The Value of Accuracy in the Patent System

Stephen Yelderman†

Because it must rely on imperfect information, the patent system will inevitably make mistakes. To determine how the system ought to err in cases of uncertainty—and whether a given mistake is worth correcting—scholars have composed a simple picture of the consequences of error in either direction. On the one hand, erroneous patent awards impose unjustified costs. On the other hand, erroneous patent denials discourage successful inventors and reduce incentives to create in the future. The result is an essentially indeterminate balancing, in which policies of overly liberal awards drive up costs, and policies of overly cautious awards drive down incentives.

As this Article will show, this conventional approach to error costs understates the role that accuracy plays in producing the benefits of the patent system. Critically, the incentives to invent created by the patent system depend on the difference between an inventor's expected returns if she invents and her expected returns if she does not invent. Erroneous patent awards do not simply increase the costs of the patent system but also narrow the expected difference between inventing and not inventing. Undeserved patent rights thus undermine the very incentives the system is intended to create.

This Article presents a framework for evaluating the value of accuracy in the patent system. As it turns out, the consequences of an undeserved patent depend significantly on a factor that has not been previously given much attention: whether the unsatisfied patentability requirement is one that seeks to influence a mutually exclusive choice. Some patentability doctrines satisfy this condition, but others do not. The result is that erroneous patent awards may in some ways be more harmful and in other ways less harmful than previously thought.

Introduction

Today, it is an almost universally accepted proposition that the patent system makes too many mistakes. Not that there are too many randomly distributed errors, but a pattern of mistakes that result in a system that is consistently (and inappropriately) biased in a pro-patent direction. Patents, it is thought, are too easy to acquire in the first place, too difficult for challengers to revoke later on, and too profitable to enforce in dubious circumstances. Countless books and law review articles have been founded on essentially this premise,[1] and in recent years both

---

[1] See, for example, Robert P. Merges, *As Many as Six Impossible Patents before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 Berkeley Tech L J 577, 589–91 (1999) (linking the "numerous incentives inside the PTO to issue rather than reject patent applications" to the problem of low-quality patents); John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U Ill L Rev 305, 314–16, 318–21; Jay P. Kesan, *Carrots and Sticks to Create a Better Patent System*, 17 Berkeley Tech L J 763, 767–68 (2002); John H. Barton, *Non-Obviousness*, 43 IDEA 475, 493–95 (2003); Adam B. Jaffe and Josh Lerner, *Innovation and Its Discontents: How Our Broken Patent System Is Endangering Innovation and Progress, and What to Do about It* 136–37 (Princeton 2004); Joseph Farrell and Robert P. Merges, *Incentives to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help*, 19 Berkeley Tech L J 943, 944–46 (2004) ("[G]iven the rapid increase in the volume of patent applications . . . , it would be astounding if patent quality had *not* suffered."); Christopher R. Leslie, *The Anticompetitive Effects of Unenforced Invalid Patents*, 91 Minn L Rev 101, 133–34 (2006) (discussing the ways in which courts err in favor of the patent holder during patent validity challenges); Doug Lichtman and Mark A. Lemley, *Rethinking Patent Law's Presumption of Validity*, 60 Stan L Rev 45, 47–48 (2007); Dan L. Burk and Mark

Congress and the Supreme Court have taken steps to weaken the power of patents having "suspect validity."[2]

Despite the bevy of proposals intended to curb the granting and enforcement of undeserved patent rights, there is surprisingly little basis to conclude that the current state of things is actually suboptimal. It is true that the system currently makes patent rights available in some cases in which the black-letter patentability requirements have probably not been satisfied.[3] But to conclude that the system ought to err differently—that is, that uncertain cases should be decided *against* patent protection—one must have some way of comparing the costs of an error in either direction. If the costs of erroneously denying patent protection are greater than the costs of erroneously providing patent protection, then there may be no reason to change the current balance of errors. In fact, doing so might actually increase the costs of errors overall, causing more harm than good.[4]

And it is here that existing theory and empirics flounder. The problem is that the conventional understanding of error costs in the patent system yields no definite conclusions. On the one hand, it is widely accepted that undeserved patents impose unjustified costs. The private value of patents necessarily comes at some expense to the public, and erroneous grants incur these costs without offsetting public benefits.[5] On the other hand, it is

A. Lemley, *The Patent Crisis and How the Courts Can Solve It* 25, 28–29 (Chicago 2009) (describing how the current system of patent litigation encourages holdup and impedes innovation); Michael J. Burstein, *Rethinking Standing in Patent Challenges*, 83 Geo Wash L Rev 498, 538–42 (2015); Roger Allan Ford, *The Patent Spiral*, 164 U Pa L Rev 827, 837–39, 858–63 (2016). See also Jonathan Masur, *Patent Inflation*, 121 Yale L J 470, 473 n 5 (2011) (summarizing this literature).

    2   See *eBay Inc v MercExchange, LLC*, 547 US 388, 397 (2006) (Kennedy concurring); *Bilski v Kappos*, 561 US 593, 608 (2010). See also Leahy-Smith America Invents Act (AIA), Pub L No 112-29, 125 Stat 284 (2011).

    3   See Part I.B.

    4   See John M. Golden, *Patentable Subject Matter and Institutional Choice*, 89 Tex L Rev 1041, 1068 (2011).

    5   See *Lear, Inc v Adkins*, 395 US 653, 670–71 (1969); Mark A. Lemley and Carl Shapiro, *Probabilistic Patents*, 19 J Econ Persp 75, 77 (Spring 2005); Jeremy W. Bock, *Does the Presumption of Validity Matter? An Experimental Assessment*, 49 U Richmond L Rev 417, 449 (2015); Gregory Mandel, *Patently Non-Obvious II: Experimental Study on the Hindsight Issue before the Supreme Court in* KSR v. Teleflex, 9 Yale J L & Tech 1, 31–32 n 129 (2007) (questioning whether granting an undeserving patent incurs more social costs than denying a deserving patent); T. Randolph Beard, et al, *Quantifying the Cost of Substandard Patents: Some Preliminary Evidence*, 12 Yale J L & Tech 240, 243–45

equally accepted that mistakes in the opposite direction—denials of patent protection to those who deserve it—undermine the private incentives to invent that are the reason for having a patent system in the first place.[6] The theory, after all, is that inventors will invest in research and development in expectation of receiving patent rights. When that promise is not kept, future inventors will expect smaller rewards from the patent system going forward and consequently will invest less in the very research and development that the patent system is supposed to encourage.

In the end, this understanding of errors reduces to a referendum on the costs and benefits of the patent system.[7] If the marginal benefits of patent-induced innovation are large compared to the marginal costs of an incremental patent, it is preferable to grant lots of them. But, if the marginal costs of each additional patent grant are large compared to the marginal benefits of patent-induced innovation, it is better to be quite stingy with patent rights.[8] And because we lack answers to these central empirical questions,[9] the conventional approach yields no clear guidance one way or the other.[10] Faced with this puzzle,

(2010); Leslie, 91 Minn L Rev at 127 (cited in note 1); Rochelle Cooper Dreyfuss, *Nonobviousness: A Comment on Three Learned Papers*, 12 Lewis & Clark L Rev 431, 434–35 (2008) (noting the existence of "structural features in the patent system" that systematically generate erroneous grants); Burstein, 83 Geo Wash L Rev at 549 (cited in note 1); Shubha Ghosh and Jay Kesan, *What Do Patents Purchase? In Search of Optimal Ignorance in the Patent Office,* 40 Houston L Rev 1219, 1228, 1244–45 (2004); Merges, 14 Berkeley Tech L J at 592–93 (cited in note 1); Shubha Ghosh, *Beyond Hatch-Waxman*, 67 Rutgers L Rev 779, 801 (2015). See also Michael J. Meurer and Katherine J. Strandburg, *Patent Carrots and Sticks: A Model of Nonobviousness*, 12 Lewis & Clark L Rev 547, 556 (2008) (noting the typicality of the cost-benefit approach to measuring erroneous patent grants).

    6   See Andres Sawicki, *Better Mistakes in Patent Law*, 39 Fla St U L Rev 735, 760 (2012); R. Polk Wagner, *Understanding Patent-Quality Mechanisms*, 157 U Pa L Rev 2135, 2141 (2009); Bock, 49 U Richmond L Rev at 448 (cited in note 5); Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 Nw U L Rev 1495, 1521 (2001).

    7   See *Mayo Collaborative Services v Prometheus Laboratories, Inc*, 566 US 66, 92 (2012); Sawicki, 39 Fla St U L Rev at 744 (cited in note 6); Bock, 49 U Richmond L Rev at 448–49 (cited in note 5).

    8   See Glynn S. Lunney Jr, *E-Obviousness*, 7 Mich Telecom & Tech L Rev 363, 385–86 (2001); Sawicki, 39 Fla St U L Rev at 744 (cited in note 6).

    9   See Oskar Liivak, *Establishing an Island of Patent Sanity*, 78 Brooklyn L Rev 1335, 1337–38 (2013); Lisa Larrimore Ouellette, *Patent Experimentalism*, 101 Va L Rev 65, 75–84 (2015).

