Essential Unity of Shareholders and the Myth of Investor Short-Termism*, 35 DEL. J. CORP. L. 97, 126–27 (2010).

[62] *See, e.g.*, SCOTCHMER, INNOVATION AND INCENTIVES, *supra* note 3, at 127–59; Jerry R. Green & Suzanne Scotchmer, *On the Division of Profit in Sequential Innovation*, 26 RAND J. ECON., 20, 20–21 (1995); Suzanne Scotchmer, "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*, 26(1): 20-33 (1995); S. Scotchmer, "*Standing on the Shoulders of Giants: Cumulative Research and the Patent Law*, 5 J. ECON PERSP. 29, 30 (1991) [hereinafter Scotchmer, *Shoulders of Giants*]; Ofer Tur-Sinai, *Cumulative Innovation in Patent Law: Making Sense of Incentives*, 50 IDEA 723, 731, 741–42 (2010).

[63] *See* Scotchmer, *Shoulders of Giants*, *supra* note 62, at 31.

[64] Green & Scotchmer, *supra* note 62, at 22.

So far we have stressed the differences between our different neoclassical innovation models. These invariably involve contingent facts (one inventor versus many, non-interacting inventors versus racing inventors) that differ from industry-to-industry. What really matters, though, are the commonalities. First, each of the foregoing literatures assumes that firms maximize profits.[65] Second, each assumes that actors pursue this goal rationally.[66] Third, each invokes the same methodology—supply, demand, and occasionally game theory. Finally, all depend on the same mathematical machinery of calculus and optimization. These commonalities guarantee that the putatively separate theories cannot disagree in any deep sense. In keeping with our physics discussion, we should instead think of these individual theories as fact-dependent special cases. This encourages us to look for unifying principles and rules of general application.[67]

### *D. Beyond Neoclassical Theory*

In keeping with the legal literature's conventional categories, we end with theories that reject the assumptions, methods, and even goals of neoclassical economics. These have enjoyed something of a renaissance in the two decades since Merges and Nelson wrote their article.

The Anticommons concept is by far the biggest development.[68] Though often denoted a theory,[69] a close reading of Professors Michael Heller and Rebecca Eisenberg's seminal paper shows that the idea is really a conjecture that—notwithstanding naïve neoclassical predictions—intellectual property owners routinely fail to negotiate licenses that would allow otherwise profitable transactions to go forward.[70] Heller and Eisenberg then offer three reasons ("structural concerns")

---

[65] *See, e.g.*, Greenspoon & Cottle, *supra* note 16, at 197 (noting that each patent theory is based on the assumption that an actor is attempting to "maximize his or her own private welfare").

[66] *See, e.g.*, *id.* (noting that each patent theory is based on the assumption that "the individuals or entities who innovate and patent are wealth-seeking rational actors").

[67] This situation is normal in law, where legislators and judges almost always act on less-than-perfect information.

[68] *See* Michael A. Heller, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621, 624 (1998) [hereinafter Heller, *Anticommons*]; Michael A. Heller, *The Boundaries of Private Property*, 108 YALE L.J. 1163, 1166–67 (1999); Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCI. 698, 698–99 (1998) (extending Anticommons concept to patent law); MICHAEL HELLER, THE GRIDLOCK ECONOMY: HOW TOO MUCH OWNERSHIP WRECKS MARKETS, STOPS INNOVATION, AND COSTS LIVES 1–22 (Basic Books, 2008) (providing a more recent and detailed account of the Anticommons conjecture).

[69] *See, e.g.*, Robert P. Merges, *Autonomy and Independence: The Normative Face of Transaction Costs*, 53 ARIZ. L. REV. 145, 146 (2011).

[70] Heller & Eisenberg, *supra* note 68, at 699–700. Despite more than a decade's effort, statistical studies have so far failed to find empirical evidence for the Anticommons conjecture. *See, e.g.,* Tom Magerman, Bart Van Looy, & Koenraad Debackere, *In Search of Anti-Commons: Patent-Paper Pairs in Biotechnology, an Analysis of Citation Flows* (Druid Soc'y, Working Paper, 2011), *available at* http://druid8.sit.aau.dk/acc_papers/ikjb8ctjuic82cfejdxmtj9ab17k.pdf; Chris Holman, Commentary, *Clearing a Path Through the Patent Thicket*, 125 CELL 629, 630 (2006) [hereinafter Holman, *Clearing a Path*]; John P. Walsh, Charlene Cho & Wesley M. Cohen, *Patents, Material*

why this could happen.[71] The first ("transaction costs" and "strategic behaviors") notes that even neoclassical theory predicts occasional deadlock.[72] Examples include so-called mixed strategy games, in which players veto profitable agreements in hopes of being bought off,[73] and the familiar dynamic that forces owners of monopoly complements to set higher prices than a cartel would.[74]

If this were all, one could still hope to unify the Anticommons with neoclassical theory. However, Heller and Eisenberg's next two concerns make this impossible. The first denies profit maximization.[75] All would be well, the authors argue, if every patent owner was trying to maximize profit.[76] But suppose the parties pursue different and perhaps incompatible goals like a desire to hurt rivals?[77] In such a world, conventional microeconomic assumptions no longer hold.

Heller and Eisenberg's final critique is even more fundamental. Neoclassical theory assumes rationality. However, "cognitive biases" could lead actors to "systematically overvalue their assets and disparage the claims of their opponents."[78] Taken literally, this implies that no two parties can ever agree on an invention's value.[79]

Given that so much of the Anticommons concept is based on rejecting neoclassical assumptions, it is hard to see how any kind of unification is possible. Still, one could imagine a situation where judges only ignore neoclassical predictions in cases where Anticommons effects are strong. But when, exactly, is that? For transactions involving two negotiators? For ten? For one hundred? The Anticommons literature does not say.

**Patent Thickets.** Neoclassical theory works best when transaction costs are small and can be safely neglected.[80] This condition is violated when patent searches are prohibitively expensive.[81] The question is, what then? Within the neoclassical framework, economists usually invoke "transactions cost" arguments as a reason to

---

*Transfers and Access to Research Inputs in Biomedical Research*, (Sept. 20, 2005), http://www2.druid.dk/conferences/viewpaper.php?id=776&cf=8. It is worth noting that Walsh et al. find that companies take longer to develop scientific opportunities in heavily patented fields. *Id.* However slender, this evidence suggests that Anticommons effects probably do exist at some level.

[71] Heller & Eisenberg, *supra* note 68, at 700.

[72] *Id.*

[73] Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 73 TEX L. *Rev.* 989, 1058–59 (1997) [hereinafter Lemley, *Improvement in IP*] (noting that rights holders may play a game of chicken in which each argues that he is "irrational enough to kill the whole deal unless you get more than your 'fair' share," and that this strategy can "result in the parties failing to come to terms at all").

[74] *See, e.g.*, SCOTCHMER, INNOVATION AND INCENTIVES, *supra* note 3, at 142–43.

[75] Heller & Eisenberg, *supra* note 68, at 700.

[76] *Id.*

[77] *Id.* For example, Professor Lemley remarks that a government agency might try to maximize public access, an author might try to suppress hostile reviews, and a business might want to hurt rivals. Lemley, *Improvement in IP*, *supra* note 73, at 1059–61.

[78] Heller & Eisenberg, *supra* note 68, at 701.

[79] *See id.* at 700–01.

[80] DOUGLASS C. NORTH, TRANSACTION COSTS, INSTITUTIONS, AND ECONOMIC PERFORMANCE 6 (1992).

[81] T. Randolph Beard & David L. Kaserman, *Patent Thickets, Cross-Licensing, and Antitrust*, 47 ANTITRUST BULL. 345, 356–57 (2002).

organize existing institutions differently. But suppose that the patent system is already (nearly) as efficient as it can be. Then transactions costs are unavoidable and we can imagine three distinct situations.

*Small Transactions Costs.* If patent thicket effects are unimportant, policymakers can ignore them.

*Moderate Transactions Cost.* Society should never invest more in an R&D project than it hopes to gain from innovation. If transactions costs are truly unavoidable, the fact that these costs involve, say, lawyers instead of bench scientists, is immaterial. From this standpoint, we should be *grateful* if transactions costs stop R&D projects from going forward.

*Large Transactions Costs.* It is possible to imagine a world where transaction costs make all R&D projects prohibitively expensive. Here, patents yield few benefits and should be abolished.

The problem, for now, is that the Patent Thicket literature—like the Anticommons—says almost nothing about how to identify these cases in practice.

**Other Theories.** One can, of course, conjure up additional theories indefinitely. In practice, this usually means suspending particular neoclassical assumptions. Professor Vertinsky has compiled an extended list of how non-neoclassical theories based on bounded rationality, imperfect information, and/or opportunism could potentially improve doctrine.[82] Strikingly, however, almost all of these insights are narrow or ambiguous. While they may sometimes extend neoclassical theory, they cannot replace it.

**Natural Law.** So far, we have focused on theories that reject neoclassical assumptions. However, the goal of neoclassical theory—maximizing society's total wealth—is also a choice. If we want to, we can set different goals. For natural law, this usually means justice to the inventor.[83]

The problem, once again, is predictive power.[84] Natural law provides a reasonably precise theory of why someone who creates a physical object should own it. But re-potting this idea in patent law leads to difficulties. How broad should patent right be? Should the patent right (unlike physical property) expire? And if it

---

[82] Vertinsky, *supra* note 40, at 221; *see also* Molly Shaffer Van Houweling, *Author Autonomy and Atomism in Copyright Law*, 96 VA. L. REV. 549, 641–42 (2010) (asserting that proliferation of works subject to copyright, number of people owning copyrights, and type/size of separately owned rights have complex and often contradictory impacts on author autonomy, publishers' ability to license, competition, diversity of expression, and distributive fairness).

[83] For a modern version of these arguments, see generally ROBERT P. MERGES, JUSTIFYING INTELLECTUAL PROPERTY 195–236 (2011). A detailed account of the 19th Century debate can be found in Janis, *supra* note 16, at 155–59.

[84] The predictive power of natural law suffers from the special problem that intuitions about fairness and moral rights are notoriously subjective. This subjectivity can infect theory to the point where it becomes too formless and *ad hoc* to offer useful guidance. Philosophers traditionally manage this problem by insisting that any fairness judgment proceed by rigorous argument from clear assumptions. Professor Merges has recently performed this service by deriving intellectual property from conventional natural rights principles already familiar from the works of Kant, Locke, and Rawls. The good news is natural law in Merges' hands is far clearer and more prescriptive. The bad news is that the gap between natural rights and conventional utility theory is, if anything, even clearer and deeper than it was before. *See* MERGES, *supra* note 83, at 31–101.

does, how long should temporary ownership last? Natural law does not answer these questions.

### *E. Privileging The Neoclassical Core*

Neoclassical theory is remarkably unified and coherent. We argue in the next section that these qualities provide clear and determinate guidance for doctrine. At the same time, we should be careful to construct doctrine on the broadest and most general foundations available. Conversely, we should only discard rival, non-neoclassical theories—some of which routinely appear in the legal literature—when unification is impossible so that we are forced to choose between neoclassical theory and its narrower and less useful rivals.

The obstacles to unification are deepest for natural law. The basic difficulty is that the elegance and power of neoclassical theory depends on a very special ("Pareto optimal") definition of economic efficiency.[85] But natural law starts from a very different ("fairness to the inventor") goal. Assuming that it could be constructed at all, any theory that combines utility and natural law is bound to be clumsy and indeterminate. And of course, American law has already made its choice. As Professor Lemley remarked, the "primacy of incentive theory in justifying intellectual property" is anchored in the language of the Constitution's Patent and Copyright Clause[86] as interpreted by case law.[87]

The prospects for unification within other, non-neoclassical branches of economics are brighter. Indeed, economists routinely construct non-neoclassical theories.[88] The problem, for now, is that the Anticommons and patent thickets are not yet theories in this sense. Should judges start to worry when the number of overlapping patents is two, or eight, or fifty? Are there simple institutional arrangements – for example cross-licenses that can overcome the Anticommons? The literature does not say.

In many ways, the situation is similar to antitrust law, which invokes neoclassical economics to offer a single cure ("more competition") to the world's ills.[89] Strikingly, judges do this even though microeconomic theory *acknowledges* that competition can sometimes lead to disaster. As the Supreme Court emphasized in

---

[85] LaRue Tone Hosmer, *Managerial Ethics and Microeconomic Theory*, 3 J. BUS. ETHICS 315, 317 (1984).

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

[87] Lemley, *Improvement in IP*, *supra* note 73, at 993; *see also* Merges & Nelson, *supra* note 17, at 852 ("While there are those who may challenge the propriety of these goals from the standpoint of economic policy, it is both realistic and necessary for courts and the patent office to pursue the goals implicit in both constitutional and statutory provisions."); Vertinsky, *supra* note 40, at 220–21 ("[T]he traditional justification for the patent regime in the United States has been largely a utilitarian one based on the public goods aspect of invention.").

[88] *See* Vertinsky, *supra* note 40, at 224–25.

[89] Kenneth G. Elzinga, *The Goals of Antitrust: Other Than Competition and Efficiency, What Else Counts?*, 125 U. PA. L. REV. 1191, 1195 (1977).

*Trinko*, such evils as do occur are best left to other bodies of law.[90] This article argues that patent doctrine should similarly build on the underlying coherence of neoclassical theory. To the extent that neoclassical assumptions fail, policymakers can always fall back on tax policy, sponsored research, and other non-patent incentives.

## II. THREE POLICY GOALS

The fact that neoclassical theories are logically consistent encourages us to think that we can write rules of general application. This Section begins the process by extracting three policy goals that appear throughout the neoclassical literature.

### *A. Goal No. 1: Avoiding Redundant Incentives*

Patents are superfluous for products that would be invented anyway.[91] This can happen for several reasons. First, inventors often obtain temporary monopolies even without formal patents. This can happen because of "first-mover advantages" or because they already possess market power over related goods. Second, society offers many incentives besides patents, including prizes, grants, and subsidies. Finally, the rise of open source communities reminds us that human beings are inherently creative.[92] This basal rate of innovation would presumably continue even if patents disappeared entirely.[93] In all of these cases, granting a patent imposes a tax on innovation to no purpose.

These observations suggest that the patent system does not have to cover every inventor's full R&D costs. The problem is deciding where to draw the line. Professors Wesley Cohen, Richard Nelson, and John Walsh have collected survey evidence showing that CEOs in most industries see patent incentives as relatively unimportant.[94] However, their work also shows that the answer varies by industry.[95]

---

[90] Verizon Comms. Inc. v. Law Offices of Curtis v. Trinko, 540 U.S. 398, 411 (2004).

[91] Phanesh Koneru, *To Promote the Progress of Useful Art[icle]s?: An Analysis of the Current Utility Standards of Pharmaceutical Products and Biotechnology Research Tools*, 38 IDEA 625, 632 (1998) (explaining that patents should not be granted where inventions would be developed anyway "for reasons unrelated to the existence of the patent reward"); Summers, *supra* note 55, at 486 (noting that "sufficient incentives for basic inventive activity [may] exist without the grant of a patent").

[92] Purists may object that open source licenses routinely assert copyright and are themselves a form of intellectual property. In practice, however, volunteers almost always join because of softer incentives like altruism or a desire for education. By comparison, copyright's role is secondary and usually focuses on deterring negative behaviors that might otherwise destabilize the collaboration. *See* Stephen M. Maurer, *The Penguin and the Cartel: Rethinking Antitrust and Innovation Policy for the Age of Commercial Open Source,* UTAH L. REV. 269, 294 (2012).

[93] The idea makes an early appearance in H. G. WELLS, THE WORLD SET FREE (1914). Wells imagines a nuclear-powered world in which people respond to increased leisure time by becoming more creative. Predictably, the novel later takes a darker turn.

[94] Wesley M. Cohen, Richard R. Nelson & John P. Walsh, *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (Or Not)* 3 (Nat'l Bureau of

Facts matter. Conversely, theory alone says very little about which inventions should and should not be patented.