    10   See Mandel, 9 Yale J L & Tech at 31–32 n 129 (cited in note 5); Bock, 49 U Richmond L Rev at 448–49 (cited in note 5).

scholars have either explicitly reserved judgment on the question of how the patent system ought to err[11] or simply fallen back on (disputed) priors about the costs and benefits of patent protection.[12]

As this Article will show, this widely adopted framework is not only indeterminate but also incomplete. The fundamental problem is that it understates the role that accuracy plays in producing the benefits of the patent system. The reason for offering patent protection is to create incentives to do particular things: to create new, useful, and nonobvious inventions, to disclose them publicly, and to do all this while complying with a number of rules designed to protect the public and other inventors. The magnitude of the patent incentives to do these things depends both on what the system rewards and on what the system does not reward—just as the power of a prize depends on granting it when it is deserved and withholding it otherwise. Erroneous patent grants narrow the difference between the expected outcome from inventing and the expected outcome from not inventing, thus reducing the marginal reward offered to do the former instead of the latter. In this way, undeserved patents do not simply impose unjustified costs, but actively undermine the very ex ante incentives that the patent system is intended to create.

Prior scholarship has noted these incentive effects only in passing and has not explored their consequences for how the patent system should err in cases of uncertainty.[13] At a minimum, adding them to the traditional account of error costs necessarily shifts the optimal balance of errors in the direction of more skepticism toward claims of patent rights. With a few basic assumptions about observability and inventors' perceptions of

---

[11]   See, for example, Daniel J. Hemel and Lisa Larrimore Ouellette, *Beyond the Patents–Prizes Debate*, 92 Tex L Rev 303, 330–31 n 134 (2013); Mandel, 9 Yale J L & Tech at 31–32 n 129 (cited in note 5); Bock, 49 U Richmond L Rev at 449 (cited in note 5); Ghosh and Kesan, 40 Houston L Rev at 1227–29 (cited in note 5).

[12]   See, for example, Dreyfuss, 12 Lewis & Clark L Rev at 435–36 (cited in note 5).

[13]   See Matthew Sag and Kurt Rohde, *Patent Reform and Differential Impact*, 8 Minn J L Sci & Tech 1, 8–9 (2007); Einer Elhauge and Alex Krueger, *Solving the Patent Settlement Puzzle*, 91 Tex L Rev 283, 294–95 (2012); Murat C. Mungan, *Reverse Payments, Perverse Incentives*, 27 Harv J L & Tech 1, 44 (2013); Stephen Yelderman, *Coordination-Focused Patent Policy*, 96 BU L Rev 1565, 1592–93 & n 129 (2016); Federal Trade Commission, *To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy* ch 5 at 1–2 (Oct 2003), archived at http://perma.cc/X8E7-UWU3.

the patent system's errors, the conclusion becomes stronger. When certain conditions hold, this incentive harm can be used to show that the probability of patentability necessary to justify patent rights must be *at least* 50 percent—a "more likely than not" standard that is more rigorous than the de facto leniency toward questionable patent rights that exists today. This is true regardless of the marginal costs and benefits of patent protection, providing a basis for increasing scrutiny of patent rights that does *not* depend on disputed empirical priors about the costs and benefits of the patent system.[14]

But there are complications as well. Most importantly, the effect of an undeserved patent turns out to depend significantly on the reason that patent was undeserved. Some of the patentability requirements are intended to influence mutually exclusive choices—failing to enforce *these* requirements not only incurs unjustified costs, but also reduces marginal rewards, thus weakening the power of the incentives created by the patent system in the future. But other patentability requirements have nothing to do with shaping incentives, and exist only to mitigate the costs of the patent system. Failing to enforce these requirements drives up the costs of the patent system, but does not harm future incentives. As a result, in cases involving mutually exclusive choices, erroneous patent grants are more harmful than previously appreciated. But in other cases, errors in favor of patentability might not be quite as detrimental as previously assumed.[15]

This Article proceeds in five parts. Part I provides background on the traditional approach to error costs in the patent system and describes how current doctrine generally breaks in favor of patent rights in close cases. Part II shows how this

---

[14]   To put this contribution in context, this Article takes substantive patent law as it stands and explores the question of how the patent system should err in the resolution of individual cases in which limited information makes the underlying facts uncertain. This should be distinguished from prior work evaluating how close questions of statutory construction in patent law should be decided, see, for example, Joseph Scott Miller, *Error Costs & IP Law*, 2014 U Ill L Rev 175, 180–82, or whether it is better to correct mistakes earlier or later in a patent's life, see, for example, Lemley, 95 Nw U L Rev at 1496–97 (cited in note 6).

[15]   Although this effect has never been explored in detail, several prior scholars have suggested that undeserved patents can reduce incentives to invent. See note 13. By bringing this generalized intuition down to specifics, the present analysis reveals not only the power of this effect but also its limits.

traditional understanding is incomplete and develops a theory of accuracy that accounts for the relationship between false positives and future incentives. Parts III and IV apply that theory to a number of patent doctrines, illustrating how the importance of accuracy is tightly bound to a particular rule's purpose. Part V then discusses implications for patent law and highlights several questions requiring further study.

I. Background

When it comes to mistakes in the patent system, prior scholarship has typically presented a simple trade-off in which erroneous grants impose unnecessary ex post costs and erroneous denials reduce ex ante incentives. Part I.A introduces this approach to error costs, showing how it has led to essentially indeterminate prescriptions for how the patent system should err in cases of uncertainty. Part I.B then introduces a number of procedural rules and structural features that tend to favor patent rights in doubtful cases, illustrating that, in practice, the patent system currently makes patents available even in cases in which there is an objectively low probability that the patentability requirements have been satisfied.

First, however, a brief note on terminology: The following discussion is focused on the question of when the system as a whole should afford patent protection. But the concept of a "patent system" is really an abstraction. Decisions about patentability are divided across multiple actors weighing in at a number of distinct stages: the US Patent and Trademark Office (PTO) performs an initial examination and, in some cases, additional rounds of administrative review; a federal district court may make findings later on; and the Federal Circuit may review the district court's (or the PTO's) work at some time after that. To address the role of accuracy on a systemic level, the following discussion evaluates the decision to award or deny protection as if it were a single, unified decision, and without differentiating based on when or by whom that decision is made. As a general rule, when this Article refers to "examining" patents or "awarding," "granting," or "denying" patent protection, it does not mean to indicate a decision made at any particular stage of this process. Part V.B, however, briefly discusses the value of correcting erroneous patent grants that have already occurred.

A. The Conventional Account of Error Costs

According to the most widely accepted theory, the reason for awarding patents is to increase prospective inventors' incentives to invent.[16] When the patent system correctly gives something of value to a successful inventor, future prospective inventors come to expect their efforts too will earn them something of value. It is this private expectation of future reward that produces the promised public benefits of the patent system—more innovation as a result of increased incentives to invest in research and development.[17]

Mistaken denials of patent protection ("false negatives") undermine this goal. When deserving inventions are denied protection, the system fails to keep its promise to reward patentable inventions through a grant of exclusive rights.[18] For the inventor in question, of course, it is too late, because the invention has already been made. But if this outcome is observable by future prospective inventors,[19] they will rationally discount the

---

[16] See *Mayo Collaborative Services v Prometheus Laboratories, Inc*, 566 US 66, 92 (2012) ("[T]he promise of exclusive rights provides monetary incentives that lead to creation, invention, and discovery."); Ward S. Bowman Jr, *Patent and Antitrust Law: A Legal and Economic Appraisal* 2–3 (Chicago 1973). See also F.M. Scherer and David Ross, *Industrial Market Structure and Economic Performance* 621–24 (Houghton Mifflin 3d ed 1990); Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 Harv L Rev 1813, 1821–22 (1984) (describing patents as the "price society pays to stimulate inventive activity").

[17] In addition to this mission of rewarding invention, courts often (and scholars sometimes) mention a goal of encouraging disclosure. See Part III.B. Whether the goal is invention or disclosure, the basic mechanism is the same: ex post rewards to create future ex ante incentives. This Article sets aside other potential functions of the patent system (such as facilitating coordination or enabling commercialization), which do not necessarily rely on a "quid pro quo" rewards mechanism. See Yelderman, 96 BU L Rev at 1575–80, 1592–93 (cited in note 13). The costs of an error in either direction will change significantly if the patent system is intended to serve some other, nonrewards function. Id at 1598–1603.

[18] See Sawicki, 39 Fla St U L Rev at 760 (cited in note 6); Wagner, 157 U Pa L Rev at 2141 (cited in note 6); Bock, 49 U Richmond L Rev at 448 (cited in note 5); Mandel, 9 Yale J L & Tech at 31–32 n 129 (cited in note 5). See also Dreyfuss, 12 Lewis & Clark L Rev at 435–36 (cited in note 5) (acknowledging this risk but suggesting any harm to innovation will be small).

[19] Observability of outcomes is a central (and common) assumption in analysis of the error costs of the patent system. See, for example, Dreyfuss, 12 Lewis & Clark L Rev at 434 (cited in note 5). If future inventors cannot observe how present inventors are being treated, then there is no reason for the government to keep its end of the patent bargain. Moreover, if inventors know this, they have no reason to trust the patent-granting authority and thus will not make investments in reliance on its promise. For

expected value of participating in the patent system to reflect the risk that they too will be denied a patent when they deserve one. The failure to offer a reward in the deserving cases of the present thus reduces the expected value of investing in patentable inventions in the future.[20]

It is important to note that the relationship between any particular false negative and future incentives to invent is indirect. The harm of a false negative is rooted in ex ante incentives—the observed treatment of present inventors affects how *other* (future) inventors will expect to be treated. Because hundreds of thousands of patents are granted every year,[21] no single false negative by itself will scuttle future incentives to invest in socially valuable projects. Rather, it is the expected rate of false negatives in general that will determine the incentives created by the patent system. A small number of errors like these may not have much effect at all. But have too many, and the expected benefits of participating in the patent system will be reduced—undermining the patent system's central goal of increasing ex ante incentives to invest in the creation of patentable inventions.