The Supreme Court famously addressed this problem in the 1960s. Its *Graham*[96] decision acknowledged the importance of "'drawing a line between the things which are worth to the public the embarrassment of an exclusive patent, and those which are not.'"[97] It also announced a test: inventions should be patentable if and only if they were "non-obvious" to a PHOSITA.[98] We argue below that this is a very useful concept. At the same time, the Court failed to explain why it put the dividing line where it did. To the contrary, basing the standard on "average" skill sounds like a confession of ignorance. Why not choose "above-average" skill or a "blinding flash of inspiration?"[99] We discuss these and other alternatives in Section 5.

### *B. Goal No. 2: Selecting the Right Level of Effort*

An inventor can almost always create products sooner or more reliably if she is willing to spend more. What she actually does depends on the reward. Here the most natural guess is that the inventor will only invest where the expected reward[100] covers her R&D cost. On the other hand, society could rationally decide that it is worth offering larger rewards to get inventions sooner or more reliably. The problem, of course, is that this implies broader and/or longer-lasting patent monopolies. Given diminishing returns, we should expect each increment of reward to provide less benefit than the one before. For this reason, theory predicts that there is some maximum patent duration and breadth beyond which further increases are counterproductive. In Professor Kitch's elegant summary, "The courts, influenced by the reward theory, view the patent system as a difficult problem of trade-offs between the incentive effects and the output constraining effects."[101] Once again, the correct tradeoff is heavily fact-dependent, so that theory provides only limited guidance.

U.S. law defines patent breadth to embrace (a) whatever the inventor actually claimed and disclosed in her patent, together with (b) certain "non-equivalent" improvements and follow-on inventions. While the latter doctrine is vague, we argue in Part VII that it almost certainly implies a PHOSITA-like cleverness standard. In practice, judges and commentators usually assume that this standard is identical to

---

Econ. Research, Working Paper No. 7552, 2000), *available at* www.dklevine.com/archive/cohen-survey.pdf.

[95] *Id.*

[96] Graham v. John Deere Co., 383 U.S. 1 (1966).

[97] *Id.* at 10 (quoting Thomas Jefferson).

[98] *Id.* at 3.

[99] *See* Cuno Eng'g Corp. v. Automatic Devices Corp., 314 U.S. 84, 90–91 (1941) (patentable device "must reveal the flash of creative genius.")

[100] The "expected value" of a reward refers to the sum of all possible rewards with each reward multiplied by its probability of occurrence. *Expected Value Definition*, MERRIAM-WEBSTER DICTIONARY, http://www.merriam-webster.com/dictionary/expected%20value (last visited June 1, 2013).

[101] Kitch, *supra* note 1, at 282.

*Graham*'s non-obviousness test. However, this is not the only choice and we will see that Congress could pick a different cleverness standard if it wanted to.

### *C. Goal No. 3: Allocating Reward Among Multiple Inventors.*

We have already noted that many inventions are "cumulative" in the sense that progress depends on insights contributed by multiple, often-independent inventors over time. Consider, for the sake of definiteness, the case where a first generation ("1G") invention establishes a new technology that receives 2G/3G/4G improvements and culminates in a saleable 5G product. Assume further that success is uncertain because the 5G product may turn out to be technically impossible or less popular than investors originally hoped. Then an economically efficient patent system must satisfy at least three separate and distinct criteria.

*Allocate Profits.* The 1G (2G, 3G . . . ) R&D programs will never be funded unless patent owners expect to share in 5G revenues. Provided that the 5G product earns enough revenue, each inventor in the chain should expect to recover his or her investment.

*Allocate Losses.* The 1G (2G, 3G . . . ) investors know that they could easily end up incurring more R&D costs than the 5G product will ever earn. In this case, someone must bear the loss. Here, the principle of "sunk costs" insists that investors at each stage should ignore past costs that have already been incurred and are now beyond recovery.[102] Instead, they should base their decision exclusively on the costs of going forward.

*Maximum Efficient Investment.* Deciding whether to invest in an R&D program is not the only question; investors must also decide how much. In general, we expect each incremental increase in R&D budget to (on average) deliver more and better products to consumers. This investment should only be made, however, if these incremental benefits exceed incremental cost. This rule correctly prices the option value of new technologies.

The economics literature provides a clear rule that gets these incentives right.[103] Each inventor should (a) retain enough revenue to cover her R&D cost,[104] and (b) pay the balance to whichever inventor immediately preceded her. Following Professor Ofer Tur-Sinai, we shall refer to this as the "Absolute Scope" rule in what follows.[105]

---

[102] *Sunk Cost Definition*, MERRIAM-WEBSTER DICTIONARY, http://www.merriam-webster.com/dictionary/sunk%20cost (last visited May 16, 2013) ("[A] cost already incurred that is not subject to variation or revision.").

[103] The proposed rule follows SCOTCHMER, INNOVATION AND INCENTIVES, *supra* note 3, at 127. In principle, there may be other rules that satisfy our three criteria.If so, I am unaware of them.

[104] More specifically, we want investors to recover revenue consistent with our first and second goals. Rewards that simply repay R&D investments may be insufficient to the extent that inventors lack, for example, access to capital.

[105] Tur-Sinai, *supra* note 62, at 758.

### *D. Reimagining Patent Doctrine*

The next six sections ask how well current patent rules conform to modern economic insights. We proceed as follows. Part IV ("Reimagining Utility") asks when an invention with no immediate end-use nevertheless offers sufficient social value to be patentable. Part V ("Reimagining the PHOSITA") argues that the *Graham* standard is best understood as a measure of cleverness and, indirectly, R&D cost. In the process, we generalize the PHOSITA concept to include the full range of cleverness standards available to Congress. Parts VI–VIII analyze how well patent doctrines that include the PHOSITA concept implement our third policy goal. Finally, Parts IX and X look at the current system's shortcomings and ask whether radical change is worth pursuing.

## III. REIMAGINING UTILITY

Any attempt to allocate reward among sequential inventors must start with the threshold question of when 1G (2G, 3G . . . ) inventors should be allowed to share in 2G (3G, 4G . . . ) revenues. Following the innovation economics literature, we argue that sharing should take place whenever inventions create "option value" for later inventors to build on.[106] We begin by reviewing judges' efforts to generalize the utility concept to inventions that have no present use and whose worth is based entirely on option value. We then suggest a revised rule that rationalizes earlier cases and brings doctrine into closer agreement with our third policy goal.

### *A. Modern Utility Doctrine*

Early U.S. law provided that only "sufficiently useful and important" inventions could be patented.[107] This seemed to say that patent doctrine requires applicants to show some minimum usefulness to consumers in the same way that *Graham's* non-obviousness standard requires them to show cleverness.[108] The approach did not last long. In 1817, Judge Story downgraded the utility requirement to capable of use and not "frivolous."[109] This implied a *de minimis* standard in which "any quantity of

---

[106] *See* Green & Scotchmer, *supra* note 62, at 22.

[107] Patent Act of 1836, ch. 357, § 7, 5 Stat. 117, 120 (1836) (amended 1870).

[108] Michael Risch, *Reinventing Usefulness*, 2010 BYU L REV. 1195, 1207, 1236 (2010) [hereinafter Risch, *Reinventing Usefulness*]. Some federal courts continued to describe non-*de minimis* utility thresholds in dictum well into the 20th Century. *Id.* The heresy was so prevalent that contemporary scholars and jurists repeatedly warned practitioners not to confuse "positive utility" with its "degree." N. Scott Pierce, *In re* Dane K. Fisher*: An Exercise in Utility*, 6 J. HIGH TECH. L. 1, 2 (2006); *see also In re* Kirk*,* 376 F.2d 936, 954 (C.C.P.A. 1967)

[109] Lowell v. Lewis 15 F. Cas. 1018, 1019 (C.C.D. Mass. 1817) (No. 8568); *see also* Bedford v. Hunt, 3 F. Cas. 37, 37 (C.C.D. Mass. 1817) (No. 1217) (requiring only that invention be "capable of use"). The *Lowell* opinion famously added that inventions should also not be "injurious to the well-being, good policy, or sound morals of society." *Lowell*, 15 F. Cas. at 1019. The Federal Circuit overruled this "moral utility" element in Juicy Whip, Inc. v. Orange Bang, Inc., 185 F.3d 1364 (Fed. Cir. 1999).

utility should suffice."[110] Judge Story justified this result on the theory that patents on worthless inventions were irrelevant in any case, since such learning would "silently sink into contempt and disregard."[111] In modern phrasing, the point seems to be that the damage from monopoly (i.e., deadweight loss) can never be larger than the monopolized product's value to users. For this reason, there is nothing to be gained from refusing to patent inventions that nobody wants. We can also glimpse a second argument in Story's comment: given that patenting costs money, no rational inventor will ever pursue rights to an invention that he expects to be "disregarded." In this situation, the inventor's willingness to invest already provides a market test of value. A utility threshold that encourages judges to second-guess the inventor overrules the very signal that patents are supposed to elicit.

For the most part, Judge Story's rules are still valid today. Starting around 1950, however, courts reintroduced a non-*de minimis* threshold for chemical and biological inventions.[112] The basic doctrinal puzzle was that these industries—unlike, say, the mechanical or electrical arts—were willing to invest in a "process for producing a product of only conjectural use" or "an intermediate that can be used to produce another class of compounds."[113] On the face of things, courts took a hard line by demanding that some use must "presently exist."[114] This, however, would have ruled out intermediate inventions that commercial firms were plainly willing to pay for. For this reason, judges softened the rule by noting that utility could still be shown where "person[s] skilled in the art" would immediately recognize the product's capabilities provided that they were not "conjecture" or else so "unpredictable" as to require "considerable experimentation."[115] Even so, the rule ignored the fact that

---

[110] *Lowell*, 15 F. Cas. at 1910; *see also* Risch, *Reinventing Usefulness*, *supra* note 108, at 1207. The invention did not have to perform better than competing prior art technologies. *In re* Ratti, 270 F.2d 810, 815 (C.C.P.A. 1959) (finding that an invention that was "no better than" the prior art was nevertheless patentable).

[111] *Lowell*, 15 F. Cas. at 1019.

[112] According to Judge Rich, pre-1950 chemical patent applications "were commonly granted although no resulting end use was stated or the statement was in extremely broad terms." *In re* Kirk, 376 F.2d 936, 950 (C.C.P.A. 1967) (Rich, J., dissenting). Judge Rich believed that the *Story* rule should apply to both intermediate and end-use discoveries. *Id.* at 1142. Like most *ad hoc* rules, the idea that courts treat utility differently for chemical and biological inventions than other arts is unsettling. Yet, as Professor Andrew W. Torrance (personal communication) remarks, it is probably true that a poorly understood gene is more likely to be rejected on utility grounds than, say, an unsittable chair. The practical reason for this is that the number of unsittable chair patents, unlike chemical patents, is vanishingly small. In these circumstances, judges may feel safer issuing a bad precedent than in invalidating a patent despite lingering doubts that the unsittable chair may, after all, possess some overlooked social value. This section argues that the best way to eliminate such doubts is to improve doctrine so that it more accurately reflects social value. Along the way, we argue that some unsittable chairs really could have social value as stepping stones to commercial (i.e., sittable) designs. If this improved understanding gives judges the confidence they need to strike down genuinely useless chair patents, so much the better.

[113] 1 DONALD S. CHISUM, CHISUM ON PATENTS § 4.02[2][e] (2012).

[114] *Ex parte* Tolkmith, 102 U.S.P.Q. 464, 466 (B.P.A.I. 1954).

[115] *Id.* at 466. The fact that the compound was new did not bar patentability provided that it "belong[ed] to a class of compounds the members of which have become well recognized to be useful for a particular purpose, and it is evident from the prior art that it is within the skill of the art to use the claimed compound for this purpose." *Ex parte* Ladd, 112 U.S.P.Q. 337, 338 (B.P.A.I. 1955).

many industries really do fund 2G projects based on unproven conjectures. In these cases, at least, doctrine failed to protect option value.

The U.S. Supreme Court tried to rationalize the situation in 1966. Its *Brenner* opinion asked whether a chemical compound, whose therapeutic effects had never been tested, should nevertheless be patentable.[116] The justices rejected this expansive result and instead announced that intermediate inventions would only qualify for protection if they possess "specific benefit" and "substantial utility" in a "currently available form."[117] The Court also advanced two policy arguments to justify its holding. First, it raised the familiar point that inventions should relate to "commerce" instead of "philosophy."[118] This distinction, however, seemed redundant because the same concepts also helped to define patentable subject matter.[119] Just what did utility add to the exercise?[120] Perhaps sensing the problem, the Court introduced a second and more novel objection: "A patent," Justice Fortas observed, should be more than a "hunting license."[121] While this sounds like a rejection of option value, the Court was more likely referring to a very different problem in which applicants patent worthless inventions in hopes of capturing after-arising technologies that they did nothing to facilitate.[122]

Courts have now spent half a century trying to generalize *Brenner* beyond its facts. The earliest cases staked out an obvious bright line rule: the utility threshold

---

[116] Brenner v. Manson, 383 U.S. 519, 520 (1966).

[117] *Id.* at 534–35.

[118] *Id.* at 536. The Court's argument continues to influence current scholars who argue that utility provides a "timing" function for determining when an invention is "ripe." Rebecca S. Eisenberg, *Analyze This: A Law and Economics Agenda for the Patent System*, 53 VAND. L. REV. 2081, 2087 (2000); Rebecca S. Eisenberg & Robert P. Merges, *Opinion Letter as to the Patentability of Certain Inventions Associated With the Identification of Partial cDNA Sequences*, 23 AIPLA Q.J. 1, 6 (1995) (Utility "serves a timing function, leaving basic research discoveries in the public domain until they have yielded tangible benefits and have thereby left 'the realm of philosophy' and entered 'the world of commerce."); Andrew T. Kight, *Pregnant with Ambiguity: Credibility and the PTO Utility Guidelines in Light of* Brenner, 73 IND. L.J. 997, 1012 (1998) ("The utility requirement operates to distinguish between basic research and applied technology."); Arti K. Rai, *Fostering Cumulative Innovation in the Biopharmaceutical Industry: The Role of Patents and Antitrust*, 16 BERKELEY TECH. L.J. 813, 839 (2001) ("In practice, however, only the utility requirement serves as a particularly good proxy for differentiating upstream from downstream research."); Risch, *Reinventing Usefulness*, *supra* note 108, at 1220 ("[B]asic science, no matter how important and valuable, does not merit protection and is therefore not useful in the patent sense.").

[119] *See* Bilski v. Kappos, 130 S. Ct. 3218, 3249 (2010) (Stevens J., concurring) (patentable subject matter does not include mere philosophical principles).

[120] There is, of course, no rule against using multiple redundant doctrines to enforce a single policy goal. Indeed, this can be forced on judges by poorly drafted statutes or tangled precedent. At the same time, the practice invites confusion and risks "using up" levers that may be needed for other policy goals.

[121] *Brenner*, 383 U.S. at 536.