In the other direction, the obvious harm caused by an erroneous award of patent protection (a "false positive") is the ex post cost of the undeserved patent itself. With or without errors, patents impose costs. In some cases, patents confer market power, causing deadweight losses—reduced consumption and higher prices for consumers—that would not exist in the absence of the patent grant.[22] Proliferation of patents can also result in higher

---

further discussion of the observability of patent outcomes, see notes 98–103 and accompanying text.

    20 See Wagner, 157 U Pa L Rev at 2141 (cited in note 6); Anup Malani and Jonathan S. Masur, *Raising the Stakes in Patent Cases*, 101 Georgetown L J 637, 650–52 (2013). See also Kaplow, 97 Harv L Rev at 1838 (cited in note 16).

    21 See US Patent and Trademark Office, *U.S. Patent Statistics Chart: Calendar Years 1963–2015* (June 15, 2016), archived at http://perma.cc/DG5B-8RUZ.

    22 See *Mayo*, 566 US at 92; Burstein, 83 Geo Wash L Rev at 539 (cited in note 1); Malani and Masur, 101 Georgetown L J at 657 (cited in note 20); Beard, et al, 12 Yale J L & Tech at 243–45 (cited in note 5); Leslie, 91 Minn L Rev at 127 (cited in note 1); Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 Tex L Rev 1031, 1059–60 (2005); Joseph Scott Miller, *Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents*, 19 Berkeley Tech L J 667, 690 (2004); Kesan, 17 Berkeley Tech L J at 767–68 (cited in note 1); Merges, 14 Berkeley Tech L J at 592 (cited in note 1); Lemley, 95 Nw U L Rev at 1517–19 & n 85 (cited in note 6) (discussing the social cost of patent holdup, which occurs when patent owners seek "to license even clearly bad patents for

transaction costs, because using a recently developed technology will require more patent searching, negotiation, and, potentially, litigation.[23] And patents can impose dynamic harm as well. For example, the exclusive rights of a patent grant can create incentives for others to design or invent "around" the patent. This work can be wasteful even when it succeeds, and sometimes it is prohibitively expensive or impossible.[24] In this way, a patent grant may not only reduce use of the patented technology itself, but also inhibit future innovation in related areas.[25]

All of these ex post costs are expected to be incurred whether or not a patent is deserved. The difference is whether these costs are justified. In the case of a meritorious grant, the costs of patenting are understood as the price to pay for rewarding invention through a system of exclusive rights. But, in the case of an unjustified grant, the full costs of a patent are incurred without any offsetting benefits.[26] Denial is preferable when a patent is undeserved, because it reduces the cost of having a patent system without harming ex ante incentives.

The patent system thus must balance the risk of failing to adequately reward invention (through erroneous denials) against

royalty payments small enough that licensees decide it is not worth going to court"); Jonathan S. Masur, *CBA at the PTO*, 65 Duke L J 1701, 1714 (2016).

[23] See *Mayo*, 566 US at 92; Kesan, 17 Berkeley Tech L J at 768 (cited in note 1); Lemley, 95 Nw U L Rev at 1502, 1507–08, 1515 (cited in note 6); Lemley, 83 Tex L Rev at 1064 (cited in note 22); Miller, 19 Berkeley Tech L J at 690 (cited in note 22); Malani and Masur, 101 Georgetown L J at 656–57 (cited in note 20).

[24] See Miller, 19 Berkeley Tech L J at 690 (cited in note 22); Leslie, 91 Minn L Rev at 127–28 (cited in note 1); Kesan, 17 Berkeley Tech L J at 767–68 (cited in note 1).

[25] See Lemley, 95 Nw U L Rev at 1516 (cited in note 6); Lemley, 83 Tex L Rev at 1060–62 (cited in note 22); Merges, 14 Berkeley Tech L J at 592 (cited in note 1); Miller, 19 Berkeley L J at 690 (cited in note 22). See also Leslie, 91 Minn L Rev at 127–28 (cited in note 1); Kesan, 17 Berkeley Tech L J at 767–68 (cited in note 1); Burstein, 83 Geo Wash L Rev at 539 (cited in note 1); Masur, 65 Duke L J at 1714–15 (cited in note 22).

[26] See *Lear, Inc v Adkins*, 395 US 653, 670 (1969) (noting the necessity of patent challenges, without which "the public may continually be required to pay tribute to would-be monopolists without need or justification"); Mandel, 9 Yale J L & Tech at 31–32 n 129 (cited in note 5); Beard, et al, 12 Yale J L & Tech at 241–42 (cited in note 5); Leslie, 91 Minn L Rev at 127 (cited in note 1); Dreyfuss, 12 Lewis & Clark L Rev at 435 (cited in note 5); Burstein, 83 Geo Wash L Rev at 549 (cited in note 1); Ghosh and Kesan, 40 Houston L Rev at 1228, 1244–45 (cited in note 5); Merges, 14 Berkeley Tech L J at 592–93 (cited in note 1); Ghosh, 67 Rutgers L Rev at 801 (cited in note 5); Masur, 65 Duke L J at 1715 (cited in note 22). See also Meurer and Strandburg, 12 Lewis & Clark L Rev at 556 (cited in note 5) (noting the typicality of this approach). For a discussion of disclosure benefits that might be lost in the case of a justified denial, see note 28.

the cost of unnecessary patents (through erroneous grants). The following table summarizes these potential outcomes:

TABLE 1. ONE-SIDED INCENTIVES MODEL

Decision	Invention Patentable	Invention Not Patentable
Award Patent	– Imposes Ex Post Costs (True Positive)	– Imposes Ex Post Costs (False Positive)
Deny Patent	– Reduces Ex Ante Incentives * No Ex Post Costs (False Negative)	* No Ex Post Costs (True Negative)

As Table 1 illustrates, the decision to award a patent (top row) imposes ex post costs whether or not the patentability requirements have been satisfied.[27] The consequences of a denial, however, turn on truth. Denying a patent always saves ex post costs—whether or not the denial was merited.[28] If the denial was deserved (bottom right quadrant), these ex post costs are avoided

---

[27]   This table and the ensuing discussion refer to the costs and benefits of these outcomes at a systemic level, without regard for whether an error is made earlier or later in time. For a discussion of the consequences of errors in either direction at the examination stage in particular, see Michael J. Meurer, *Patent Examination Priorities*, 51 Wm & Mary L Rev 675, 689–701 (2009); Sawicki, 39 Fla St U L Rev at 746–49, 753–58, 767–76 (cited in note 6).

[28]   As noted above, this analysis explicitly sets aside any ex post *benefits* that might follow from a grant of patent protection—such as facilitating coordination or increasing incentives to commercialize. See note 17. Nonetheless, it is possible that denying a patent will impose ex post costs, as in some cases the public will be unable to use the information disclosed in the rejected patent application. While this harm is theoretically possible, in practice it is quite rare for a denial of patent rights to result in any case-specific loss of disclosure at all. By default, patent applications become public eighteen months after filing, which will usually occur well before any final decision has been made about an applicant's entitlement to a patent. See 35 USC § 122(b)(1); 37 CFR § 1.211. Under certain conditions applicants may request nonpublication, see 35 USC § 122(b)(2)(B), but this option is infrequently exercised. In 2009 (the last year for which data is publicly available), nearly 95 percent of patent applications were published within eighteen months of filing. See Tegernsee Experts Group, *Study Mandated by the Tegernsee Heads: 18-Month Publication* *17 (Sept 2012), archived at http://perma.cc/NE8H-VZZ6. And even those 5 percent of applications not published at eighteen months can still become public if at any point they ultimately result in an issued patent. See 35 USC § 153. So, in the end, an erroneous denial results in the loss of that inventor's proffered disclosure only in the rare case in which: (a) the inventor elected nonpublication; (b) the erroneous denial occurred during examination (as opposed to postgrant proceedings or litigation); *and* (c) that erroneous denial by an examiner was at no point corrected through a request for continued examination, see 37 CFR § 1.114, or appellate review by the Patent Trial and Appeal Board, see 35 USC § 134.

with no downside. But if the denial was erroneous because the patentability requirements *were* satisfied (bottom left quadrant), these ex post cost savings are tempered by a reduction in future ex ante incentives to invent.

Under this framework, the certainty required to justify patent rights depends on the relationship between the ex ante incentives and the ex post costs of the patent system. To put the task of balancing false positives and false negatives in more formal terms, consider an objective examiner applying the substantive patentability requirements to the facts known to her in order to estimate the probability that a specific application satisfies those requirements.[29] Call this application-specific estimate q, which as a probability estimate ranges between zero and one. The task of the patent policymaker is to set a threshold T for all patent applications, such that applications with $q > T$ will result in patent rights and applications with q equal to or below that threshold will be denied patent rights.

Taking substantive patent law and the examiner's level of information as givens,[30] the ideal probability threshold T can be determined by comparing the expected consequences of an action in either direction. Whether deserved or not, granting a patent can be expected to impose ex post costs—call these estimated incremental costs C. Going the other direction, denying a patent avoids these ex post costs with certainty, but carries a risk of undermining ex ante incentives going forward. When the rejected invention really was patentable, a denial reduces prospective inventors' expectations of receiving patent rights when they deserve them in the future. The resulting diminished ex ante incentives reduce the public benefits of the patent system by some amount I_{FN}.[31] Unlike ex post costs, however, this harm

---

[29]   The term "objective examiner" is used here to distinguish this theorized probability estimation step from the way patent examiners and district-court judges actually scrutinize patents under present arrangements. As discussed in Part I.B, these decision-makers review patents and patent applications through the lens of various presumptions, and may themselves be subject to certain biases.