[122] Because all legal systems are inherently imperfect, patent owners can sometimes assert accidents of language to claim inventions that were actually developed by others. Judges have long resisted this outcome. Dr. Zuhn points out that common law mining law similarly prevents miners from claiming whatever minerals lie beneath randomly selected plots of land. Donald L. Zuhn, Jr. *DNA Patentability: Shutting the Door to the Utility Requirement*, 34 J. MARSHALL. L. REV. 973, 973 (2001). *See also* William E. Ridgway, *Realizing Two-Tiered Innovation Policy Through Drug Regulation*, 58 STAN. L. REV. 1221, 1221 (2006) (setting the utility threshold too low leads to inventors "stockpiling" instead of "developing" innovations).

excludes products whose only known use is to make other products that similarly lack known uses.[123] This rule was most recently applied in *In re Fisher*, which denied patentability to a list of genetic sequences—so-called "Expressed Sequence Tags" or "ESTs"—that lack known functions.[124] The reason, according to the Federal Circuit, is that ESTs are mere "research intermediates" presenting no "assurance that anything useful will be discovered in the end."[125]

The case is harder for inventions that claim specific uses. Here, post-*Brenner* utility analyses usually devolved into arguments over whether the evidence was sufficient to prove the asserted capability.[126] Significantly, the required proof of utility was much lower than the human trials that were needed to obtain FDA approval and achieve "commercial usefulness"[127] Instead, utility could be established by animal tests and chemical similarity to previously-known drugs[128] so long as these methods offered a "reasonable correlation" or were "generally predictive" of end-use capabilities.[129] Not surprisingly, judges identified such proofs with the PHOSITA concept, i.e., whether "one of ordinary skill in the art would accept appellant's claimed utility in humans as valid and correct"[130] or else would not doubt the utility

---

[123] *In re* Kirk, 376 F.2d 936, 945 (C.C.P.A. 1967) (rejecting patent claims for insufficient utility, noting that "[i]t is not enough that the specification disclose that the intermediate exists and that it 'works,' reacts, or can be used to produce some intended product of no known use"); *In re* Joly, 376 F.2d 906, 908 (C.C.P.A. 1967) (rejecting utility argument based on "mere disclosure that a claimed chemical compound may be used as an intermediate to make other compounds, without regard for the usefulness of the latter compounds"). Significantly, the *Kirk* court's holdings were based on the fact that many steroids "possess no activity whatsoever" and that the patented compound's "actual uses-or possible lack of uses" could only be determined by additional experimentation. *Kirk*, 376 F.2d at 942. These uncertainties did not, of course, show that the discoveries lacked sufficient "option value" to justify 2G research.

[124] *In re* Fisher, 421 F.3d 1365, 1367 (Fed. Cir. 2005).

[125] *Id.* at 1373 (quoting Brenner v. Manson, 383 U.S. 519, 535 (1966)). Judge Rader's exasperated dissent cogently argued that any research which took society "one step closer to identifying and understanding a previously unknown and invisible structure" had social value and that this value existed whether or not the particular invention represented "the final step of a lengthy incremental research inquiry." *Id.* at 1380 (Rader, J., dissenting).

[126] *Id.* at 1373 ("Fisher has not presented any evidence . . . showing that the claimed ESTs have been used in either way."); *see also In re* Brana, 51 F.3d 1560, 1556 (Fed. Cir. 1995).

[127] *In re* Langer, 503 F.2d 1380, 1393 (C.C.P.A. 1974); *accord In re* '318 Patent Litigation, 583 F.3d 1317, 1324–25 (Fed. Cir. 2009); *Brana*, 51 F.3d at 1568. The distinction between commercial and actual usefulness makes considerable sense. Absent FDA regulation, adventurous (or desperate) consumers might well purchase drugs based on "actual usefulness" shown by animal test results and/or a molecular resemblance to known compounds. "Commercial usefulness," on the other hand, suggests that the drug can be legally sold to consumers, i.e., that it is sufficiently well documented to survive FDA testing. *See id.* (contrasting safety and efficacy with patent law "utility").

[128] *In re* Jolles, 628 F.2d 1322, 1327 (C.C.P.A. 1980).

[129] Cross v. Iizuka, 753 F.2d 1040, 1050 (Fed. Cir. 1985).

[130] *Jolles*, 628 F.2d at 1327–28; *see also* U.S. Patent & Trademark Office, Utility Examination Guidelines, 66 Fed. Reg. 1092, 1093 (Jan. 5, 2001) (arguing that genes should be patentable where "a person of ordinary skill in the art would immediately appreciate why the invention is useful"); *Ex parte* Tolkmith, 102 U.S.P.Q. 464, 466 (B.P.A.I. 1954) (arguing that value of proposed parasiticide is measured by whether its capabilities are apparent to "one skilled in the art").

based on animal testing.[131] While this lent the analysis an aura of scientific objectivity, it also begged the question of how a PHOSITA would decide just when the numeric odds of success were large enough to be deemed "reasonable" or "generally predictive."[132] By comparison, a complete answer would have required some explicit judgment about which 2G projects are worth pursuing—an inquiry that necessarily includes non-technical factors like researcher salaries and consumer demand. Not for the first time, the PHOSITA standard let judges sweep these economic variables under the rug.[133]

Despite doctrinal gymnastics, courts continue to tie utility to presently-existing capabilities that—however hypothetically—consumers might actually want to use. Strangely, judicial policy arguments are seldom so narrow. For example, the Court of Customs and Patent Appeals argued in *Cross* that a low utility threshold is needed to "marshal resources and direct the expenditure of effort to further *in vivo* testing . . . thereby providing an immediate benefit to the public."[134] Similarly, the Federal Circuit has held that *Brenner* utility "necessarily includes the expectation of further research and development."[135] Such arguments seem indistinguishable from option value.

### *B. A Rationalized Utility Standard*

The traditional utility test for ordinary "present value" inventions is simple: does the invention do something that consumers might value? The problem, as we have seen, is that industries often find commercial reasons to invest in inventions that have no known end-use and hence no value to consumers.[136] This puts judges and bureaucrats in the awkward position of telling industry that its most promising discoveries lack "real world" value.[137]

Following Judge Story, we argue that it is simpler—and also more economically efficient—to define value with reference to actual consumers and markets. The question is, *which* markets and *which* consumers? Current doctrine invariably identifies "markets" and "consumers" with end-users. This, however, requires fictitious inquiries into how much consumers would value inventions that will never

---

[131] *In re Brana,* 51 F.3d 1560, 1567 (Fed. Cir. 1995); *but see Ex parte* Aggarwal, 23 U.S.P.Q.2d 1334, 1338 (B.P.A.I. 1992) (finding that evidence that compounds treated tumors in animals did not provide sufficient evidentiary support for human applications).

[132] *Cross*, 753 F.2d at 1050; *see also Jolles*, 628 F.2d at 1327–28 (holding that evidence that inventor's pharmaceutical compositions were analogous to existing anti-cancer agents was sufficient to establish utility).

[133] Courts also debate how "specific" the claimed end use must be. While the requirement is clearly met by compounds that treat particular diseases, judges have stretched the concept to include "known pharmacological activities" like reducing blood pressure. Nelson v. Bowler, 626 F.2d 853, 855–56 (C.C.P.A. 1980).

[134] Cross v. Iizuka, 753 F.2d 1040, 1051 (Fed. Cir. 1985).

[135] *In re Brana,* 51 F.3d 1560, 1568 (Fed. Cir. 1995).

[136] *See supra* note 122 and accompanying text.

[137] U.S. PAT. & TRADEMARK OFF., REVISED INTERIM UTILITY GUIDELINES TRAINING MATERIALS 4 (1999), *available at* http://www.uspto.gov/web/menu/utility.pdf (noting that materials used for industrial R&D lack "real world" value).

be sold to the public. But this fiction is unnecessary. After all, there really is a market for 1G results. Following the innovation literature, we argue that a 1G invention only has value if someone, somewhere, is willing to develop it further.[138] This transforms the issue of what end-users might or might not value into an objective investment decision. The rule is also consistent. Inventions that fail to attract 2G investors are automatically worthless to follow-on (3G, 4G, . . . ) developers, end-users, and society at large.

This option value viewpoint is already implicit in judicial tests that measure value from the PHOSITA's standpoint. However, economic theory sharpens this instinct. Under our rationalized rule, an intermediate invention has value any time an inventor would pay a non-zero (but possibly *de minimis*) fee to develop it. Assuming profit-maximization,[139] this judgment normally turns on three well-defined variables.

*Probability of Success.* The 2G program must have some chance of success. We have already noted that current patent doctrine stresses this scientific judgment.

*Expected R&D Cost.* The 2G project must not be "too costly" to develop. Naïvely, this will be true whenever the expected reward is even slightly greater than expected costs. This neatly replicates Judge Story's *de minimis* rule that inventions should be patentable whenever their economic value is even slightly greater than zero.

*Expected Revenue.* The 2G decision ultimately depends on an economic judgment that consumer demand for the end product is large enough to cover R&D costs.[140]

One benefit of using a PHOSITA standard is that it avoids making these inquiries directly. Instead, following Judge Story, the PHOSITA's judgment that further investment is worthwhile serves as a conclusive test of value. While litigants can address our three economic variables directly, it will often be enough to submit less direct evidence. This could include the opinions of knowledgeable industry observers or objective evidence that similar ideas have been funded in the past. At the same time, our explicitly economic rule reminds courts that such showings can be rebutted by evidence that option value has changed over time. This notably includes cases where R&D costs fall or consumer demand increases. Finally, courts should recognize that option value can sometimes exist even where the 1G inventor cannot name a single 2G application.[141] This can happen where the 1G inventor makes a fundamental discovery but knows almost nothing about other industries or else

---

[138] In the case of research tools, one might equally say to use the tool to conduct at least one experiment. We ignore this essentially semantic distinction in what follows.

[139] Our rule does not really need this assumption. The chance that a 1G invention may lead to further development by non-commercial actors like universities, national laboratories, and open source collaborations is also included in option value.

[140] This judgment may, of course, depend recursively on a perception that 3G inventors are willing to pay licensing fees. But in that case the 3G inventor must believe that a 4G inventor would pay, and so on. Logically, all such chains must eventually end in some judgment about consumers' willingness to pay.

[141] The best-known example is almost certainly the laser, which many contemporaries ridiculed as a "solution looking for a problem." Charles H. Townes, *The First Laser*, *in* A CENTURY OF NATURE: TWENTY-ONE DISCOVERIES THAT CHANGED SCIENCE AND THE WORLD 107–12 (Laura Garwin & Tim Lincoln eds., 2003).

reasonably believes that some exceptionally clever inventor can find applications she never thought of.

## IV. REIMAGINING THE PHOSITA

Economists and lawyers speak different languages. This naturally leads to problems. We have seen that each of our three economic principles is defined by R&D cost. Yet, patent law hardly even mentions this concept.[142] This is not necessarily fatal because law often uses proxies for quantities that are difficult or impossible to measure directly. But if a proxy does exist, we should be able to name it. Fortunately, the choices are limited. Open any introductory textbook and you will see that patent law is constructed around a dozen or so core concepts.[143] Of these, only two concepts—"non-obviousness" and "utility"—say anything about what the inventor has accomplished or, implicitly, invested. Furthermore, we have already seen that utility measures value from the consumer's perspective. This is a very different concept from the inventor's effort or cost. In the end, we are left with just one possible proxy: Non-obviousness.[144]

This section introduces *Graham*'s PHOSITA test and explains how the idea can be generalized to measure different levels of cleverness and, indirectly, effort and cost. We then examine the PHOSITA test's strengths and weaknesses as a cost-proxy. We close by commenting on how the proxy limits the doctrine's ability to implement economically efficient incentives.

### *A. The PHOSITA Standard*

The Federal Circuit's efforts to make patent law more rule-based and predictable have encrusted the PHOSITA with properties that no actual person possesses. On the one hand, she is a polymath who possesses "a complete and thorough knowledge of all legally pertinent prior art."[145] On the other, she is an uninspired drudge who thinks almost entirely "along the line of conventional wisdom" and brings little or no originality to R&D.[146] Clearly, we should not expect

---

[142] As we shall see, the concept of R&D cost plays little or no role in the utility (Part VII), non-obviousness (Part V.B), or blocking patents (Part VIII.B) doctrines. It also plays no role in damages: none of the fifteen *Georgia-Pacific* factors so much as mentions it. Georgia-Pacific Corp. v. United States Plywood Corp., 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970).

[143] *See, e.g.*, ROBERT P. MERGES, PETER S. MENELL AND & MARK A. LEMLEY, INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE x–xii (4th ed. 2006).

[144] We find further encouragement in the fact that judges tend to invoke the PHOSITA concept in the same contexts where innovation economists stress R&D cost. This includes reward size, non-obviousness, and even utility. *See infra* Parts VI.B, VI.C, and VII.

[145] Michael Abramowicz & John F. Duffy, *The Inducement Standard of Patentability*, 120 YALE L. J. 1590, 1606 (2011).

[146] *Id.* at 1606–07 (quoting Standard Oil Co. v. Am. Cyanamid Co., 774 F.2d 448, 454 (Fed. Cir. 1985)). As Professors Abramowicz and Duffy write:

to meet this person in real life. Like the proverbial "reasonable man," she is best seen as a construct or "ghost of the law."[147] We return to these points below.

For now, we take a different path. The phrase "ordinary skill" can be readily identified with "average" or perhaps "median" skill.[148] Compared to our judge-made "ghost of the law," these are objective and indeed measurable characteristics. In what follows, we will think of our PHOSITA as a real person who has a name, a home address, pays taxes, and goes to work most mornings. If we wanted to, we could find her and put her on the witness stand.[149]

We can also generalize the concept. As Professor Suzanne Scotchmer points out, there is no economic reason why the same PHOSITA concept should be used to define *both* non-obviousness *and* patent breadth.[150] Indeed, dividing the PHOSITA concept into two distinct tests would solve various policy problems.[151] Identifying the PHOSITA with an actual person makes this surprisingly easy to do. Recall that *Graham* defined non-obviousness in terms of a PHOSITA having "average skill in the art," meaning that a non-obvious invention is cleverer than fifty percent of all R&D projects.[152] In what follows we will call this threshold "PHOSITA-50," where "50" denotes the invention's percentage cleverness rank compared to all other inventions

---

> [T]he courts have defined the person of ordinary skill to be a rather extraordinary creature, an idiot savant with extraordinary knowledge and virtually no creativity. The mind of this hypothetical person comes equipped with a complete and thorough knowledge of all legally pertinent prior art, far more knowledge than could be possessed by any average or actual researcher . . . . Meanwhile, the Federal Circuit at one time described the person of ordinary skill as someone "who thinks along the line of conventional wisdom in the art and is not one who undertakes to innovate."

*Id.*

[147] Panduit Corp. v. Dennison Mfg., Co., 810 F.2d 1561, 1566 (Fed. Cir. 1987) ("[T]he decisionmaker confronts a ghost . . . not unlike the 'reasonable man' and other ghosts in the law."); *see also,* Kimberly-Clark Corp. v. Johnson & Johnson, 745 F.2d 1437, 1452, 1454 (Fed. Cir. 1984) (noting that the PHOSITA is a "hypothetical person" and "an imaginary being" of the courts' "own devising") (emphasis omitted).

[148] *See, e.g.,* Burk & Lemley, *Biotechnology's Uncertainty Principle*, *supra* note 28, at 713 ("The PHOSITA standard contemplates some median or ordinary skill.").

[149] Courts sometimes recognize this possibility. *Id.* at 710. ("Courts have on occasion equated the knowledge of a given individual, such as a patent examiner, with that of the PHOSITA.").

[150] *See* SCOTCHMER, INNOVATION AND INCENTIVES, *supra* note 3, at 117–18.

[151] *Id.* at 146–149; *see also infra* Part VII.D. Professors Burk and Lemley note that it is doctrinally possible to use different PHOSITA standards in different legal contexts. Burk & Lemley, *Biotechnology's Uncertainty Principle*, *supra* note 28, at 712 ("The PHOSITA for purposes of obviousness may not necessarily be the PHOSITA for purposes of enablement, written description, definiteness, or equivalence.").

[152] Remarkably, our PHOSITA standard has a recursive quality. The cleverness it prescribes cannot be specified in isolation and differs from market to market. To see this, suppose that a PHOSITA-50 standard is implemented in a particular market on Day One. Rational inventors whose inventions fall just below the threshold will immediately invest in making their products cleverer since patentability promises a large jump in value for very little effort. Now, however, this new investment raises the average cleverness level so that some inventions that satisfied the PHOSITA-50 standard on Day One no longer qualify. This encourages a second round of Day Two investments which leads to further Day Three investments and so on. The net result is that the PHOSITA standard improves the average cleverness of all inventions, including many that are never patented.

in the art. But in that case, why not generalize the concept to define additional thresholds for, say, "PHOSITA-33" or "PHOSITA-95"? Obviously, the number of available tests is infinite. We will use the term "PHOSITA-X" to denote this family of standards.[153]

### *B. How Good is the Proxy?*

The PHOSITA-X concept lets juries determine cleverness from the face of the patent. The deeper question is whether cleverness is itself a proxy for R&D cost. We argue that there are at least four good reasons to link cleverness with cost.