[30]   For a discussion of this choice, see notes 32–34 and accompanying text.

[31]   An important feature of this framework is that the objective examiner can make no individualized assessments of C and I_{FN}; she may apply only generally applicable patent law to the facts of the case and make the decision to grant or deny based on the resulting output q. This constraint reflects the longstanding principle of patent law that the validity of patent rights is a legal question, not a matter of agency or court discretion. See 35 USC § 102 ("A person shall be entitled to a patent unless . . . ."). See also Sawicki, 39 Fla St U L Rev at 744–45 (cited in note 6).

is incurred only in cases in which patent rights were actually deserved. The expected cost of denying patent rights can thus be written as $q * I_{FN}$—the probability-discounted harm to future ex ante incentives in the event of a false negative.

Patent rights should be awarded when the cost of doing so is less than the cost of denying them—that is, when $C < q * I_{FN}$. A patent should thus be granted if and only if $q > T$, where

(1) $$T = \frac{C}{I_{FN}}$$

As this equation illustrates, under the conventional view, the probability necessary to justify awarding patent protection turns on the relationship between the ex post costs C of granting an additional patent and the public harm I_{FN} from reduced ex ante incentives following an erroneous denial. If the ex post costs of an incremental patent grant are low compared to the public benefits that will be lost from a diminution of ex ante incentives, the minimum probability necessary to justify patent rights will be low as well, and the patent system ought to grant promiscuously. But if the ex post costs of an incremental patent grant are large (or the public benefits that will be lost as a result of an erroneous denial are small), the probability of patentability necessary to justify a patent will be quite demanding.

Because this basic framework (and similar variations) will be used throughout the Article, it is worth pausing to recognize its limitations. This is a model only for setting the probability of patentability necessary to justify a patent grant, not for minimizing the total error costs of the patent system. The decision to take (a) substantive patent law and (b) the objective examiner's level of information as givens results in serious constraints on a policymaker's ability to maximize the public benefits of the patent system.[32] These constraints are embraced here for consistency with the literature suggesting that the patent system grants too many invalid patents,[33] and they will prove useful for

---

[32]   For a generalized model of error costs in the patent system, see Golden, 89 Tex L Rev at 1065–74 (cited in note 4).

[33]   The literature on patent quality almost inevitably accepts substantive patent law as it stands. See, for example, Sean B. Seymore, *Patent Asymmetries*, 49 UC Davis L Rev 963, 990–91 (2016); Sawicki, 39 Fla St U L Rev at 745 (cited in note 6); Malani and Masur, 101 Georgetown L J at 639 n 7 (cited in note 20); Lichtman and Lemley, 60 Stan L Rev at 47 (cited in note 1); William Alsup, *A District Judge's Proposal for Patent Reform: Revisiting the Clear and Convincing Standard and Calibrating Deference to the Strength of*

bringing the incentives effects of false positives into sharp relief in the next Part. But the use of this highly constrained model is not meant to suggest that the public benefits of the patent system can be maximized solely through modulations in the probability-of-patentability threshold, or that substantive patent law is necessarily set correctly as it presently stands.[34]

For reference throughout this Article, this particular approach to balancing errors in the patent system is called the "one-sided incentives" model. This term refers to the fact that, in this framework, more-liberal patent awards consistently lead to greater incentives to invent. Incentives are one-sided because more is always more: a lower probability-of-patentability threshold means more patents, which in turn means greater incentive to invent in the future. The only constraint on the patent free-for-all is the matter of cost. Future rewards would be maximized through prolific patent granting; it is only the ex post costs of those grants that create a need for balance.[35]

---

the Examination, 24 Berkeley Tech L J 1647, 1650 (2009); Merges, 14 Berkeley Tech L J at 589 (cited in note 1). See also Christi J. Guerrini, *Defining Patent Quality*, 82 Fordham L Rev 3091, 3098–99 (2014) (observing that "patent quality" is typically synonymous with validity under existing law). Proposals to encourage more patent litigation can, however, be understood as calls to increase the level of information available to the patent system.

[34] Some readers might also question why the threshold T does not depend on the objective examiner's priors about the probability that an application is valid prior to examination, as is common in total error cost models. As Professor Louis Kaplow explains, this kind of Bayesian analysis is inappropriate in cases in which future actors are likely to behave differently depending on how T is set. In other words, the ratio of valid to invalid applications coming into the patent office is likely endogenous to T. See Louis Kaplow, *Burden of Proof*, 121 Yale L J 738, 748, 783–86 (2012). An additional question some readers may have is whether the examiner's task should be viewed as an inquiry into probabilities or likelihood ratios. See Louis Kaplow, *Likelihood Ratio Tests and Legal Decision Rules*, 16 Am L Econ Rev 1, 5–13 (2014) (explaining this distinction). This discussion is framed around probabilities of patentability to enable more straight-forward comparisons of the costs of each type of error. Once these error costs are understood, a given level of scrutiny may in some cases be more profitably implemented through an analysis of comparative likelihoods. See id at 20–25, 34–36.

[35] To be clear, under this model, more-generous patent awards are only certain to increase incentives to create the first generation of a technology. At some point, overly expansive patent rights might well inhibit subsequent innovation, as future inventors would have to contend with these existing patent rights, reducing their reward for making improvements to the original technology. See Suzanne Scotchmer, *Innovation and Incentives* 134 (MIT 2004). This complication is not actually relevant here, though: the claim that liberal patent awards would "maximize" incentives to invent is made only to distinguish the two-sided incentives model presented in Part II, in which overly generous patent rewards can reduce incentives to make even the first generation of a technology.

This tug-of-war between ex ante incentives on the one hand and ex post costs on the other means that, under the one-sided incentives framework, *any* level of probability is a potentially appropriate threshold for justifying patent rights. When the patent system creates significant public benefits at low cost—that is, when I_{FN} is much greater than C—the threshold necessary to justify patent rights dives toward zero, and patent rights can be justified even for inventions that appear quite unlikely to meet the legal standards of patentability. And when the patent system has high costs and yields modest public benefits—that is, when C approaches I_{FN}—the threshold necessary to justify patent rights moves toward one, and patent rights should be available only when it appears highly likely that they are actually deserved. The one-sided incentives model has no internal bounds. On its own terms, the best policy could be to give patents to everyone or to no one.

These extremes are possible because the expected harm of erroneously awarding a patent C could be many times larger than the harm of erroneously withholding a patent I_{FN} and vice versa.[36] If the expected cost of granting a patent and the expected cost of denying a patent were approximately equal, the threshold for awarding patent protection would hover somewhere near 50 percent.[37] But there is no reason to assume that the patent system's ex post costs and ex ante incentives align in the way necessary to make this happen. It is entirely possible that one term dominates the other, and nothing requires the model to converge to its midpoint.[38]

---

[36] If $C > I_{FN}$, then the threshold for granting a patent requires more than 100 percent probability—in other words, patents should be granted to no one, no matter how certain it appears they've satisfied the patentability requirements. It might seem that these values are unlikely, given the fact that we do have a patent system, but unfortunately the available empirics do not even allow one to rule out such extreme possibilities. See note 39 and accompanying text.

[37] This occurs when the expected incremental cost of a patent grant is half the expected incentive harm from a false negative—that is, $I_{FN} = 2C$.

[38] See Kaplow, 121 Yale L J at 784–86 (cited in note 34) ("[W]e should instead be troubled by the notion that it may make sense, even as an approximation, to employ a single threshold . . . (such as fifty percent) to make important decisions in a wide range of contexts in which the consequences vary dramatically."). Note that there is sometimes an intuitive appeal to making decisions around a 50 percent probability threshold because doing so minimizes the total number of errors that are made. For example, if a coin has a 51 percent chance of coming up heads, a guesser would be correct most often by always calling "heads." But that hardly means that is the best strategy. If the penalty for incorrectly calling "tails" is losing $1 and the penalty for incorrectly calling "heads" is

Unfortunately, the empirical studies necessary to narrow this range of plausible cutoff points are sorely lacking and not likely to be completed soon. To determine even approximately how permissive to be in granting patents, one would first need to estimate the direct ex post costs attributable to an incremental patent grant. One would then need to estimate the public benefits that might be lost as the result of an individual erroneous patent denial. This is, empirically, a nonstarter. No one really knows how much public benefit is produced by the patent system *in general*.[39] Attempting to put a number on how individual decisions to grant or deny affect first ex post costs and then ex ante incentives seems an impossible undertaking. The result is that the one-sided incentives model yields essentially no conclusions as to how the patent system ought to err in cases of uncertainty.

This indeterminacy is a likely reason why the appropriate balance of errors in the patent system has been addressed so glancingly.[40] Under the conventional framework, there simply isn't much to say. More scrutiny saves ex post costs but reduces ex ante incentives; more permissiveness increases ex ante incentives and drives up ex post costs. And the magnitudes of the effects on either side of this balancing are simply unknown.

B.   How the System Errs Presently

While prevailing theory and the available empirics are equivocal as to how patent cases should be decided in the face of uncertainty, the current patent system is decidedly stacked in just one direction: in cases of doubt, patents are available. This

---

losing a finger, the same guesser might reasonably prefer to always call "tails." Minimizing the costs of errors requires considering not just the probability that a guess will prove correct but also the consequences of being wrong. See Alan Devlin and Michael Jacobs, *Antitrust Error*, 52 Wm & Mary L Rev 75, 97–98 & n 91 (2010).