*Brute Force Research.* "Genius," as Thomas Edison remarked, is "one percent inspiration" and "ninety-nine percent perspiration."[154] He should know. Edison's search for a workable light bulb required tests of more than 3,000 filaments.[155] While cleverness may have saved him from testing even more candidates, Edison himself was doubtful.[156] Similar brute force searches are common in many modern industries including pharmaceuticals.[157]

*Corporate Research.* Team research is inevitable in any industry– for example, airplanes or computer software – where innovation requires more man-hours than any human lifetime can supply. The fact that companies routinely hire R&D teams shows that cleverness really can be bought and sold as a commodity.

*Self-Employed Inventors.* Many inventors who work at home could be earning wages in commercial R&D projects. Given this opportunity cost, even garage-based R&D programs have an implicit dollar value.

*New Ideas.* The link between money and innovation is thinnest when we ask how innovators get R&D ideas in the first place. Even so, Pasteur reminds us that,

---

[153] As before, we use "X" to denote the invention's percentage cleverness rank compared to all other inventions in the art.

[154] *Genius is One Percent Inspiration, Ninety-Nine Percent Perspiration*, QUOTEINVESTIGATOR.COM (Dec. 14, 2012), http://quoteinvestigator.com/2012/12/14/genius-ratio.

[155] S*ee generally* Consol. Elec. Light Co. v. McKeesport Light Co., 159 U.S. 465, 468–69 (1895) (discussing Thomas Edison's incandescent lamp research).

[156] As Edison later recalled:

> I would construct a theory and work on its lines until I found it was untenable . . . . I speak without exaggeration when I say that I have constructed 3,000 different theories in connection with the electric light, each one of them reasonable and apparently likely to be true. Yet only in two cases did my experiments prove the truth of my theory.

DANIEL STARCH, EDUCATIONAL PSYCHOLOGY 377–78 (1919).

[157] *See* Cath O'Driscoll, *A Virtual Space Odyssey*, NATURE 2 (Apr. 4, 2012), http://www.nature.com/horizon/chemicalspace/background/pdf/odyssey.pdf (noting that pharmaceutical companies test roughly 5,000 candidates for each drug approved); *About Asterand,* ASTERAND, http://www.asterand.com/Asterand/about/index.htm (last visited May 16, 2013) (estimating 5,000-10,000 candidates per successful drug). Readers can find an extended discussion of the drug discovery pipeline in Solomon Nwaka & Robert G. Ridley, *Virtual Drug Discovery and Development for Neglected Diseases Through Public-Private Partnerships*, 2 NATURE, 919, 919 (2003).

"[c]hance favors the prepared mind."[158] This suggests that paying people to think about a problem can increase the supply of "prepared minds" and, implicitly, the flow of ideas.

These observations confirm the intuition that money can induce cleverness. At the same time the cleverness/cost link can fail. This is because some valuable ideas really do come as "blinding flashes of inspiration" that cost nothing at all. In these cases, our proxy will overprice R&D costs. This is a fundamental problem, though not nearly as bad as under-rewarding R&D so that innovation never occurs at all.

*Limitations.* Using cleverness as a cost proxy has at least three drawbacks. First, R&D costs vary from industry to industry. This means that the PHOSITA standard cannot be used to compare costs across industries. Patent law has long acknowledged this limitation.[159] Second, the standard can also vary for different types of research within the same industry. This problem is particularly noticeable for pharmaceuticals, where costs usually increase several-fold at each step along the R&D "pipeline."[160] Third, the PHOSITA test only provides information on *relative* costs. While we know that some projects are costlier than others, we cannot say by how much. Finally, R&D costs are unpredictable. This suggests that the *ex post* cost of R&D programs may often be larger or smaller than inventors anticipate. This should not matter in most cases. This is because innovation economics almost always depends on *ex ante* incentives.

### *C. Uncertain Outcomes*

Identifying the PHOSITA with cost is only half the battle. An economically efficient patent law must also be flexible. In order to implement our third policy goal, for instance, courts must be able to allocate revenue in all possible ratios, from 100:0 to 0:100. On the face of things, the patent doctrine offers just four possible outcomes: 0:0 (both patents invalid), 50:50 (both patents upheld), 100:0 (first patent valid, second invalid), or 0:100 (first patent invalid, second valid).[161] Like the proverbial stopped clock, it seems that doctrine can only be right by accident.

Fortunately, closer inspection reveals a loophole. As Professors Mark Lemley and Carl Shapiro have emphasized, real inventors make their investment decisions

---

[158] Peter Buffett, *Chance Favors the Prepared Mind*, FORBES.COM (May 21, 2010, 2:48 PM), http://www.forbes.com/2010/05/21/life-luck-work-success-opinions-book-excerpts-peter-buffett.html.

[159] Burk & Lemley, *Policy Levers*, *supra* note 16, at 1650. Professors Burk and Lemley said:

> As the name suggests, PHOSITA-based analysis is specific to the particular art in which the invention is made. Courts measure most significant patent law doctrines against a benchmark that varies by industry. . . . Overwhelming evidence indicates that the application of the PHOSITA standard varies by industry, leading for example to fewer, but broader, software patents and more, but narrower, biotechnology patents.

*Id.*

[160] Published estimates for each stage of the process are collected in STEPHEN M. MAURER, THE RIGHT TOOL(S): DESIGNING COST-EFFECTIVE STRATEGIES FOR NEGLECTED DISEASE RESEARCH, (2005) (commissioned by WHO). *See also* O'Driscoll, *supra* note 157, at 2.

[161] *See* discussion *infra* Part VI.C.

*before* litigation based on "expected value."[162] In theory, at least, these probabilities can take on any value from 1 to 100. This means that suitably chosen probabilities can generate whatever allocations our Absolute Scope Rule requires. From an *ex ante* incentives standpoint, patent law is more flexible than it looks.

How patent law uses this flexibility is another matter. Plainly, the actual outcomes depend on doctrine. In Part VI, we show how doctrine translates the inherent uncertainty in jurors' PHOSITA-X determinations to allocate reward among successive inventors.

### *D. Lessons for Doctrine*

Any doctrine that uses the PHOSITA concept as a cost proxy is bound to make errors. These errors can be roughly divided into (a) the way the concept is defined and (b) its lack of predictability from one jury to the next. These simple observations already provide useful insights for doctrine.

*Definitional Limits.* There are many things that the PHOSITA concept cannot do even in principle. We have argued that sequential innovation is most efficient under an Absolute Scope Rule where each inventor (a) retains enough revenue to cover her actual R&D cost, and then (b) pays any excess to her immediate predecessor. But a jury's PHOSITA-X determination does not measure actual R&D costs. It only provides a relative judgment that some inventions are cleverer—and, implicitly, more expensive to develop—than others. It follows that no PHOSITA based doctrine can ever say when a particular inventor has covered her costs and should transfer any remaining revenue upstream. This limitation deters investment by increasing the odds that at least one inventor will fail to recover her costs. For this reason, PHOSITA-based incentives can never be as efficient as our ideal Absolute Scope Rule.

*Uncertainty.* Lawyers are trained to seek clear and replicable rules. However, we have argued that the PHOSITA proxy works best when jury determinations are uncertain. This is obviously troubling. At the same time, we should also know when to admit defeat. The Federal Circuit has tried and failed to make patent outcomes predictable for over thirty years now.[163] In these circumstances, it may be wiser to admit that the problem will never be solved and look for silver linings.

The good news is that unpredictability has only a minor impact on economic efficiency. If investors are risk-neutral,[164] incentives only depend on average

---

[162] Mark A. Lemley & Carl Shapiro, *Probabilistic Patents*, 19 J. ECON. PERSP. 75, 77–78 (2005). Professors Green and Scotchmer also examine models in which patent validity is probabilistic. Green & Scotchmer, *supra* note 62, at 22–32. The main difference is that their model focuses on uncertainty in R&D outcomes rather than litigation. *Id.*

[163] *See, e.g.*, *Phillips v. AWH Corp.*, 415 F.3d 1303, 1330 (Fed. Cir. 2005) (en banc) (Mayer, J., dissenting). As Judges Mayer and Newman protested, "In the name of uniformity, [we have] held that claim construction does not involve subsidiary or underlying questions of fact . . . . What we have wrought, instead, is the substitution of a black box, as it so pejoratively has been said of the jury, with the black hole of this court." *Id.*

[164] Rebecca S. Rudnick, *Enforcing the Fundamental Premises of Partnership Taxation*, 22 HOFSTRA L. REV. 229, 284 (1993) ("Risk neutrality is the condition where, given the choice of a

expected reward. If investors are risk averse, on the other hand, they can readily manage the problem by investing in so many companies (and implicitly, patent portfolios) that any fluctuations cancel.

The main costs of uncertainty involve individual justice, *i.e.* the certainty that some inventors will receive windfalls while others are penalized. Still, even this objection is weaker than it sounds. First, everyone knows that lawsuits are uncertain, and patent owners necessarily accept this risk.[165] Second, we have already said that investors can defend themselves by diversifying. Finally, the actual number of cases that go to trial in any well-functioning system is small.[166] This means that most rewards will be paid in the form of licensing fees rather than court judgments. Theory suggests that most of these payouts will cluster somewhere near the average litigated outcome.[167]

*Reforming Doctrine.* This Section began by remarking that the PHOSITA doctrine is considerably more complicated than our simple PHOSITA-X discussion suggests. Some of these additional rules are designed to make the jury inquiry more routinized and predictable[168] and would presumably improve our proxy. However, we have also seen that the PHOSITA doctrine has become encrusted with "reasonable man"-type requirements that have nothing to do with actual inventors. Judges who take our arguments seriously should want to clear away this underbrush.

In the meantime, the PHOSITA standard continues to evolve. By far the most important development in recent years has been the Supreme Court's *KSR* decision, which adds "ordinary creativity" to the PHOSITA's attributes.[169] Previously, the PHOSITA had been more or less incapable of having new ideas. Instead, she was limited to teachings, suggestions, or motivations ("TSM") that already existed—and could be documented—in the prior art.[170] The *KSR* decision expanded the TSM universe to include trivial ideas that no one had previously bothered to articulate.[171] This gave defendants many more chances to assert obviousness.[172]

---

certain return of $100, and an investment which has a 50 percent probability of a return of $200, a risk neutral investor would be indifferent.").

[165] *See, e.g., Patent Litigation-Risks and Rewards*, NAT'L L. REV. (Oct. 25, 2010), http://www.natlawreview.com/printpdf/2900.

[166] Jay P. Kesan & Gwendolyn G. Ball, *How are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 244–45 (2006).

[167] Theorists traditionally assume that parties who know the likely litigation outcome will settle to avoid unnecessary transactions costs. The range of bargained outcomes depends on the size of those costs. The theory is potentially testable because it predicts that the parties will only litigate close questions. George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 26 (1984). Unfortunately, the prediction is highly fact-dependent and may not hold in practice. *See* Keith N. Hylton, *Asymmetric Information and the Selection of Disputes for Litigation*, 22 J. LEGAL STUD. 187, 187–88 (1993); Steven Shavell, *Any Frequency of Plaintiff Victory at Trial is Possible,* 25 J. LEGAL STUD. 493, 494 (1996).

[168] For a comprehensive survey, see Phillips v. AWH Corp.*,* 415 F.3d 1303, 1330–35 (Fed. Cir. 2005) (en banc) (Mayer, J., dissenting).

[169] KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 421 (2007).

[170] *Id.* at 399.

[171] *Id.* at 400–04.

[172] *See, e.g.*, Amy L. Landers, *Ordinary Creativity in Patent Law: The Artist Within the Scientist*, 75 MO. L. REV. 1, 73 (2010); Andrew B. Dzeguze, *The Devil in the Details: A Critique of*

The Supreme Court justified the PHOSITA's new attribute based on our first policy goal of declining to reward inventions that would arrive anyhow "in the ordinary course," of events.[173] Our economic analysis confirms and clarifies this argument. We have already argued that the *Graham* test is best understood as a proxy for the inventor's R&D costs, and that creative ideas are often less costly than other R&D inputs.[174] But this leads to an obvious problem: if the optimal threshold for average research inputs is PHOSITA-50, then the test for unusually cheap inputs, ideas, must be higher, for example, PHOSITA-70 or PHOSITA-80. The problem with the old TSM rule is that it was inconsistent in the wrong direction, if an idea could not be found in the prior art, it was automatically deemed creative. This implied something like a PHOSITA-0 standard for creativity. Whatever its flaws, *KSR* has almost certainly made the PHOSITA a better cost proxy.

## V. HOW DOCTRINE ALLOCATES REWARD (A): LITERAL INFRINGEMENT AND BLOCKING PATENTS

We have argued that the PHOSITA concept supplies a reasonable proxy for R&D costs. However, doctrine must still provide rules that translate jurors' PHOSITA determinations into expected rewards for 1G and 2G inventors. U.S. patent law currently does this through a two-step process. First, the jury must find that the 1G patent is valid, *i.e.* satisfies *Graham*'s "non-obviousness" test.[175] Second, the jury must find that the 2G invention infringes, *i.e.* falls within the 1G patent's "breadth." Strangely, this second inquiry is governed by two separate sets of rules – "Literal Infringement"[176] and the "Doctrine of Equivalents."[177] This Section analyzes the simplest Literal Infringement case. This typically occurs where the 1G patent literally describes all or part of the 2G invention. We discuss more complex doctrinal problems involving non-literal infringement and affirmative defenses in Parts VII and VIII.

---

KSR*'s Unwarranted Reinterpretation of "Person Having Ordinary Skill,"* 10 COLUM. SCI. & TECH. L. REV. 1, 46–48 (2009). Dzeguze objects that "creativity" and "skill in the art" are two distinct concepts, so that a mechanic could easily possess the latter while lacking the former. *Id.* The present article solves this difficulty by identifying the PHOSITA with actual inventors. Indeed, requiring both "average creativity" and "average skill" is no more inconsistent than asking what a person of "average height and weight" looks like.

[173] *KSR*, 550 U.S. at 417.

[174] *See supra* Part V.B.

[175] Graham v. John Deere Co., 383 U.S. 1, 17 (1966).

[176] Jurgens v. McKasy, 927 F.2d 1552, 1560 (Fed. Cir. 1991) (Literal infringement occurs when "all claim limitations are present in the accused device exactly [so that the] claims 'read on' the accused device.").

[177] Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 732 (2002) ("The scope of a patent is not limited to its literal terms but instead embraces all equivalents to the claims described"); Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 35 (1997).

### *A. Literal Infringement and Blocking Patents*

The boundary between Literal Infringement and the Doctrine of Equivalents runs like a fissure through patent doctrine.[178] As the Supreme Court's *Graver Tank* decision explains, "resort must be had in the first instance to the words of the claim. If accused matter falls clearly within the claim, infringement is made out and that is the end of it."[179] This process necessarily excludes any opportunity to present evidence of the 2G invention's cleverness.[180]

Literal Infringement frequently leads to "Blocking Patent" scenarios in which anyone practicing the 2G invention must obtain licenses from *both* the 1G and 2G patent owners.[181] Blocking Patents occur whenever both inventions are patentable and the second patent "includes the first."[182] This usually involves situations where the 1G and 2G inventions are both patentable and a complete description of the 2G invention includes (a) all of the underlying 1G invention's original claims elements, and (b) additional claims elements supplied by the 2G inventor.[183] Conversely, 2G inventions that modify one or more elements of the 1G patent escape literal infringement and are analyzed under the Doctrine of Equivalents.[184]

### *B. What Incentives Do Investors See?*

We now ask how the *Graham* standard allocates patent revenue between successive innovators. This Section introduces the basic logic by analyzing a simple scenario in which a 2G inventor patents improvements to a previously patented 1G

---

[178] *See generally* 5B CHISUM, *supra* note 113, § 18.04[4][a] ("Federal Circuit [cases] regularly refer to 'literal infringement' as one of the two species of infringement, the other being infringement under the doctrine of equivalents.").

[179] Graver Tank & Mfg. Co. v. Linde Air Prod. Co., 339 U.S. 605, 607 (1950).

[180] For example, evidence showing that the 2G invention has received a patent is inadmissible in literal infringement cases. See, for example, 5B CHISUM, *supra* note 113, § 18 and cases cited therein.

[181] Robert P. Merges, *Brief Note on Blocking Patents and Reverse Equivalents: Biotechnology as an Example, A*, 73 J. PAT. OFF. SOC'Y 878, 878–79 (1991).

[182] Cantrell v. Wallick, 117 U.S. 689, 694 (1886).

[183] Atlas Powder Co. v. E.I. Du Pont de Nemours & Co., 750 F.2d 1569, 1580 (Fed. Cir. 1984) ("[I]f Atlas patents A + B + C and Du Pont then patents the improvement A + B + C + D, Du Pont is liable to Atlas for any manufacture, use, or sale of A + B + C + D."). Chisum similarly explains:

> The patentability of an accused product or process often stems from what may be characterized as an additive or selective improvement, adding elements, features, or functions to the earlier patented invention or discovering preferred species for generic elements in the earlier invention, rather than from a mere substitution of one element for another. In such cases, Federal Circuit decisions recognize that an accused product or process does not avoid infringement by adding functions or features if it contains literally or by equivalents all the elements of patent claim in question.

Donald S. Chisum, *The Scope of Protection for Patents After the Supreme Court's* Warner-Jenkinson *Decision: Fair Protection--Certainty Conundrum*, 14 SANTA CLARA COMPUTER & HIGH TECH. L.J. 1, 36 (1998).

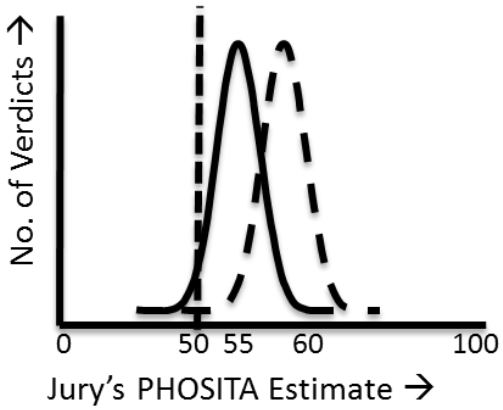
[184] Jurgens v. McKasy, 927 F.2d 1552, 1560 (Fed. Cir. 1991).

technology. For simplicity, we assume that the 1G product is pre-commercial and earns no revenue. The 1G inventor sues for infringement and each party argues that the other's patent is invalid under *Graham*'s non-obviousness test.

*Analysis.* We begin by reviewing the possible legal outcomes. The *Graham* "non-obviousness" test tells juries to (a) determine each invention's PHOSITA-X rank, and (b) invalidate patents that do not exceed PHOSITA-50. Because invalid patents receive nothing, our hypothetical lawsuit can end in 0:0, 100:0, and 0:100 allocations.[185] If both patents are upheld, our blocking patents assumption produces 50:50 sharing.

In order to decide the case, the jury must determine PHOSITA-X values for both the 1G and 2G patents. We start with the 1G patent. In theory, at least, we can imagine empaneling 1,000 different juries to determine the 1G patent's PHOSITA-X value. Suppose that we do this and the resulting verdicts have, say, a mean of fifty-five. Since all 1,000 juries have seen the same evidence and received the same instructions, it is reasonable to think that most juries will reach similar results. This suggests that juries will pick rankings close to fifty-five (e.g. fifty-eight) more often than distant ones (e.g. seventy-five). The solid line in Figure 1 depicts a typical example of this "verdict distribution." Loosely speaking, readers can think of the vertical axis as showing the probability that a jury will determine each rank from 1 to 100. [186]

**FIGURE 1**




[185] An invention which is obvious cannot be patented and therefore receives no revenue. In our 0:0 case both patents are obvious and invalid and therefore receive nothing. Our 100:0 and 0:100 cases occur when only one of the two patents is obvious and invalid.

[186] Mathematically-inclined readers will recognize that one cannot really speak of "the probability of x" for continuous values. Properly speaking, Fig. 1 depicts a "probability density."

Different inventions will usually have different verdict distributions.[187] How does this affect the allocation of reward? We know that the expected allocation is determined by the ratio of how often the 1G invention is upheld to how often the 2G invention is upheld. Geometrically, this corresponds to how much of the area under each verdict distribution falls to the right of *Graham*'s PHOSITA-50 standard. On the other hand, a PHOSITA-60 invention (Figure 1, dashed line) will usually have more area to the right of the PHOSITA-50 threshold than a PHOSITA-55 invention (Figure 1, solid line). More clever inventions will therefore be upheld more often, and have greater expected value, than less clever ones. This is encouraging.

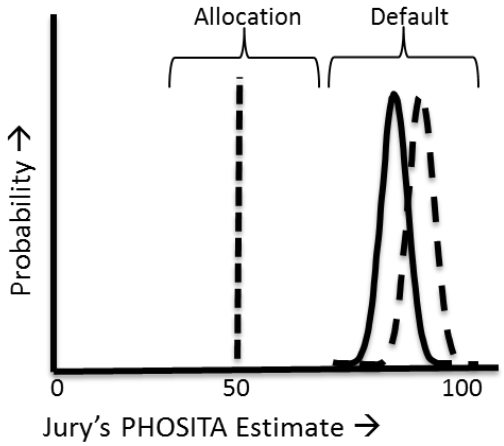
As usual, the devil is in the details. Ideally, we would like to use our verdict distributions to allocate reward according to cost.[188] But this requires proportionality so that a one percent increase in cleverness improves an invention's chances of receiving revenue by roughly one percent. In practice, this ideal is never reached. A glance at Figure 1 shows why. Our example started with a verdict distribution that peaked near *Graham*'s PHOSITA-50 threshold standard. Because the distribution is so tall at this point, even a slight increase in cleverness drags a large slice of area into the right hand side. In this situation, the patent system tends to over-reward small cleverness increments. On the other hand, suppose that distribution's peak is far to the right of the threshold. Here, a one percent increase in cleverness generates too little expected reward. Figure 2 depicts the extreme case where both inventions are extremely clever. Here, the chances that the jury's PHOSITA-X determination will be less than fifty are already tiny. For this reason, further increases in cleverness have almost no impact on expected revenue. Instead, our 50:50 blocking patents default rule dictates the allocation.

---

[187] If our legal rules were completely determinate, all 1,000 juries would return a verdict of 55. In this case, Figure 1 would show an infinitely narrow verdict distribution, i.e., a vertical line. More realistically, we expect real verdicts to contain a random component as well. Where this component is small, we expect juries to deliver "mostly determinate" verdicts. This will produce narrow verdict distributions with high peaks. Where the random component is larger, we expect verdict distributions to be broader and flatter.

[188] We have already said that this prescription follows the neoclassical economics literature. It is worth noting, however, that some legal scholars suggest that reward should reflect value to consumers. *See, e.g.*, Amy L. Landers, *Patent Claim Apportionment, Patentee Injury, and Sequential Invention*, 19 GEO. MASON L. REV. 471, 472 (2012). For them, tying reward allocation to the PHOSITA cost proxy is a disaster. The problem could presumably be fixed by replacing *Graham*'s PHOSITA with an "Ordinary Consumer" who defines average utility. U.S. law has repeatedly rejected such proposals. Risch, *Reinventing Usefulness*, *supra* note 108, at 1236–37.

**FIGURE 2**




For convenience, we will refer to regions close to the PHOSITA-50 as the "Allocation Zone" in what follows. Conversely, we will refer to regions dominated by the 50:50 rule as the "Default Zone."

Putting these results together, we distinguish several cases. First, both patents may fall inside the Allocation Zone. Here, the expected allocations really do depend on comparative cleverness. Second, both patents may fall far to the right of the PHOSITA-50 threshold. Within this Default Zone, we expect a Blocking Patents Rule to enforce 50:50 sharing even when one patent is significantly cleverer than the other. Finally, readers should consider the case (not shown) where one patent falls in the Default Zone and the other falls inside the Allocation Zone. In this case, the final result will depend almost entirely on small differences in the Allocation Zone patent's cleverness while ignoring similar nuances in the Default Zone patent.

Finally, we have drawn Figs. 1 and 2 so that they feature narrow verdict distributions with high, sharp peaks. For large uncertainties, the verdict distributions become so broad that the Default Zone disappears entirely.

### *C. Which World Do We Live In?*

Our argument suggests that—far from being a flaw—litigation uncertainty is not necessarily a flaw because it provides an essential bridge between patent law's formal winner-take-all outcomes and the need for practical incentives that promise adequate reward to each inventor in the R&D chain. Clearly, much depends on the shape of the verdict distributions. If they are broad, we should expect large Allocation Zones.[189] In this case, the *Graham* non-obviousness standard can be

[189] *Supra* Part VI.B.

trusted to link expected rewards to relative R&D costs for most inventions[190] so that *Graham*'s basic PHOSITA standard provides reasonable allocations. Conversely, narrow distributions imply large Default Zones.[191] In this situation, the *Graham* standard often functions poorly by producing 50:50 allocations when one invention is markedly cleverer than the other.[192] Here, legislators and judges will often decide that it is better to create additional Allocation Zones which reduce or eliminate the number of default outcomes. This typically involves adopting affirmative defenses that implement additional PHOSITA-X tests for highly clever inventions.[193]

Which world do we actually live in? In principle, the verdict distribution can be measured. For example, scholars could play the same videotaped patent trial to 100 average Americans and ask them to deliver verdicts. This experiment, of course, has never been done. Still, we can guess what the verdicts would be. Given that we are all prospective jurors, this might involve little more than asking readers how consistently they could detect the difference between, say, a PHOSITA-75 invention and its PHOSITA-50 counterpart.

In what follows, we will assume moderate verdict distributions that leave substantial Default Zones. This provides a strong argument for creating affirmative defenses that invite juries to conduct additional comparative cleverness determinations far from PHOSITA-50.

### *D. Implications for Doctrine*

Our Literal Infringement example is a special case. Nonetheless, it contains important lessons for doctrine. First, it reminds us of unintended consequences. Recall that the Supreme Court invented the *Graham* standard to address our first policy goal, i.e., screen out inventions that are not worth the "embarrassment of an exclusive patent."[194] In the process, however, the Court inadvertently created a process for allocating reward among successive inventors. Though facially reasonable, this rule falls short of our benchmark Absolute Scope rule. We will see that this defect can be ameliorated—though not eliminated—by grafting additional PHOSITA-X inquiries onto the doctrine.

Finally, we have so far assumed that litigation is a single one-shot event. But many patents face multiple litigations.[195] Under principles of collateral estoppel this means that a patent which fails to establish *Graham* non-obviousness in, for example, the third litigation will lose *every* subsequent lawsuit.[196] More generally, patent owners facing multiple litigations will receive systematically less expected income than they would if their claims were tried in a single action. It is difficult to see how this problem can be solved short of abolishing collateral estoppel altogether.

---

[190] *See supra* Part VI.B.

[191] *See supra* Part VI.B.

[192] *Graham*, 383 U.S. at 17.

[193] *Infra* Part VIII.

[194] Graham v. John Deere Co., 383 U.S. 1, 9–10 (1966).

[195] *See, e.g.*, *Collateral Estoppel Bars UpJohn's Patent Suit*, 14 ANDREWS PHARMACEUTICAL LITIG. REP. 15, 15 (1999) (Upjohn patent faced two earlier lawsuits).

[196] *See id.* (invoking collateral estoppel to bar re-litigation of validity issues).

## VI. HOW DOCTRINE ALLOCATES REWARD (B): THE DOCTRINE OF EQUIVALENTS

Blocking Patents do a surprisingly good job of allocating reward among successive inventors. If our third policy goal were the only one, we would expect the same 50:50 rule to operate throughout patent law. However, we must also consider our second goal. It insists that the 1G patent monopoly should have limited duration and breadth. This implies that sufficiently clever 2G inventors should sometimes be allowed to build on 1G insights and still escape infringement.

This Section extends our probabilistic patents analysis to include the Doctrine of Equivalents. Part VII.A sets the stage by reviewing the sometimes incoherent case law that surrounds the Doctrine of Equivalents. We argue that attempts to rationalize the Doctrine will almost certainly require juries to assess the 2G invention's cleverness. Part VII.B analyzes how a rationalized Doctrine of Equivalents would allocate reward for different 2G cleverness standards. Significantly, we find that very strong versions of the Doctrine of Equivalents approximate our Blocking Patents rule. Part VII.C concludes by discussing the prospects for combining Literal Infringement and the Doctrine of Equivalents into a single unified doctrine.

### *A. The Doctrine of Equivalents*

The Doctrine of Equivalents was originally designed to prevent copyists from using trivial changes to evade verbal "metes and bounds" descriptions of the patent monopoly.[197] This eventually led to a "Function-Way-Result" Test which asked whether the accused device "perform [sic] substantially the same function in substantially the same way to achieve substantially the same result."[198] But how should judges define "substantially"? If the answer is more than *de minimis*, then any changes beyond trivial and calculated evasion should escape infringement. Precedents that describe the Doctrine of Equivalents as a deterrent against "pirating," "mere imitation," "insubstantiality,"[199] "minor variations," and "unimportant and insubstantial changes"[200] tend to reinforce this view. More recently, however, the Supreme Court has emphasized that this doctrine is not limited to cases of copying and piracy but instead defines "the notion of identity between a patented invention and its equivalent."[201] This seems to recognize expansive, although sometimes vague, judicial statements that 2G inventors cannot escape so long as "the gist" of their invention is the same and there are no

---

[197] *See* Kenneth D. Bassinger, *Unsettled Expectations in Patent Law:* Festo *and the Moving Target of Claim Equivalence*, 48 HOW. L.J. 685, 691 (2005) ("Recognizing that language is often a dull instrument with which to define the intellectual nuances of an invention, the courts created the doctrine of equivalents to provide an escape from the confines of the restrictive literalism of claim language.").

[198] *Id.* at 694.

[199] Bassinger, *supra* note 197, at 691, 695.

[200] Alan L. Durham, "Patent Symmetry," 87 B.U.L. REV. 969, 1004-05 (2007).[201] Warner Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 34–35 (1997).

[201] Warner Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 34–35 (1997).

"unexpected or substantially different results.".[202] These standards bring the Doctrine of Equivalents much closer to innovation theory's familiar concept of patent breadth.

That said, the content of words like "equivalent" and "gist" is elusive. Historically, most efforts to clarify the Doctrine of Equivalents have involved arguments that modifications which use an "interchangeable" element cannot be substantial and are therefore equivalent.[203] "Interchangeability," in turn, is almost always framed in terms of the expectations of a skilled artisan.[204] This implies a cleverness test and bears an obvious resemblance to *Graham's* PHOSITA standard. At the same time, the "skilled artisan" phraseology predates and remains distinct from the usual formula of a "Person Having Ordinary Skill in the Art." Maddeningly, the nature of these differences is obscure. On the one hand, assertions that the notional artisan must have "reasonable skill" hint that the Doctrine of Equivalents' cleverness threshold could be lower than the PHOSITA's "average" or "ordinary" test.[205] On the other hand, statements that the "advancement" must be more than "routine"[206] point in the opposite direction.