[39]   See Liivak, 78 Brooklyn L Rev at 1337–38 (cited in note 9) ("Though this underlying purpose is simple to state, it has created an intractable cost–benefit analysis that resists either justification or, alternatively, falsification."); Ouellette, 101 Va L Rev at 75–84 (cited in note 9).

[40]   Indeed, the most specific conclusion prior scholars have offered is that the optimal number of false positives is not zero. See Sawicki, 39 Fla St U L Rev at 745–46 (cited in note 6); Merges, 14 Berkeley Tech L J at 593 (cited in note 1); Wagner, 157 U Pa L Rev at 2139 (cited in note 6). A number of commentators have explicitly set aside the question of how false positives should be balanced against false negatives. See Hemel and Ouellette, 92 Tex L Rev at 330–31 n 134 (cited in note 11); Mandel, 9 Yale J L & Tech at 31–32 n 129 (cited in note 5); Bock, 49 U Richmond L Rev at 448–49 (cited in note 5).

outcome is not the product of any single policy lever but rather the result of a number of procedural rules and structural features of both the examination and enforcement stage. These characteristics in combination cause the patent system to err consistently in favor of patent rights in cases of uncertainty.

The preference for patent rights begins the moment an application shows up at the door of the PTO. By long-standing rule, patent applications are reviewed with a presumption of patentability; the burden rests on the examiner to show why a patent should *not* issue.[41] The standard of proof here is moderate: the examiner need only show unpatentability by a preponderance of the evidence.[42] Nonetheless, the effect of this burden assignment is that ties are broken in favor of patentability.[43]

This presumption would assure that patents issue in close cases even if the PTO had perfect information. But its effects are compounded by the limited information available at the examination stage.[44] Because patentability often turns on the non-existence of prior art,[45] the fact that information is missing will tend to inure to the benefit of a patent applicant, making unpatentable inventions appear patentable rather than the other way around. And the information deficit at the PTO is substantial. Patent examiners on average have fewer than twenty hours to read a patent application, search the prior art, and render a written decision.[46] Moreover, entire categories of relevant

---

[41] See 35 USC § 102(a) ("A person shall be entitled to a patent unless . . . ."); *In re Oetiker*, 977 F2d 1443, 1444–45 (Fed Cir 1992).

[42] See *Oetiker*, 977 F2d at 1444–45; *In re Caveney*, 761 F2d 671, 674 (Fed Cir 1985).

[43] See *Oetiker*, 977 F2d at 1449 (Plager concurring); Sean B. Seymore, *The Presumption of Patentability*, 97 Minn L Rev 990, 997–99 (2013).

[44] See Dreyfuss, 12 Lewis & Clark L Rev at 434 (cited in note 5); Seymore, 49 UC Davis L Rev at 995–96 (cited in note 33). See also Michael D. Frakes and Melissa F. Wasserman, *Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents?: Evidence from Micro-Level Application Data*, 99 Rev Econ & Stat *8–9, 41 (forthcoming 2017), archived at http://perma.cc/8HFC-6GK6 (finding that "as examiners are given less time to review applications . . . , the less prior art they cite, the less likely they are to make time-consuming rejections, and the more likely they are to grant patents").

[45] See Frakes and Wasserman, 99 Rev Econ & Stat at *7–8 (cited in note 44).

[46] See id at *8–10 (estimating that patent examiners spend an average of nineteen hours per application); Melissa F. Wasserman, *The PTO's Asymmetric Incentives: Pressure to Expand Substantive Patent Law,* 72 Ohio St L J 379, 414 n 135 (2011) (citing an email from a PTO official stating that the average examination time allotted ranged from fourteen to thirty-two hours depending on the complexity level of the art); Ford, 164

information are largely unavailable to patent examiners. For example, an invention can become unpatentable if it is in public use or on sale for too long before an application is filed. But activities like these do not usually produce the kind of written records an examiner can discover, so unless the applicant herself knows about such activity and discloses it, the examiner likely will be left in the dark.[47] And because applications are presumed patentable, the default outcome in a case of missing information is an issued patent.

Imperfect incentives at the PTO further lower the threshold of patentability in practice. As others have noted, the agency has a financial interest in granting rather than rejecting: as a direct result of an issued patent, the PTO can expect to receive substantial renewal (or "maintenance") fees in the future.[48] These maintenance fees are essentially pure profit for the agency—the PTO gets to keep these receipts and incurs only trivial marginal costs in the process.[49] Asymmetries in the appeal process further distort the agency's incentives—a PTO decision to deny an application may be swiftly appealed to the courts, but the agency is never required to defend its decision to grant a patent.[50] So, to the extent the agency wants to avoid appeals in general and

---

U Pa L Rev at 860 (cited in note 1) (estimating a similar average examination time using a different methodology).

[47] See Christopher R. Leslie, *Antitrust, Inequitable Conduct, and the Intent to Deceive the Patent Office*, 1 UC Irvine L Rev 323, 327–28 (2011); Mark D. Janis, *Rethinking Reexamination: Toward a Viable Administrative Revocation System for U.S. Patent Law*, 11 Harv J L & Tech 1, 56–57 (1997); Kesan, 17 Berkeley Tech L J at 766–67 (cited in note 1) ("Hence, the Patent Office is unlikely to be well informed about the relevant prior art, creating an asymmetry between the patentee's information and the information possessed by the Patent Office.").

[48] See Michael D. Frakes and Melissa F. Wasserman, *Does Agency Funding Affect Decisionmaking?: An Empirical Assessment of the PTO's Granting Patterns*, 66 Vand L Rev 67, 70, 78 & n 35 (2013).

[49] Id at 79–80. See also Michael D. Frakes and Melissa F. Wasserman, *Does the U.S. Patent and Trademark Office Grant Too Many Bad Patents?: Evidence from a Quasi-Experiment*, 67 Stan L Rev 613, 629–30 (2015) (discussing how maintenance fees cross-subsidize examination activities).

[50] See Masur, 121 Yale L J at 487 (cited in note 1). To be clear, members of the public can challenge the agency's decision to grant a patent using the inter partes and post-grant review procedures created by the AIA. See 35 USC §§ 311–19, 321–29. But these are processes within the agency, not really appeals, and in any event the patent holder typically does the work of defending the agency's decision. See 37 CFR §§ 42.120, 42.220; Gregory Dolin, *Dubious Patent Reform*, 56 BC L Rev 881, 914–20 (2015).

reversals in particular, it has reason to err in favor of granting in cases of uncertainty.[51]

Examiners, too, have reasons to favor allowance over rejection. The PTO closely monitors examiner productivity using a "count system," wherein each examiner is required to complete a specific number of work units every two weeks.[52] For productivity-measurement purposes, rejections and allowances are considered equivalent, although examiners report that it is usually less work (and takes less time) to issue an allowance than to level a rejection.[53] As a result, examiners are quietly incentivized to grant rather than reject patents in close cases.

To be sure, the examiner's is not always the final word. A patent granted by the PTO can still be challenged through litigation in a federal district court, and a final judgment of invalidity will preclude any further assertion of the patent.[54] But, as at the initial examination stage, litigation burdens and structural incentives work to the benefit of patent rights in cases of doubt.

Once issued, a patent enjoys a statutory presumption of validity. Reversing an erroneous grant by the PTO requires a challenger to show that the patent is invalid by "clear and convincing evidence"—a higher bar than the "preponderance of the evidence" standard that usually applies in civil litigation.[55] This is so even if the argument for invalidity is based on evidence that the PTO demonstrably lacked at the time of examination.[56] In this way, litigation is not a straightforward mechanism to counterbalance the information deficit and pro-grant biases of

---

[51]   See Masur, 121 Yale L J at 489–99, 505–07 (cited in note 1).

[52]   See Clarisa Long, *The PTO and the Market for Influence in Patent Law*, 157 U Pa L Rev 1965, 1990–91 & n 128 (2009); Jaffe and Lerner, *Innovation and Its Discontents* at 133–38 (cited in note 1).

[53]   See Jaffe and Lerner, *Innovation and Its Discontents* at 133–38 (cited in note 1). See also Sag and Rohde, 8 Minn J L Sci & Tech at 19 (cited in note 13); Lemley, 95 Nw U L Rev at 1496 n 3 (cited in note 6); Merges, 14 Berkeley Tech L J at 590 (cited in note 1).

[54]   See 35 USC § 282(b); *Blonder-Tongue Laboratories, Inc v University of Illinois Foundation*, 402 US 313, 329–30 (1971).

[55]   See 35 USC § 282(a); *Microsoft Corp v i4i Limited Partnership*, 564 US 91, 95, 99 (2011).

[56]   See *Microsoft*, 564 US at 109–10; *Dow Chemical Co v Nova Chemicals Corp (Canada)*, 809 F3d 1223, 1227 (Fed Cir 2015) (Moore concurring in denial of rehearing en banc).

the examination stage. To the contrary, in close cases, juries are explicitly instructed to preserve validity.[57]

An additional limitation of postgrant error correction is that it relies on private parties investing in legal process. Challenges do not come automatically or for free—instead, any given challenge depends on two or more parties having some reason to see the dispute through to completion. Patent litigation is expensive, and in many cases it may be more profitable for a firm to settle a patent claim rather than fight it.[58] As a result, it is likely that some challenges to patents of questionable validity will not be litigated to final judgment, if they are brought at all.