Suggestions for reforming the Doctrine of Equivalents usually fall into three categories. [207] The first would leave the Function-Way-Result Test unchanged but demand more explicit argument and evidence.[208] This would presumably weaken the Doctrine while avoiding any policy discussion of what patent breadth is or ought to

---

[202] *See* Atlas Powder Co. v. E.I. Du Pont de Nemours & Co., 750 F.2d 1569, 1580 n.3 (Fed. Cir. 1984).

[203] Frustratingly, the test is not absolute. *See, e.g.*, Chiuminatta Concrete Concepts, Inc. v. Cardinal Indus., Inc., 145 F.3d 1303, 1309 (Fed. Cir. 1998) ("Moreover, a finding of known interchangeability, while an important factor in determining equivalence, is certainly not dispositive."); *see also* Hon. Paul R. Michel, *The Role and Responsibility of Patent Attorneys in Improving the Doctrine of Equivalents*, 40 IDEA 123, 128 (2000) (remarking that "known interchangeability" is currently a mere factor instead of "a true test").

[204] There are several cases that invoke the concept of a skilled artisan without using the traditional "Person Having Ordinary Skill in the Art" standard. *Graver Tank*, 339 U.S. at 609 ("An important factor is whether a person reasonably skilled in the art would have known of the interchangeability of an ingredient not contained in the patent with one that was."); Interactive Pictures Corp. v. Infinite Pictures, Inc., 274 F.3d 1371, 1383 (Fed. Cir. 2001) ("[T]he known interchangeability test looks to the knowledge of a skilled artisan to see whether that artisan would contemplate the interchange as a design choice."). Commentators routinely paraphrase these cases as invoking a PHOSITA standard. *See, e.g.*, Bassinger, *supra* note 197, at 695. Durham notes that the Supreme Court has variously invoked "persons skilled in the art" and "skilled practitioners" and also asked whether the 2G substitution was "routine" or else required "a further advancement in the art." Durham, *supra* note 200, at 1011–12.

[205] *See* Durham, *supra* note 200, at 1011.

[206] *Id.*

[207] *See* Michel, *supra* note 203, at 124–25; *see also* M. Scott Boone, *Defining and Refining the Doctrine of Equivalents: Notice and Prior Art, Language and Fraud,* 43 IDEA 645, 646 (2003).

[208] *See* Michel, *supra* note 203, at 129. According to Judge Michel, the approach would build on *Malta v. Schulmerich Carillons, Inc.*, 952 F.2d 1320, 1327 (Fed. Cir. 1991) (criticizing "offhand and conclusory statements" that left jury "to its own imagination on the technical issue of equivalency") and *Lear Siegler, Inc. v. Sealy Mattress Co. of Mich., Inc.*, 873 F.2d 1422, 1426 (Fed. Cir. 1989) (stating that the jury must receive "particularized testimony and linking argument" for each element of the Triple Identity Test).

be.[209] The second proposal would expand "known interchangeability" into a "true test."[210] While the details of this proposal are obscure, it is reasonable to think that the word "known" would continue to be measured by the awareness of skilled artisans. This clearly implies a PHOSITA-X threshold test, though not necessarily a PHOSITA-50 test. The final and arguably most popular suggestion would be to identify breadth with *Graham*'s non-obviousness standard. This approach is already widely used overseas,[211] and has been endorsed by various U.S. commentators[212] and judges.[213] As the late Judge Nies argued in a celebrated concurrence, "[a] substitution in a patented invention cannot be both nonobvious and insubstantial."[214] Significantly, Judge Nies may not have realized that *Graham*'s PHOSITA-50 standard is just one choice among many. In fact, her argument applies to any PHOSITA-X value greater than PHOSITA-50.[215]

### *B. Analysis*

Given these uncertain precedents, it is hard to predict how judges will ultimately reform and clarify the Doctrine of Equivalents. At the same time, practically all courts invoke PHOSITA-like cleverness judgments. No reform is likely to change this. The only question is whether the corresponding PHOSITA-X threshold will be large or small. This section examines how the various possible choices change reward allocations compared to *Graham*'s baseline non-obviousness test.

We begin by reminding the reader of our Literal Infringement/Blocking Patents analysis. In that case, doctrine asked juries whether the 1G and 2G patents were non-obvious. This meant that both patents were compared against the same PHOSITA-50 standard and, through that standard, each other. The Doctrine of Equivalents adds a second test over and above the basic *Graham* non-obviousness inquiry. Now, jurors must also compare the 2G invention's cleverness against the 1G patent's breadth. Doctrinally, this second breadth threshold can potentially take on any PHOSITA-X value from 1 to 100. For reasons that appear below, it is enough to consider two examples.

*Narrow Breadth.* Figure 3 shows how much reward the 2G invention receives where breadth is defined by a PHOSITA-25 standard. Unlike previous Figures, we plot the 2G invention's cleverness against its expected allocation share. For clarity,

---

[209] More explicit evidence and argument by counsel might, of course, lead to clearer judicial decisions.

[210] Michel, *supra* note 203, at 129; *see also* Boone, *supra* note 207, at 654.

[211] *See* Mark D. Janis, *Who's Afraid of Functional Claims? Reforming the Patent Law's § 112, ¶ 6 Jurisprudence*, 15 SANTA CLARA COMPUTER & HIGH TECH. L.J. 231, 267 (1999).

[212] *See* Durham, *supra* note 200, at 982; Scott P. Zimmerman, *The Doctrine of Equivalents: A Call for Congressional Reinvigoration*, 40 IDEA 599, 623 (2000); Michael T. Siekman, *The Expanded Hypothetical Claim Test: A Better Test for Infringement for Biotechnology Patents Under the Doctrine of Equivalents*, 2 B.U. J. SCI. & TECH. L. 6, 10 (1996).

[213] *See* Michel, *supra* note 203, at 128–29; Roton Barrier, Inc. v. Stanley Works, 79 F.3d 1112, 1128 (Fed Cir. 1996) (Nies, J., concurring).

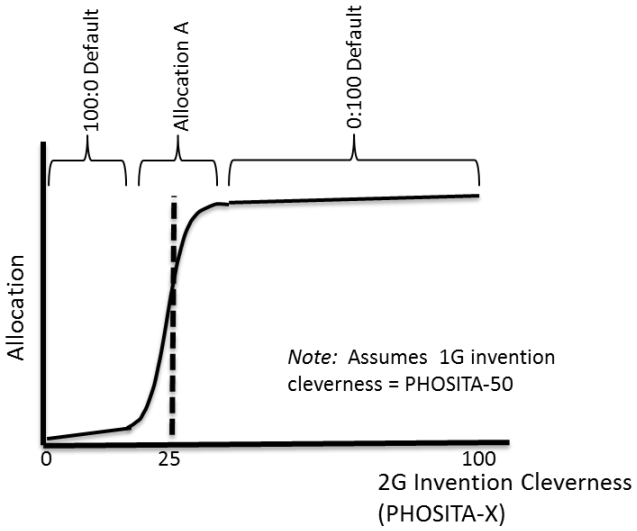
[214] *Roton Barrier*, 79 F.3d at 1128.

[215] *See supra* note 152 and accompanying text.

the Figure illustrates the specific case in which the 1G invention's verdict distribution peaks at PHOSITA-50.

We start on the far left-hand side. For very low cleverness inventions the 2G patent has almost no chance of exceeding the PHOSITA-25 threshold. For this reason, it earns nothing in expectation. This leads to a 100:0 Default Zone.[216] Now examine the far right-hand side. Here, inventions that are significantly cleverer than PHOSITA-25 escape the 1G patent and pay nothing. This implements our second policy goal but also leads to very extreme 0:100 allocations that fail to reflect the parties' relative R&D costs. This region is shown as the "0:100 Default Zone" in Figure 3.

**FIGURE 3**




Finally, consider the middle region near PHOSITA-25. Unlike our blocking patents case, the 1G patent is measured against *Graham*'s PHOSITA-50 standard while the 2G patent is measured against our PHOSITA-25 breadth definition. This distorts the calculation so that the 2G inventor receives roughly three times as much for her cleverness as the 1G inventor.[217] This allocation can only be right for rare cases where R&D costs rise steeply at each successive development step.

---

[216] The label is a matter of convenience. Because the 1G patent has a 50 percent chance of failing the non-obviousness test, 0:0 outcomes are equally likely.

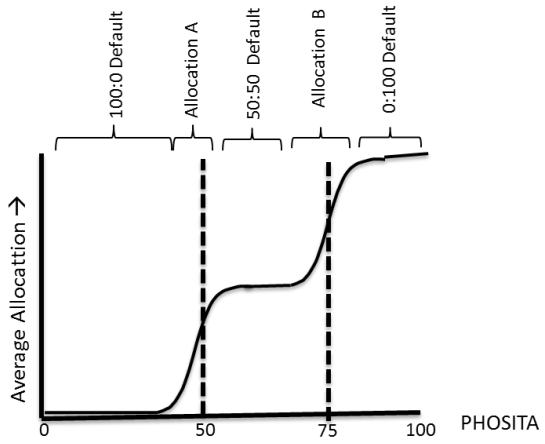
[217] To see this, consider the case where the 1G invention's verdict distribution peaks at PHOSITA-50 and the 2G curve peaks at PHOSITA-25. Then we expect the 1G patent to be upheld 50% of the time. Even in these cases, however, there is still a 50% chance that the 2G invention will exceed PHOSITA-25 and escape infringement. We therefore expect the final allocation to be 25:75 on average.

So far, we have assumed that the 1G invention's verdict distribution peaks at PHOSITA-50. If the 1G patent is significantly cleverer than the *Graham* non-obviousness juries will almost certainly find it valid. In this case, the 2G invention's expected allocation share will be smaller at each point in Figure 3. Similarly, less clever 1G invention will produce higher 2G rewards.

*High PHOSITA-X.* We now consider the opposite case in which most follow-on inventions infringe the 1G patent. Figure 4 depicts this situation for the PHOSITA-75 case. As in Figure 3, Figure 4 depicts the specific case where the 1G invention's verdict distribution peaks at PHOSITA-50.

Not surprisingly, everything to the left of the PHOSITA-75 standard resembles our Blocking Patent analysis in Part VI. We start by examining the left-hand side where the 2G invention is markedly less clever than the PHOSITA-50 threshold needed to create a blocking patent. Since the 2G invention receives nothing this produces a "100:0 Default Zone" in the figure. Now consider the region near the center of the Figure (Allocation Zone A) where the 2G invention is close to PHOSITA-50. Here, revenue is allocated in proportion to relative cleverness. There is also a second, 50:50 Default Zone for 2G inventions whose verdict distributions fall in the broad middle region between PHOSITA-50 and PHOSITA-75.[218]

**FIGURE 4**




[218] The zone disappears in the special case where breath and non-obviousness standards are identical. As Professor Adams remarks, combining the two standards sets up a syllogism. Charles W. Adams, *Allocating Patent Rights Between Earlier and Later Inventions*, 54 ST. LOUIS U. L.J. 55, 60 (2009). If the 1G patent would enable a PHOSITA to make the 2G invention, the latter falls within the 1G patent and should not be patentable. But if the 1G patent does not enable a PHOSITA to make the 2G invention, the latter is patentable and falls outside the 1G patent's breadth. This may explain why most blocking patent cases involve Literal Infringement and not the Doctrine of Equivalents.

Finally, the regions on the right resemble our Low Breadth case. These include a probabilistic region ("Allocation B") near the PHOSITA-75 threshold that defines non-infringing 2G improvements and a 0:100 region for ultra-clever 2G inventions. As before, making the 1G invention cleverer reduces the 2G invention's expected allocation at each point in the figure.

Putting these observations together, we see that everything to the left of the PHOSITA-75 threshold replicates our "Blocking Patents" allocation rule. For infinite (PHOSITA-100) breadth Literal Infringement and the Doctrine of Equivalents produce identical results.

### *C. Lessons for Doctrine*

The deep divide between Literal Infringement and The Doctrine of Equivalents outcomes is puzzling. Certainly, nothing in innovation economics suggests that inventors who add new elements require fewer incentives than those who rethink existing ones. But this leaves us with an uncomfortable choice. Either the two rules share some underlying unity, or one must surely wrong.

The fact that large PHOSITA-X versions of the Doctrine of Equivalents produce much the same allocation outcomes as Literal Infringement provides an important clue. If the two doctrines really do converge, the difference may be less about substance than judicial economy. After all, cleverness inquiries require expensive jury determinations. But how clever can a 2G invention be if it fails to escape the 1G patent's literal claims language? Judges could well decide that it is better to outlaw cleverness inquiries in Literal Infringement cases altogether. Like all bright line rules, this is bound to produce mistakes. Indeed, Part VIII.A will discuss one such example at length. Still, this may not matter much if defenses like the "Reverse Doctrine of Equivalents" provide a safety valve for radically clever improvements.[219]

The doctrinal fissure is harder to rationalize for low PHOSITA-X versions of the Doctrine of Equivalents. Here, Literal Infringement really does produce markedly different outcomes. In theory, Congress could re-unify doctrine by abolishing Literal Infringement and mandating a single Doctrine of Equivalents test for all patent disputes. But in that case, what PHOSITA-X value should Congress use to define breadth? In an ideal world, legislators could estimate the optimal PHOSITA-X thresholds for our second and our third goals separately and split the difference. In reality, this procedure would consist mostly of guesswork.

## VII. HOW DOCTRINE ALLOCATES REWARD (C): DEFENSES

We saw in Part VI that any patent doctrine based on the *Graham*'s PHOSITA test automatically establishes baseline rules for allocating reward among inventors. But we also showed that these results could be improved by adopting doctrinal architectures that invite juries to make additional inquiries. This Section reviews

---

[219] *See infra* Part VIII.A.

various infringement defenses that have been proposed over the years and compares their potential for improving *Graham*'s baseline allocation rule.

We begin by reviewing the often controversial defenses that courts have developed to address multi-inventor issues. These include the Reverse Doctrine of Equivalents (Part VIIIA), the Pioneer Patents Doctrine (Part VIII.B) and the Written Description Requirement (Part VIII.C). We then compare these doctrines against our third goal.

### *A. Reverse Doctrine of Equivalents*

The Reverse Doctrine of Equivalents was introduced more than 100 years ago[220] and remains (apparently[221]) good law today. However, it is expressly limited to literal infringement cases[222] and is seldom applied.[223] Despite this, it remains

---

[220] Boyden Power Brake Co. v. Westinghouse, 170 U.S. 537, 568 (1898).

[221] The *Westinghouse* rule languished for over half a century before the Supreme Court revived it in *Graver Tank & Mfg. Co. v. Linde Air Prods., Inc*., 339 U.S. 605, 608 (1950). Thereafter, it remained uncontroversial, if seldom used, for another fifty years. This changed in 2002 when the Federal Circuit suggested that the *Graver Tank* holding had been superseded by 35 U.S.C § 112 (2012). If so, the reverse doctrine could no longer be invoked to constrict the "literal language when it is clearly claimed." Tate Access Floors, Inc. v. Interface Architectural Res., Inc., 279 F.3d 1357, 1368 (Fed. Cir. 2002). However, this argument has been widely criticized. Mark A. Lemley, *The Changing Meaning of Patent Terms*, 104 MICH. L. REV. 101, 104 (2005) (arguing that Reverse Doctrine of Equivalents is distinct from written description because it is measured at the time of infringement rather than filing); Cyrill P. Rigamonti, *Conflicting Theories of Equivalence: 35 U.S.C. § 112, ¶ 6 in the Supreme Court and the Federal Circuit*, 40 IDEA 163, 193–94 (2000). *But see* Martin J. Adelman & Gary L Francione, *The Doctrine of Equivalents in Patent Law: Questions that* Pennwalt *Did Not Answer*, 137 U. PA. L. REV. 673, 724 n.221 (1989) (Section 112 "should be viewed as a legislative codification of the reverse doctrine of equivalents."). Moreover, more recent Federal Circuit decisions continue to mention the defense. *See* Biogen, Inc. v. Berlex Labs, Inc., 318 F.3d 1132, 1140 (Fed. Cir. 2003); Plant Genetic Sys., N.V. v. DeKalb Genetics Corp., 315 F.3d 1335, 1340–41 (Fed. Cir. 2003); Amgen, Inc. v. Hoechst Marion Roussel, Inc., 314 F.3d 1313, 1351 (Fed. Cir. 2003); DePuy Spine, Inc. v Medtonic Sofamor Daenk, Inc., 567 F.3d 1314, 1338–39 (Fed Cir. 2009) ("The Supreme Court has recognized it to be a viable defense, even if it is rarely asserted.").