In combination, these features allow applicants to obtain and enforce patents in at least some cases in which an objective examiner would assess the probability of patentability to be less than 50 percent.[59] A number of prior commentators have criticized this arrangement, proposing reforms so that the system will err against patent rights in questionable cases—or at least not quite so heavily in favor of them. For example, scholars have proposed things like revising the presumption of patentability, changing the compensation structure for patent examiners,

---

[57] See Lichtman and Lemley, 60 Stan L Rev at 47–48 (cited in note 1). See also Dreyfuss, 12 Lewis & Clark L Rev at 434–35 (cited in note 5) (discussing the combination effects of the presumption of validity and juror biases).

[58] For analysis of the private incentives to challenge patents, see Farrell and Merges, 19 Berkeley Tech L J at 948–55 (cited in note 1); Jay P. Kesan and Andres A. Gallo, *Why "Bad" Patents Survive in the Market and How Should We Change?—The Private and Social Costs of Patents*, 55 Emory L J 61, 80–85 (2006); Sag and Rohde, 8 Minn J L Sci & Tech at 22–28 (cited in note 13). See also Jay P. Kesan and 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, 265–69 (2006) (collecting data on case dispositions, including settlements).

[59] See Seymore, 49 UC Davis L Rev at 971–73 (cited in note 33). There are a couple of features of the patent system that might somewhat mitigate these pro-patent biases. First, during patent examination, claims are given their "broadest reasonable construction," 37 CFR § 42.100(b), which can sometimes result in a claim being denied even though it might have been upheld in litigation. See Meurer, 51 Wm & Mary L Rev at 702–03 (cited in note 27). But the effect of this rule is likely small, given that applicants can usually amend their claims to embrace the narrower (valid) meaning explicitly. See *Cuozzo Speed Technologies, LLC v Lee*, 136 S Ct 2131, 2145 (2016). Second, the AIA created new administrative review procedures in which a patent may be revoked under a preponderance of the evidence standard—thus avoiding the heightened presumption of validity that issued patents normally enjoy. See AIA § 6(a), (d), 125 Stat at 302–03, 308–09, 35 USC §§ 316(e), 326(e). However, these procedures still require the challenger to carry the burden of showing unpatentability, and are themselves to subject to various procedural and substantive limitations. See note 199.

tweaking the statutory presumption of validity, and encouraging more patent litigation.[60] All of these have been rooted in a goal of rebalancing the distribution of errors *against* patentability—that is, to require a greater probability that an invention really is patentable to justify the costs of patent protection.

But, problematically for these proposals, it is not clear that the current bias in favor of patent protection is actually suboptimal. As discussed above, under the conventional, one-sided incentives model, the probability of patentability necessary to justify patent rights could be anywhere between zero and one.[61] In fact, it is possible that despite the structural and procedural biases toward patent rights, the current system actually makes it *too difficult* to obtain patent protection. If the probability of patentability needed to get a patent is currently 0.4, maybe it should be 0.3.

As a result, there is a logical step missing in the recent calls to rebalance the errors of the patent system in a more patent-skeptical direction. It is true that some patents issue even though they are more likely invalid than valid. It is also true that some of these patents may be successfully enforced, or otherwise impose ex post costs, despite objective probabilities of patentability below 50 percent. But without relying on disputed priors about the costs and benefits of patent protection, it is not necessarily true that increasing precautions against undeserved patent rights would be socially beneficial. It could just as easily work harm.[62]

II. ACCURACY AND INCENTIVES

This Part explores the role that accuracy plays in determining the rewards offered by the patent system. The central insight here is that the incentives created by the patent system depend on accuracy in two directions: the probability of awarding

---

[60]   See, for example, Seymore, 97 Minn L Rev at 1022–31 (cited in note 43) (proposing shifting the burden of persuasion in patent examination); Merges, 14 Berkeley Tech L J at 607–09 (cited in note 1) (proposing changing the compensation structure for examiners); Miller, 19 Berkeley Tech L J at 704–11 (cited in note 22) (proposing offering a bounty to encourage more patent litigation); Lichtman and Lemley, 60 Stan L Rev at 59–65 (cited in note 1) (proposing weakening the presumption of validity); Alsup, 24 Berkeley Tech L J at 1649–54 (cited in note 33); Burstein, 83 Geo Wash L Rev at 538–48 (cited in note 1) (proposing expanding standing to enable more patent challenges).

[61]   See notes 27–35 and accompanying text.

[62]   See Golden, 89 Tex L Rev at 1068 (cited in note 4).

patents when they are deserved and the probability of withholding them when they are undeserved. This fundamentally alters the optimal balance between false positives and false negatives, and yields a "two-sided incentives" model that suggests a significantly higher probability of patentability should be required to justify patent rights than the conventional one-sided incentives model would indicate.

Parts II.A and II.B start by introducing the role that accuracy plays in determining the ex ante incentives created by prize and punishment systems in general. Part II.C then revises the conventional account of error costs in the patent system to include this understanding of the value of accuracy. Finally, Part II.D shows that the answer to the question *how should the patent system err?* depends significantly on whether patent law is seeking to influence mutually exclusive choices, a condition that has not previously received much attention.

A. Marginal versus Absolute Rewards

To understand the role that accuracy plays in the patent system, it is helpful to take a step back and explore the theory for how patent rights provide public benefits in the first place. According to the dominant account, the purpose of awarding patents is to incentivize future investments in research and development.[63] Patents are a form of public subsidy to encourage private actors to engage in a specific socially desirable activity.[64] In this way, patents are not so different from prizes or even punishments—all are publicly funded mechanisms to encourage private behavior that benefits the public.[65]

Prize and punishment systems operate using the same fundamental model. At heart, they are based on either a promise or

---

[63] See note 17 and accompanying text. While the goal of rewarding invention is the most commonly accepted justification for the patent system, many of the principles expounded in this Part can be applied to other functions of the patent system as well. See Parts III.B–C.

[64] See Peter S. Menell and Suzanne Scotchmer, *Intellectual Property Law*, in A. Mitchell Polinsky and Steven Shavell, eds, 2 *Handbook of Law and Economics* 1473, 1476–77 (Elsevier 2007); Benjamin N. Roin, *Intellectual Property versus Prizes: Reframing the Debate*, 81 U Chi L Rev 999, 1028 (2014).

[65] See Hemel and Ouellette, 92 Tex L Rev at 312 (cited in note 11); Roin, 81 U Chi L Rev at 1021 (cited in note 64). Unlike patents and criminal penalties, prizes can be offered by nongovernmental entities as well. See Michael J. Burstein and Fiona E. Murray, *Innovation Prizes in Practice and Theory*, 29 Harv J L & Tech 401, 419–23 (2016).

a threat, as in: "If you do X, you will get Y." In the case of a prize, X is usually some socially desirable activity, and Y is some privately valuable reward—as in, "If you cure cancer, you will receive $100 million." In the case of a punishment, X is some socially *un*desirable activity, and Y is some privately dreaded result—as in, "If you commit plagiarism, you will be expelled." Patents are a complex form of prize, in which the offer is, "If you make an invention that satisfies the patentability criteria, you will receive some time-limited exclusive rights, the value of which will depend (at least in part) on the value of your invention."[66]

A critical component of threats and promises like these is often left unspoken. Implicit in the offer, "If you do X, you will get Y," is a promise to carry out the inverse statement: "If you do *not* do X, you will *not* get Y." For purposes of inducing the desired conduct, this silent, negative promise is as important as the articulated, affirmative promise. The logic here is straightforward: if the target of the promise will receive Y either way, then offering Y provides no additional incentive to do X. The lure of the offer depends on Y being bestowed if *and only if* the target does X.

The effectiveness of a prize or punishment regime depends on reliable enforcement of both the affirmative and the negative promises. Consider a prize that offers a fixed payout S to anyone who performs some specified task.[67] The lure of such a prize will depend, first, on the size of the prize itself and, second, on an actor's chance of receiving the prize if she in fact performs the task—call this probability p_A. Holding the size of the prize constant, the higher the probability that the prize will be awarded when deserved, the greater the incentive created by the offer of the prize. (Conversely, the higher the probability that the prize will be wrongfully withheld, the smaller the incentive created by the prize offer.) An actor who performs the specified task receives S with probability p_A, for an expected value of $p_A * S$.

But there is an important complication lurking here. The expected value given in the prior paragraph describes the absolute reward the actor can expect to receive by performing the specified task. But it does not give the actor's marginal reward

---

66 See Adam Smith, *Lectures on Jurisprudence* 83 (Oxford 1978) (R.L. Meek, D.D. Raphael, and P.G. Stein, eds) (suggesting that patents are preferable to simple prizes because their value can change depending on the value of the underlying invention).

67 Importantly, in the patent system S is *not* fixed—a complication that will be dealt with below.

for performance, because it does not account for the possibility that she might (erroneously) be given the prize either way. Suppose that, even if she does not perform the specified task, the target of the promise has some chance p_0 of receiving the prize by mistake. The additional (or marginal) reward offered to perform the desired conduct is the difference between the target's expectations when she performs the task and her expectations when she does not perform the task. In the terms given above, the marginal reward M is given by $p_A * S - p_0 * S$, or:

$$(2) \qquad M = S(p_A - p_0)$$

Examining this equation, it becomes clear that there are not two but three levers a prize administrator can pull to maximize the lure created by a prize: (1) increase the size of the payout S; (2) increase the probability p_A of awarding the prize when it is deserved; and (3) *decrease* the probability p_0 of awarding the prize when it is not deserved. While it is perhaps easiest to capture public attention by offering a prize of large absolute magnitude (lever #1),[68] accuracy in awarding that prize (levers #2 and #3) can affect marginal rewards just as much. Indeed, in an extreme case, in which the probability of receiving the prize is the same whether or not the desired task is performed, the offer will provide no marginal reward at all, no matter how large its absolute dollar value may be.[69]

These three levers will be familiar to scholars of criminal law. I.P.L. Png recognized an analogous relationship between false convictions and false acquittals back in 1986.[70] In fact, the general point that false convictions (the corollary to undeserved prize awards) can impair marginal deterrence is now well accepted

---

[68]   See, for example, *Philips Claims L Prize Victory*, 41 Lighting Design & Application 10, 10 (Sept 2011) (reporting Philips's victory of a $10 million prize for a 60-watt LED replacement lightbulb); Steve Russell, *DARPA Grand Winner: Stanley, Stanford University's Robot Car,* 183 Popular Mechanics 36, 36–39 (Jan 2006) (describing a $2 million prize for creation of an autonomous vehicle); *$25,000 Orteig Prize Presented to Flier* (NY Times, June 17, 1927).

[69]   This is the case of $p_A = p_0$. When this is true, $S(p_A - p_0)$ will be zero no matter the size of S.

[70]   See I.P.L. Png, Note, *Optimal Subsidies and Damages in the Presence of Judicial Error*, 6 Intl Rev L & Econ 101, 101–02 (1986).

in the law-and-economics literature.[71] But when it comes to *encouraging* certain conduct (such as with patents or prizes), the incentive effects of undeserved rewards have been largely overlooked.[72]

To illustrate this principle more concretely, imagine public health officials want to encourage the world's leading medical researchers to focus their efforts on creating a vaccine against the Zika virus. One way to do this would be to offer a $1 million prize to anyone who creates a successful vaccine. As a first step to making the prize effective, the administrators would want to make sure their promise to award the prize to those who deserve it is credible.[73] This much is obvious—if prospective researchers do not trust the prize administrators (or think they will otherwise be wrongfully denied the award), they will discount the possibility of receiving the prize after making their investments in the vaccine. Reducing the risk that the prize will be wrongly denied after a researcher has sunk costs to develop the vaccine will increase the absolute expected reward available from such an investment, given by $p_A * S$.

But, in addition, it is important for the prize administrators to make a credible commitment *not* to award a prize to someone who has *not* created a Zika vaccine. For example, the prize administrators might want to announce the testing criteria they are planning to use to ensure they do not give a prize based on a vaccine that does not stop Zika—such as a vaccine that reduces the appearance of Zika-like symptoms, but has no effect on Zika

---

[71]   See A. Mitchell Polinsky and Steven Shavell, *Legal Error, Litigation, and the Incentive to Obey the Law,* 5 J L Econ & Org 99, 100 (1989); Louis Kaplow and Steven Shavell, *Accuracy in the Determination of Liability*, 37 J L & Econ 1, 2–3 (1994); Joel Schrag and Suzanne Scotchmer, *The Self-Reinforcing Nature of Crime*, 17 Intl Rev L & Econ 325, 326 (1997); Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 Stan L Rev 1477, 1483–84 (1999) ("[G]reater accuracy in the determination of guilt increases the returns to being innocent."). But see Daniel Epps, *The Consequences of Error in Criminal Justice,* 128 Harv L Rev 1065, 1126–28 (2015) (challenging the premises of this argument in the criminal context).

[72]   Patent scholars have, however, suggested that undeserved patent rewards may inadvertently encourage wasteful activity—"diverting [ ] resources out of productive activities and into the 'patent game.'" Merges, 14 Berkeley Tech L J at 592 (cited in note 1). This is a harm to future ex ante incentives as a result of erroneous patent grants, but it is not the same as a reduction in future ex ante incentives to invent.

[73]   See Burstein and Murray, 29 Harv J L & Tech at 413–14, 444 (cited in note 65) (observing the importance of commitment mechanisms); Roin, 81 U Chi L Rev at 1067–68 (cited in note 64); Menell and Scotchmer, *Intellectual Property Law* at 1532 (cited in note 64).

itself. If a flu researcher thinks she has a good chance of pawning off an existing flu vaccine as a Zika prophylactic without the prize administrators being able to tell the difference, then she will have little incentive to drop her current research and switch to Zika. Efforts to avoid giving the prize in error will decrease the expected prize winnings available without actually making the desired vaccine, given by $p_0 * S$. This in turn increases the marginal reward for stopping Zika and thus the effectiveness of the prize regime.

The incentive-reducing effects of erroneous awards can be dramatic. For example, suppose a researcher estimates she has a 60 percent chance of claiming the prize if she pauses her flu research and instead focuses her efforts on a Zika inoculation (that is, $p_A = 0.6$). A 60 percent chance of claiming a $1 million prize might seem like a strong inducement. But this absolute reward is not the complete picture. To determine how much the researcher has to gain by researching Zika, one must consider her probability of receiving the prize undeservedly. If the odds of successfully pawning off a flu vaccine as a Zika vaccine are 50 percent (that is, $p_0 = 0.5$), then the marginal expected reward available by focusing on Zika is only $100,000.[74] So, in the end, the seemingly large $1 million absolute prize actually provides a (comparatively small) $100,000 of marginal reward.[75]

B. The Need for a Mutually Exclusive Choice

The distinction between marginal and absolute rewards might seem a simple point, but it is actually not as straightforward as it first appears. In fact, there is a subtle but critical limitation at work here that has previously escaped attention: the difference between marginal and absolute rewards is relevant only when a prize is being offered to influence a mutually exclusive choice. Without a mutually exclusive decision at stake, the

---

[74]   As discussed above, the expected prize reward available by focusing on Zika is given by $S(p_A - p_0)$. In this example, $S = \$1$ million, $p_A = 0.6$, and $p_0 = 0.5$, so her expected reward for pursuing Zika is $100,000.

[75]   The flip side of this example is that precautions to avoid undeserved awards can go a long way. For example, if the prize administrator can introduce a test that would reduce the probability of an erroneous prize award to 10 percent, the original $1 million cash prize would provide $500,000 in marginal reward. Separately, these increased precautions will result in fewer prize payments overall—a savings in the "ex post costs" category discussed below.

distinction between marginal and absolute rewards collapses, and mistaken awards do not have the incentive-reducing effect described above.

To illustrate, return to the researcher considering whether to seek a $1 million prize for developing a Zika vaccine. As before, assume the competition is run rather sloppily: focusing on Zika gives the researcher a 60 percent chance of claiming the prize, while instead continuing her flu research gives her a 50 percent chance of claiming the prize undeservedly. For the reasons discussed above, these haphazard prize awards result in only weak incentives to stop work on the flu and instead switch to Zika.

But suppose this particular researcher can have it both ways. For example, say the Zika vaccine prize program permits multiple entries and promises a separate reward to each chemically distinct submission that appears to reduce the rate of Zika infection. And suppose further that this researcher has the time and resources to pursue *both* Zika and the flu simultaneously. The decision to develop and submit a Zika vaccine is thus one she can make independently of the decision to develop and submit a flu vaccine.

In this situation, the possibility of an undeserved award does not diminish the researcher's marginal rewards for pursuing Zika. Consider each step of the researcher's decision-making process sequentially. Putting aside ethical or reputational concerns, the decision to submit a flu vaccine to the contest is an easy one: she was planning to develop it anyway, and attempting to claim the prize has an expected value of $500,000.[76] With that decision made, the researcher must then choose whether to develop an actual Zika vaccine. Doing so brings her a (separate) 60 percent chance of claiming the $1 million prize, for an expected value of $600,000. In this case, that is *both* the absolute and marginal reward available to her from seeking the Zika prize legitimately. Notably, the magnitude of the marginal reward here does not change depending on her chances of winning the prize undeservedly. Without a mutually exclusive choice at stake, erroneous rewards do not affect the researcher's incentives to perform the desired activity.

---

[76]   The expected value is $500,000 because, as stated above, the researcher has a 50 percent chance of claiming the $1 million prize undeservedly.

The difference between this case and the one presented in the prior Section is the existence of opportunity cost. When an actor faces a mutually exclusive choice, selecting one path means forfeiting the ability to pursue the other paths. The value that could have been obtained by selecting one of those other paths is an opportunity cost.[77] In the case of a vaccine researcher who must pick between pursuing Zika or the flu, spurious rewards for the flu raise the opportunity cost of selecting Zika, thereby reducing the very marginal incentives that the prize was intended to create. But, critically, this effect is present *only* when the researcher must choose between one path and the other. Without a mutually exclusive decision, there can be no opportunity costs,[78] and the incentives to select the desired path are the same whether or not undesired paths also receive some unintended rewards. If the researcher can pursue both the flu and Zika simultaneously, the incentive harm from spuriously awarding the flu vaccine disappears.[79]

As a result, the consequences of an error depend significantly on whether the prize is being offered in hopes of swaying a mutually exclusive decision. Erroneous awards always impose extra costs—either way the prize administrators have written an unneeded $1 million check—but there is an additional incentive harm present only in cases involving mutually exclusive choices.

C. False Positives in the Patent System

Although the patent system is in some important ways different from the simple prize system described in the prior Section, false positives have the same basic potential to reduce the marginal rewards offered to successful inventors. This Section revisits the traditional account of error costs in the patent system in

---

77   See David W. Pearce, ed, *The MIT Dictionary of Modern Economics* 315 (4th ed 1992) ("[T]he opportunity cost of an action is the value of the foregone alternative action.").

78   See id ("Opportunity cost can only arise in a world where the resources available to meet wants are limited. . . . If resources were limitless no action would be at the expense of any other . . . and the opportunity cost of any single action . . . would be zero."). See also Steven N. Durlauf and Lawrence E. Blume, eds, 6 *New Palgrave Dictionary of Economics* 198–201 (Macmillan 2008).