[222] Tex. Instruments, Inc. v. U.S. Int'l Trade Comm'n, 846 F.2d 1369, 1371 (Fed Cir. 1988). We will see that the Pioneer Patents doctrine provides similar rules in Doctrine of Equivalents cases.

[223] *But see* Gardner v. Ford Motor Co., No. C85-711WD, 1990 U.S. Dist. LEXIS 21052, at *31–35 (W.D. Wash. 1999) (finding that the second invention was far removed from the "principle, structure, and operation" of claimed invention); U.S. Steel Corp. v. Phillips Petroleum, Co., 865 F.2d 1247, 1253 n.9 (Fed. Cir. 1989); Am. Standard, Inc. v. Pfizer, Inc., 722 F. Supp. 86, 104 (D. Del. 1989) (endorsing "principle"-based test and criticizing tests based on whether invention functions "in a substantially different way"); Smithkline Diagnostics, Inc. v. Helena Labs. Corp., 859 F.2d 878, 889 (Fed. Cir. 1988) (finding that the defendant could establish non-infringement by showing that his device performs a similar function in a substantially different way); Precision Metal Fabricators Inc. v. Jetstream Sys. Co., 693 F. Supp. 814, 819 (N.D. Cal. 1988) (holding that infringement did not occur where defendant's machines did "not operate on the same principle" and "[w]hatever similarities exist are incidental and do not enhance the operation of defendants' machines"); Leesona Corp. v. United States, 530 F.2d 896, 906 (Ct. Cl. 1976) ("It is well settled that more than a literal response to the terms of the claims must be shown to make out a case of infringement.").

surprisingly popular among scholars.[224] This suggests that it could be readily revived given a clear policy reason to do so.

Strangely, most of what we know about the Doctrine still comes from the original *Westinghouse* opinion. The case involves – and describes in prolix, obsessive detail – air brakes for trains.[225] Westinghouse had developed a system in which each car's brakes were driven by a local tank which was, in turn, pressurized and controlled by lines from the locomotive.[226] This required an elaborate system of "triple" and "auxiliary valves."[227] Boyden invented a new triple valve that made the auxiliary valves unnecessary. Despite remarkable cleverness, however, the new arrangement still fell within the literal language of Westinghouse's patent.[228] This might have encouraged the Court to declare a blocking patent. In fact, the Court went much further by declaring that Boyden's invention had "so far changed the principle of the device that the claims of the patent, literally construed, have ceased to represent his actual invention."[229] Left unsaid was how much change would be required to trigger the defense. Most modern scholarship[230] states that the change

---

[224] Extended discussions and proposed extensions of the doctrine can be found in, *inter alia*, R. Scott Roe, *Nanotechnology: When Making Something Smaller is Nonobvious*, 12 B.U. J. SCI. & TECH. L. 127 (2006); Carrier, *supra* note 35, at 118; Christopher D. Hazuka, *Supporting the Work of Lesser Geniuses: An Argument for Removing Obstructions to Human Embryonic Stem Cell Research*, 57 U. MIAMI L. REV. 157, 157 (2002); Anand Gupta, *Patent Law: The Supreme Court Reinforces the Validity of the Doctrine of Equivalents in* Warner-Jenkinson Co. v. Hilton Davis Chemical Co.," 23 S. ILL. U. L. J. 123, 123 (1998); Lemley, *Improvement in IP*, *supra* note 87, at 1010–13; Robert P. Merges, *Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents*, 62 TENN. L. REV. 75, 101–03 (1994); Michael S. Greenfield, *Recombinant DNA Technology: A Science Struggling with the Patent Law*, 44 STAN. L. REV. 1051, 1051 (1992); Laura A. Handley, *Refining the* Graver Tank *Analysis With Hypothetical Claims: A Biotechnology Exemplar*, 5 HARV. J. LAW & TECH. 31, 31 (1991).

[225] Boyden Power Brake Co. v. Westinghouse, 170 U.S. 537, 537–39 (1898).

[226] *Id.* at 545.

[227] *Id.* at 553, 558–59.

[228] *Id.* at 571–73.

[229] *Id.* at 568, 572. The Supreme Court famously reaffirmed the rule in 1950:

> Thus, where a device is so far changed in principle from a patented article that it performs the same or a similar function in a substantially different way, but nevertheless falls within the literal words of the claim, the doctrine of equivalents may be used to restrict the claim and defeat the patentee's action for infringement.

Graver Tank & Mfg. Co., Inc. v. Linde Air Prods. Co., 339 U.S. 605, 608–09 (1950).

[230] Peter Lee, *The Accession Insight and Patent Infringement Remedies*, 110 MICH. L. REV. 175, 185–86 (2011) (noting that radical improvements may avoid liability under literal infringement); Roe, *supra* note 224, at 141 (Doctrine prevents 1G patent owners from stifling "radical improvements."); Simone A. Rose, *On Purple Pills, Stem Cells, and Other Market Failures: A Case for a Limited Compulsory Licensing Scheme for Patent Property*, 48 HOW. L.J. 579, 584 (2005) ("Moreover, the reverse doctrine of equivalents allows for infringing acts, which produce radically pioneering inventions"); Hazuka, *supra* note 224, at 212; Janice M. Mueller, *No "Dilettante Affair": Rethinking the Experimental Use Exception to Patent Infringement for Biomedical Research Tools*, 76 WASH. L. REV. 1, 45 n.228 (2001) (Reverse Doctrine of Equivalents triggered by "radical improvement."); Maureen A. O'Rourke, *Toward a Doctrine of Fair Use in Patent Law*, 100 COLUM. L. REV. 1177, 1193 (2000) (Reverse Doctrine of Equivalents operates where 2G inventor is "radical improver" of underlying invention.); Robert P. Merges, *A Brief Note on Blocking Patents and the*

must be "radical." However, the matter is not free from doubt and some scholars have suggested that the 2G advance need only be radical compared to the underlying 1G invention.[231] Still others argue that mere "substantial" or "significant" changes suffice.[232] In terms of our present analysis, then, existing precedent would support any 2G cleverness standard ranging from "substantial" to "radical."[233]

The Court also suggested an alternative theory.[234] The outcome might have been different, the justices argued, if Westinghouse's method "would naturally have suggested the device" invented by Boyden.[235] But in fact, "the Westinghouse patent, *if [Boyden] had had it before him*, would scarcely have suggested the method he adopted to accomplish the same results. Under such circumstances, the law entitles him to the rights of an independent inventor."[236] This implies that a 2G innovator should be allowed to claim a defense where she "can truly assert that her claimed invention was not made possible (i.e., enabled) by what was described by the prior

---

*Reverse Doctrine of Equivalents in Biotechnology Cases*, 73 J. PAT & TRADEMARK OFF. SOC'Y 878, 883 (1991) (safety valve that prevents patent owners form stifling radical improvements).

[231] Alan Devlin, *The Misunderstood Function of Disclosure in Patent Law*, 23 HARV. J.L. & TECH. 401, 434 n.166 (2010) (Reverse Doctrine of Equivalents operates when subsequent invention represents "an overwhelming technological leap" beyond the original patent disclosure.); Carolyn Abbot & David Booton, *Using Patent Law's Teaching Function to Introduce an Environmental Ethic into the Process of Technical Innovation*, 21 GEO. INT'L ENVTL. L. REV. 219, 249 n.130 (2009) (finding no infringement where "an improvement exceeds the contribution made by the patented invention"); Mark A. Lemley, *Improvement in IP*, *supra* note 87, at 1010, 1012 (finding that the doctrine applies to "sufficiently radical" improvers and is "most likely" where the "value of the improvement greatly exceeds the value of the original invention").

[232] Donald S. Chisum, Bilski v. Kappos*: Everything Old is New Again: Weeds and Seeds in the Supreme Court's Business Method Patents Decision: New Directions for Regulating Patent Scope*, 15 LEWIS & CLARK L. REV. 11, 24 (2011) ("The prevailing standard for equivalency - substantial change - can take into account whether, on the one hand, an accused equivalent represents a merely inconsequential design around, or on the other, represents a significant invention."); *see also,* Ben Depoorter, *The Several Lives of Mickey Mouse: The Expanding Boundaries of Intellectual Property Law*, 9 VA. J.L. & TECH. 4, 51 (2004) (Reverse Doctrine of Equivalents requires "a significant contribution that takes the invention outside of the original, allegedly infringed-upon patent."); Alan Devlin, *Patent Law's Parsimony Principle*, 25 BERKELEY TECH. L.J. 1693, 1738 n.220 (2011) ("The reverse doctrine of equivalents serves to free a sufficiently large leap forward in the prior art from the claims of an earlier patent that would otherwise block use of the new technology."); Tex. Instruments, Inc. v. U.S. Int'l Trade Comm'n, 846 F.2d 1369, 1371 (Fed Cir. 1988) (applying doctrine where defendant's device is "sufficiently different" from plaintiff's).

[233] Although most commentators stress cleverness, a few argue that value should be measured by economic success or performance. Lemley, *Improvement in IP*, *supra* note 87, at 1065 (arguing that Reverse Doctrine of Equivalents should apply where improvement "is such a major advance . . . that its value is much greater than the original"); Devlin, *supra* note 232, at 1739 (arguing that "significant leaps" can be measured by whether the underlying patent's technology is "commercially defunct"). Implementing these user-centered concepts would force juries to make full-blown UTILITY-X judgments assessing the absolute utility of different inventions. U.S. law has long disfavored this notion. N. Scott Pierce, *In re* Dane K. Fisher: *An Exercise in Utility*, 6 *J. High Tech. L.* 1, 4 (2006) (surveying case law and scholarship distinguishing "positive utility" from "degree").

[234] Boyden Power Brake Co. v. Westinghouse, 170 U.S. 537, 573 (1898).

[235] *Id.*

[236] *Id.* at 573 (emphasis added).

inventor."[237] This sounds very much like a rule that 2G inventors should escape infringement where the 1G patent possessed no option value for their specific invention.

### *B. Pioneer Patents*

The Reverse Doctrine of Equivalents grew out of an earlier rule that "pioneer" inventions should receive broader patent protection than mere "improvements."[238] Unlike the Reverse Doctrine of Equivalents, the concept of Pioneer Patents expressly extends to non-literal infringement. Like the Reverse Doctrine, its status is unclear. While the Federal Circuit seemed to overrule it in the late 1980s,[239] more recent courts[240] and commentators[241] continue to recognize it.

There is also a split of authority on how the rule should be phrased. Most courts hold that pioneer patents must represent a radical improvement over prior art.[242] This standard is variously expressed in phrases like "broad breakthrough," "major advance," "basic operational concept," "broadly new," "devoid of significant prior art," "primary," "basic," "generic," "original," or "key." [243] However, some courts insist that the "pioneer" and "improvement" concepts only mark the ends of a continuum[244] so that breadth should be continuously adjusted for even small differences in cleverness. These alternatives are similar to those at issue in Reverse Doctrine cases and should presumably be decided the same way.

---

[237] Robert A. Bohrer & John T. Prince, *A Tale of Two Proteins: The FDA's Uncertain Interpretation of the Orphan Drug Act*, 12 HARV. J.L. & TECH. 365, 407 (1999). The view is also implicit in arguments that the Reverse Doctrine of Equivalents is an equitable doctrine designed to prevent unjust windfalls.

[238] *Westinghouse,* 170 U.S. at 561–62.

[239] Tex. Instruments, Inc. v. U.S. Int'l Trade Comm'n, 846 F.2d 1369, 1370 (Fed Cir. 1988); *see generally* Brian J. Love, *Interring the Pioneer Invention Doctrine*, 90 N.C.L. REV. 379, 394–95 (2012).

[240] *See* Love, *supra* note 239, at 395–96.

[241] Favorable discussions of the doctrine include John R. Thomas, *The Question Concerning Patent Law and Pioneer Inventions*, 10 BERKELEY HIGH TECH. L.J. 35 (1995); Michael J. Meurer & Craig A. Nard, *Invention, Refinement and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents* 93 GEO. L.J. 1947, 1989 (2005), and Georgia E. Kralovic, *The Principle of Fair Notice: Is It Prudent Guidance for the Future of Patent Law?*, 26 PEPP. L. REV. 89, 104–05 (1998) (remarking that pioneer status is "given little weight," but "make[s] sense intuitively.")

[242] *Tex. Instruments,* 846 F.2d at 1370 (quoting Boyden Power-Brake Co. v. Westinghouse, 170 U.S. 537, 561-62 (1898)).

[243] John R. Thomas, *The Question Concerning Patent Law and Pioneer Inventions*, 10 HIGH TECH. L.J. 35, 48 (1995).

[244] *See* Hughes Aircraft Co. v. United States, 717 F.2d 1351, 1362 (Fed. Cir. 1983); Cont'l Oil Co. v. Cole, 634 F.2d 188, 198 n.7 (5th Cir. 1981); Price v. Lake Sales Supply R.M., Inc., 510 F.2d 388, 394 (10th Cir. 1974); Shields v. Halliburton Co., 667 F.2d 1232, 1238 (5th Cir. 1982); Corning Glass Works v. Anchor Hocking Glass Corp., 374 F.2d 473, 476 (3d Cir. 1967).

### *C. Written Description*

In its most general form, the Written Description requirement dates from the Supreme Court's 1822 pronouncement that a patent applicant has an obligation "to describe what his own improvement is, and to limit [the] patent to such improvement."[245] However, this function became redundant with the rise of claims pleading and lay "dormant for many years."[246] In 1967, the Court of Customs and Patent Appeals' *Ruschig* decision[247] held that a patent whose disclosure potentially explained how to make "something like half a million possible compounds"[248] could not be amended to add a specific drug called "chlorpropamide."[249] The reason, the court explained, was that the inventor had provided no "motivation for wanting to make the compound in preference to others" and failed to disclose the compound "as something appellants actually invented."[250]

On the face of things, the *Ruschig* rule was limited to the narrow question of when amended claims could relate back to the original filing date.[251] In 1997, however, the Federal Circuit's *Eli Lilly* decision broadened the Written Description requirement to include original claims.[252] In particular, it held that inventors trying to patent a particular gene had to disclose an atom-by-atom description of the corresponding DNA molecule.[253] Most commentators interpreted this as a new and *ad hoc* "super-enablement" requirement for biotechnology inventions.[254] Formally, however, the Court seemed to be announcing the more general principle that courts should ignore literal infringement when the 1G patent failed to show that the inventor had actually "possessed" the 2G invention.[255] This view has since been confirmed by cases that apply the doctrine to non-biotechnology technologies like furniture design.[256]

*Lilly* was widely criticized by scholars[257] and some Federal Circuit judges.[258] In 2010, the Federal Circuit wrote the en banc *Ariad* decision to rationalize and explain the defense.[259] The case turned on a drug patent that purported to claim a method of

---

[245] Evans v. Eaton, 20 U.S. (7 Wheat.) 356, 435 (1822).

[246] Burk & Lemley, *Policy Levers*, *supra* note 16, at 1653.

[247] *In re* Ruschig, 379 F.2d 990 (C.C.P.A. 1967).

[248] *Id.* at 993.

[249] *Id.* at 995.

[250] *Id.*

[251] Moba, B.V. v. Diamond Automation, Inc., 325 F.3d 1306, 1319 (Fed. Cir. 2003) (explaining that the doctrine is designed to "prevent the addition of new matter to claims").

[252] Regents of the Univ. of Cal. v. Eli Lilly & Co., 119 F.3d 1559, 1566–67 (Fed. Cir. 1997).

[253] *Id.* at 1567.

[254] Christopher M. Holman, *Is* Lilly *Written Description a Paper Tiger?: A Comprehensive Assessment of the Impact of* Eli Lilly *and its Progeny in the Courts and PTO*, 17 ALB. L.J. SCI. & TECH. 1, 17 (2007) [hereinafter Holman, *Eli Lilly*].

[255] *Eli Lilly*, 119 F.3d at 1569 (finding no infringement where the claims were invalid for inadequate written description).

[256] Gentry Gallery, Inc. v. Berkline Corp., 134 F.3d 1473, 1479 (Fed. Cir. 1998).

[257] *See, e.g.*, Holman, *Eli Lilly*, *supra* note 254, at 17.

[258] *See, e.g.*, *id.* at 18.

[259] *Ariad Pharmaceuticals, Inc. v. Eli Lilly & Co.*, 598 F.3d 1336 (Fed. Cir. 2010) (*en banc*).

regulating cellular responses by interfering with the binding of certain molecules.[260] The court held that the Written Description requirement barred the claim because the patentee had failed to disclose any molecules capable of achieving the desired result.[261] Conceding that the "possession" test had "never been very enlightening," the court added that the patent had to "describe an invention understandable to [a] skilled artisan" and, further, "show that the inventor invented the invention claimed."[262] This would prevent patent owners from "merely recit[ing] a description of the problem to be solved while claiming all solutions to it and . . . cover any compound later actually invented and determined to fall within the claim's functional boundaries—leaving it to the pharmaceutical industry to complete an unfinished invention."[263]

Barring Supreme Court intervention, *Ariad* will likely remain definitive for decades. For our purposes, it makes three important points. First, Written Description is a universal defense that applies to literal and non-literal infringement alike. Second, Written Description focuses on what the 1G inventor actually knew. This is subtly different from the Reverse Doctrine of Equivalents, which asks whether the 2G inventor would have benefitted from reading the 1G patent. Finally, the test is objective and asks how a PHOSITA would have interpreted the 1G inventor's disclosure. This presumably guards against obscure hints that could only be understood by a genius-level (*e.g.* PHOSITA-95) inventor.

### *D. Analysis (Pt. 1): Option Value Revisited*

The Reverse Doctrine of Equivalents, Pioneer Patent, and Written Description defenses ask juries to make a second PHOSITA-X determination beyond *Graham's* baseline PHOSITA-50 inquiry. We now ask how well these additional inquiries narrow the gap between doctrine and our third social goal.

We begin by asking which defenses best fit our option value theory of utility. We have already said that option value often depends on the 2G inventors' ability to spot unusually clever (*e.g.* PHOSITA-95) applications and improvements. The Reverse Doctrine captures such cases by asking whether the 1G patent would have helped the 2G inventor achieve her insight or, conversely, "would scarcely have suggested the method . . . adopted."[264] By comparison, the Written Description test is narrowly focused on what the 1G inventor actually knew and communicated in terms that would have been understandable to a PHOSITA-50. This necessarily excludes hints to 2G geniuses that often possess significant option value. Judges who accept our option value arguments should want to replace Written Description with a reinvigorated Reverse Doctrine of Equivalents defense.

---

[260] *Id.* at 1340–42 ("disclosing their discoveries and claiming methods for regulating cellular responses to external stimuli by reducing NF-kB activity in a cell").

[261] *Id.* at 1355 ("Thus, to satisfy the written description requirement for the asserted claims, the specification must demonstrate that Ariad possessed the claimed methods by sufficiently disclosing molecules capable of reducing NF-kB activity.").

[262] *Id.* at 1351.

[263] *Id.* at 1353.

[264] Boyden Power-Brake Co. v. Westinghouse, 170 U.S. 537, 573 (1898).

### *E. Analysis (Pt. 2): Allocation Rules*

We have seen that courts and commentators have interpreted the Reverse Doctrine of Equivalents, Pioneer Patents, and Written Description defenses in different ways over the years. Here, we examine the various interpretations and ask which one best serves our third social goal.

*Absolute Cleverness.* Some courts have argued that the Reverse Doctrine of Equivalents and Pioneer Patent defenses should apply whenever the 2G patent is exceptionally clever in absolute terms. However, this ignores situations where the 1G patent's cleverness (and, implicitly cost) is comparable to or exceeds the 2G patent's. Courts that accept our analysis should reject these decisions out of hand.

*Direct Comparison.* Some versions of the Reverse Doctrine of Equivalents and Pioneer Patent defenses hold that the 2G patent should escape infringement when its cleverness is only slightly larger than the 1G patent's. Conceptually, this is equivalent to comparing the 2G patent to a PHOSITA-X threshold located at the peak of the 1G invention's verdict distribution. After that, the analysis is the same as our discussion of strong Doctrine of Equivalents rules.[265] In particular, we expect the defense to produce improved allocations whenever the 1G and 2G inventions are comparably clever. Conversely, inventions that possess markedly different cleverness will usually receive 50:50 (or 0:100) allocations without regard to cleverness or relative cost.

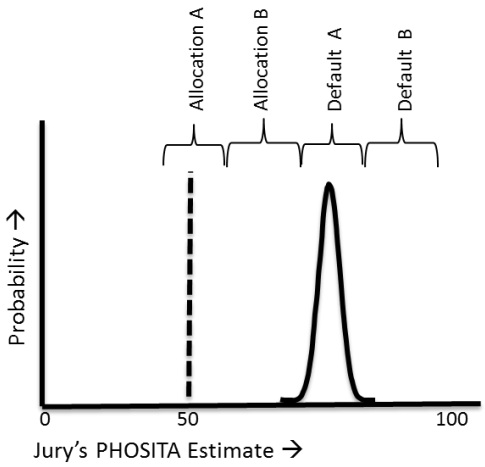
*Radical Improvements.* Most Reverse Doctrine of Equivalents and Pioneer Patent cases specify that the 2G invention must represent a "radical" improvement over its 1G predecessor to escape infringement. Figure 5 summarizes the situation. Reading from right to left, the PHOSITA-50 standard establishes a probabilistic zone ("Allocation A") that ranges from 50:50 to 0:100 as before. Now, however, there is a second probabilistic zone ("Allocation B") where the 2G patent is radically cleverer than its 1G predecessor. Because the radical improvement rule does *not* apply where the two inventions are similarly clever, our first Default A zone is centered on the 2G patent. This puts the 50:50 default blocking outcome exactly where it should be: where the two inventions are already known to have similar cleverness. Finally, there may also be a second ("Default B") zone that provides 0:100 allocations in cases where the 2G patent is much cleverer than its 1G predecessor.[266]

---

[265] *See supra* Figure 4 and accompanying text.

[266] We should not insist on such radical improvements that the new Allocation B zone overlaps the original PHOSITA-50 region and becomes redundant. This could happen, for example, if juries interpreted "radical" improvement in such extreme terms that the defense was only available to, say, PHOSITA-99 inventions.

**FIGURE 5**




Based on the foregoing, we argue that Pioneer Patent/Reverse Doctrine of Equivalents defenses that require "radical" improvement rules are the best way to implement our third policy goal. Courts that accept our analysis should discard the Written Description test and concentrate on reviving these doctrines.

## VIII. DOCTRINE VS. THEORY: ASSESSING THE SHORTFALL

We have now completed our survey of how patent doctrine implements modern neoclassical theory. On the positive side, we have seen that doctrine does a reasonably good job of tracking our three neoclassical goals. At the same time, any patent system that includes *Graham*'s PHOSITA test is unavoidably imperfect. This leaves us in an awkward position. If the current system worked exceptionally well – or badly – we would immediately know whether to keep it. Instead, the question stands on a knife's edge.

This Section begins by recapitulating the strengths and weaknesses of existing doctrine. We then ask whether a radically different system that discarded the PHOSITA proxy and tried to estimate R&D costs directly could provide a better fit.

### *A. How Good is the Proxy?*

We have argued that the PHOSITA standard is best understood as a proxy for R&D cost. Like all proxies, this introduces various possible errors.

*Imperfect Proxy*. No proxy is perfect. We have seen that there are a significant number of cases where cleverness says little or nothing about R&D cost. In other cases, incremental changes in cleverness can receive disproportionately large (or small) rewards.

*Ignoring Profit.* Because our PHOSITA proxy only tracks relative R&D costs, it cannot detect the moment when each inventor recoups her absolute cost and begins to earn a profit. For this reason, no PHOSITA-based doctrine can ever fully implement our Absolute Scope rule by redistributing profit upstream.

*Systematic Error.* Our probabilistic PHOSITA mechanism often generates 50:50 or 0:100 default allocations that are divorced from cleverness or, implicitly, cost. Adding a Reverse Doctrine of Equivalents defense mitigates but does not solve this problem.

*Under-Rewarding Multi-Step Inventions.* Technologies based on multiple patents often lead to multiple lawsuits. In these cases, collateral estoppel will truncate patent rewards compared to a single, one-shot lawsuit.

*A Cushion for Mistakes?* We have argued that any *Graham*-based doctrine is inherently imperfect. That being said, policymakers can still influence which kinds of mistakes it makes. More specifically, it is always better to have an over-priced invention than no invention at all. This suggests that legislators and judges should err on the side of broad patent monopolies and liberal rewards. The argument is even stronger in the multi-inventor case, where breaking just one link in a long development chain can kill promising technologies.

The good news is that current doctrine allocates reward on the basis of *relative* costs. This means that each inventor will still be able to cover her costs provided that the *absolute* reward is large enough. At the same time, a deliberately inflated reward violates our second policy goal. Given current state of the art, any attempt to adjust this tradeoff is largely a matter of guesswork.

*Industry-Specific Rules.* This article has pursued the traditional intuition that global rules are attractive and worthwhile. But real judges do not need to be purists . Instead, industry-specific rules can and do provide a safety valve when global doctrine fails to fit to conditions on the ground.

Pharmaceutical R&D provides a spectacular example of such ad hoc-ery.[267] This is not surprising since the industry often features (a) inventions that have no immediate use, and (b) R&D costs that typically rise several-fold for each step along the development pipeline. These unusual facts make industry-specific rules particularly attractive. At the same time, Big Pharma is uniquely well-funded and able to protect its interests. How would we know if other, poorer industries were encountering similar problems? Unlike industry-specific rules, a sensible global rule protects rich and poor alike. For this reason, judges should adopt *ad hoc* solutions reluctantly and then only as a last resort.

### *B. The Road Not Taken*

It is not enough to say that the PHOSITA standard is "good" or "bad." We must also ask, "Compared to what?" Here, the obvious alternative is asking courts to

---

[267] For example, courts may be more willing to apply the Reverse Doctrine of Equivalents in biotechnology than other industries. *See* Handley, *supra* note 224, at 45–50; Scripps Clinic & Research Found. v. Genentech, Inc., 927 F.2d 1565 (Fed. Cir. 1991); Amgen, Inc. v. Chugai Pharm. Co., 706 F. Supp. 94, 96 (D. Mass. 1989).

estimate R&D cost directly. Doctrinally, this means moving the allocation inquiry from liability to damages. Existing case law suggests at least two ways to do this.

*Apportionment.* The most direct solution is to apportion revenue according to each patent's respective "contribution" to products purchased by end-users. The idea is not as strange as it sounds. Indeed, some district courts apportioned damages in the Nineteenth and early Twentieth Centuries[268] while Congress debated the idea as recently as 2009.[269] There is also considerable scholarly support for the idea. Professor Amy Landers makes a compelling argument that reasonable royalty damages would divide revenues more fairly among patent owners.[270] Similarly, Professor Peter Lee argues that the related equitable doctrine of "accession" creates "a middle zone of flexibility" between 0:100 and 100:0 outcomes.[271] This might "embolden" courts to allocate reward differently between pioneer and improving inventors.[272]

From our standpoint, neither of these proposals is quite right. The reason is that they focus on each patent's "contribution" in providing value (i.e., utility) to end-users. This concept is decidedly different from asking what costs each inventor incurred to advance the technology. Even so, shifting the allocation inquiry to R&D costs sounds like a minor detail.

*Compulsory Licenses.* Compulsory licenses provide a second way to apportion revenue. Professor Tur-Sinai has argued that this is the best way to implement the Absolute Scope rule.[273] However, modern courts are notoriously hostile to the idea.[274] That being said, the concept remains fairly mainstream and continues to appear in academic articles.[275]

*New Headaches.* The foregoing approaches would require courts to estimate corporate R&D costs directly. Naively, this sounds straightforward. On reflection, though, one could say the same thing about public utility commissions' efforts to set fair rates of return. In practice, direct cost estimation would require specialized bureaucracies and make frequent errors. As the U.S. Supreme Court has pointed out in the antitrust context, regulation is a "daunting task"[276] which courts should avoid.[277]

*Replacing* Graham. We have argued that any doctrine that includes *Graham*'s non-obviousness test automatically contains a baseline allocation rule for successive inventors. On the other hand, our first social goal still requires a threshold standard. If we overrule *Graham* we must replace it with something. The least radical solution is to retain *Graham* and only estimate costs for whatever patents survive the non-

---

[268] Landers, *supra* note 188, at 499–500 (2012).

[269] *See* Patent Reform Act of 2009, H.R. 1260, 111th Cong. (2009).

[270] Landers, *supra* note 188, at 505–07. (arguing that the current reasonable royalty calculation standards are too restrictive, but could be reworked to provide a more fair outcome).

[271] Peter Lee, *supra* note 230, at 239.

[272] *Id.*

[273] Tur-Sinai, *supra* note 62, at 743.

[274] Merges & Nelson, *supra* note 17, at 840.

[275] Nielsen & Samardzija, *supra* note 16, at 535 (describing a recent survey of the compulsory licensing as an alternative to damages in both theory and practice); *see also*, Rose, *supra* note 230, at 618–24 (advocating "limited" compulsory licenses where "market failure" exists).

[276] Verizon Commc'ns., Inc. v. Law Offices of Curtis v. Trinko, 540 U.S. 398, 414 (2004).

[277] *Id.* at 415.

obviousness screen. However, this would only improve incentives for 2G inventions that are much cleverer than PHOSITA-50. An ambitious reform would replace *Graham* itself. Under this rule, R&D projects that incurred above-average costs would be deemed non-obvious. This, however, would expand the inquiry still further: in addition to estimating the parties' own R&D costs, the court would also have to estimate the average cost of similar projects throughout the industry.

## CONCLUSION

Nearly a half-century has passed since the Supreme Court's *Graham* decision built neoclassical economics into the foundations of patent doctrine. But while innovation economics has moved on, most legal scholars still treat the literature as a grab bag of conflicting theories. This gloomy assessment seems to put global rules out of reach. It is also unreasonable. Far from being fragmented, the various neoclassical insights share the same fundamental assumptions and methods. For this reason, they are best seen as special cases of a deeper and more general theory. This article has identified and explored three social goals that any such theory must have. Two of these are generalized versions of principles (non-obviousness, reward size) that were already familiar in the 1960s. The third specifies how patent reward should be divided among successive inventors.

Remarkably, current doctrine uses *Graham*'s PHOSITA standard to address all three goals. This practically guarantees that attempts to address one goal will have unintended impacts on the others. We have argued that separate PHOSITA-X standards can relax this link for our first and second principles. Future attempts to reform the Doctrine of Equivalents should take advantage of this fact. At the same time, *any* doctrine based on a PHOSITA non-obviousness test *automatically* implements economically inefficient allocations. A revived Reverse Doctrine/Pioneer Patents defense would go some distance toward narrowing this gap.

No doctrine that includes the PHOSITA concept can ever be perfect. Full implementation of our third social goal would require judges to scrap the current system in favor of estimating corporate R&D costs directly. This is bound to be a large and error-prone undertaking. By comparison, the PHOSITA standard—despite its flaws—is at least reasonably accurate and can be implemented by judges and juries. Radical change, if it comes at all, should be left to Congress.