79   Though framed in different language, the criminal-deterrence literature has struggled with a similar complication—that it is difficult to say what counts as a distinct "opportunity" to commit a crime. See generally Henrik Lando, *Does Wrongful Conviction Lower Deterrence?*, 35 J Legal Stud 327 (2006).

light of this previously unappreciated harm caused by erroneous patent grants.

The move from the prize context to the patent context introduces a number of complications. One important wrinkle—explored in detail in the next Part—is that there is not one single *thing* (like curing Zika) patent law asks inventors to do. Rather, the patent system presents inventors with an elaborate framework of rewards (and, sometimes, punishments) to encourage them not simply to create inventions, but to invest in specific kinds of research and development, to carry out that work in acceptable ways, and to disclose the results of that work to the public in a transparent and timely fashion. The multifaceted nature of the patent system's goals will result in several caveats and conditions in the final analysis.[80]

But before introducing these nuances, a simple example may help to illustrate in broad strokes how false positives can affect incentives to invent. Consider a manager choosing how to allocate a firm's research and development budget. One of the manager's options (call it Plan A) is to pursue a program of small-scale, incremental improvements to an existing technology. By assumption, the capital outlay required to perform this kind of work is small. And, because these will be improvements to an existing technology, the firm may be able to use its incumbent advantages to capture much of the benefits of its investment. As a result, the firm may be able to recover the costs of Plan A without patents even coming into the picture.[81]

The manager's other option (Plan B) is to pursue an ambitious, disruptive technology. From a social perspective, this groundbreaking project is a better use of resources than the program of incremental improvements—Plan B costs more, but promises much greater benefits. Nonetheless, because this

---

[80]   See Part III; Part IV.

[81]   For example, suppose the project will cost the firm $100,000 and that the firm will enjoy private rewards of $500,000 through its incumbent/first-mover advantage. An additional $200,000 of social benefits will not be captured by the firm. Putting aside patent rewards and opportunity costs, the manager would choose to undertake Plan A, because the private benefits exceed the private costs. The numbers selected throughout this example are for purposes of illustration only. It should also be noted that this example draws heavily from the "carrots and sticks" theory of obviousness, see Meurer and Strandburg, 12 Lewis & Clark L Rev at 558–65 (cited in note 5); Lunney, 7 Mich Telecom & Tech L Rev at 404–12 (cited in note 8), which bring complications of its own, see Part III.A.

would be a disruptive technology, the firm will not be able to capture nearly as much benefit from its investment as it would under Plan A. As a result, it is possible that without some form of subsidy, the firm would not be able to recover the costs necessary to bring about Plan B's groundbreaking innovation.[82]

One reason for offering patent protection is to encourage managers like this one to undertake this latter, more innovative kind of project. Patents do this, the theory goes, by offering inventors who solve particularly challenging problems a share in the benefits that they would not otherwise capture.[83] The promise of a patent reward can push a socially valuable but privately unprofitable project like Plan B into the black, thereby steering private capital in the direction of more ambitious undertakings.[84]

But it might not be enough to make such inventions profitable in a vacuum. If the firm manager can pick only one of these two research paths, the question is not simply whether Plan B can be made profitable, but whether it can be made *more profitable* than Plan A. In order to encourage groundbreaking innovation, the patent value offered to highly inventive projects must be sufficiently large to offset the opportunity costs of foregoing less-inventive projects.[85] This may require subsidizing Plan B at a level that goes beyond making the project narrowly profitable for the firm.

In addition to affecting the magnitude of the needed reward, the goal of encouraging highly inventive over less-inventive projects heightens the need for accuracy in the administration of that reward. No matter how generous the patent prize, the marginal incentive to choose Plan B over Plan A will depend on

---

[82] For example, suppose the project will cost the firm $300,000 and that the firm will enjoy private rewards of $250,000 through its incumbent/first-mover advantage. An additional $1 million of social benefits will not be captured by the firm. Without patent protection or some other inducement, the manager would not choose to undertake this project because the private costs exceed the private benefits.

[83] See Lunney, 7 Mich Telecom & Tech L Rev at 412 (cited in note 8).

[84] For example, suppose the offer of patent protection increases the firm's expected return from Plan B by $100,000. Assuming no opportunity costs, the firm will now have sufficient incentives to undertake the project, because the expected rewards ($350,000) exceed the expected costs ($300,000).

[85] And indeed, the patent protection described in note 84 provides inadequate incentives for the firm manager to select Plan B over Plan A. Plan A remains the more profitable option to the firm, because it yields expected rewards of $500,000 and costs only $200,000, while even with the promise of patent protection Plan B yields expected rewards of $350,000 and costs $300,000.

successful discrimination between the two projects. If the firm manager expects that *both* the incremental improvement and the disruptive technology will receive patent protection (and that protection will have equal value),[86] then the promise of patent rights will result in no additional incentive to choose the latter.[87] Creating marginal incentives to pursue Plan B instead of Plan A requires accuracy not only in awarding patents to firms investing in projects like Plan B—something prior commentary takes as a given—but also in denying patents to firms investing in projects like Plan A.[88]

As with simple prizes, this effect depends on the existence of a mutually exclusive choice. If the manager has the option of pursuing both research paths simultaneously, then patent law can fully satisfy its mission simply by making the groundbreaking innovation path (Plan B) profitable for the firm, without attending to its value in comparison to Plan A. If there is no mutually exclusive choice at stake, marginal rewards are equal to absolute rewards, and the conventional, one-sided incentives model presented in Part I.A correctly balances errors between false positives and false negatives.

But in cases in which a mutually exclusive choice *is* present, the conventional account overlooks an important harm flowing from patents that are mistakenly granted. When undeserving inventions are given protection, the system fails to keep its implicit, negative promise: that ordinary incremental improvements (like those of Plan A) will *not* be rewarded by a grant of exclusive rights. In other words, false positives increase the probability p_0 that a future inventor will receive a

---

[86] It's possible that a patent on the disruptive technology might naturally be worth more than a patent on the incremental improvement. But the opposite could be true as well, especially given the risks attendant to emergent, disruptive technologies. For a discussion of the consequences of this complication, see note 117.

[87] For example, if the manager expects that both Plan A and Plan B will earn patent protection worth $400,000, her calculation will remain unchanged.

[88] For example, offering patent protection worth $400,000 to Plan B and no patent protection to Plan A would, at last, create private incentives to pursue the more socially valuable project. With the patent offer, Plan B's expected rewards ($650,000) would not only exceed its expected costs ($300,000), but would also offer sufficiently large profits to justify selecting Plan B over Plan A (which continues to offer $500,000 of rewards at a cost of $200,000).

patent undeservedly, thereby reducing marginal rewards to pursue highly inventive projects in the future.[89]

Once these effects on marginal incentives are considered, the potential outcomes of the decision to grant or deny patent protection can be summarized as follows:

TABLE 2. TWO-SIDED INCENTIVES MODEL

Decision	Invention Patentable	Invention Not Patentable
Award Patent	– Imposes Ex Post Costs (True Positive)	– Imposes Ex Post Costs – *Reduces Ex Ante Incentives* (False Positive)
Deny Patent	– Reduces Ex Ante Incentives * No Ex Post Costs (False Negative)	* No Ex Post Costs (True Negative)

As Table 2 illustrates, the decision to award a patent (top row) imposes ex post costs whether or not the patentability requirements have been satisfied. But granting a patent undeservedly imposes an additional harm missing from the conventional account: a reduction in future ex ante incentives. In this way, a false positive is a doubly expensive mistake: it both imposes ex post costs and reduces ex ante incentives.

This framework is dubbed the "two-sided incentives" model, because it recognizes that ex ante incentives are not only affected by errors in a single direction. Instead, when a mutually exclusive choice is at stake, ex ante incentives depend on *both* the successful granting of patents when they are deserved and the successful denial of patents when they are not deserved.

Unsurprisingly, the addition of this new error cost term in the case of a false positive changes how the patent system ought to err in cases of uncertainty. As before, consider an objective examiner applying the substantive patentability requirements to the facts known to her to estimate the probability q that a specific application satisfies those requirements. The task of the patent policymaker is to set a threshold T' for all patent

---

[89] As with false negatives, the relationship between any particular false positive and future incentives is indirect—the question is how *future* prospective inventors can expect to be treated. Ex ante harms do not come as an immediate result of individual errors, but from changes to the overall expected rate of false positives and false negatives. See notes 19–21 and accompanying text.

applications, such that applications with $q > T'$ will result in patent rights and applications with q equal to or below that threshold will be denied patent rights.

The ideal probability threshold T' can be determined by comparing the expected consequences of an action in either direction. Whether deserved or not, granting a patent can be expected to impose ex post costs—these estimated incremental costs remain C. But, depending on the merits of the case, there may be an additional harm to granting as well. When the patentability requirements were not actually satisfied, a grant increases prospective inventors' expectations of receiving patent rights without deserving them, thereby reducing marginal rewards in the future. The resulting diminished ex ante incentives reduce the public benefits of the patent system similar to the way false negatives do—call this analogous incentive harm from a false positive I_{FP}. The decision to grant thus incurs social costs equal to $C + (1 - q) * I_{FP}$—the ex post costs incurred as a result of the patent grant, plus the probability-discounted harm to ex ante incentives as a result of an incremental false positive.[90]

Consistent with the one-sided incentives model, denying a patent results in no ex post costs, but produces an effect on ex ante incentives that depends on whether or not that decision was correct. If the invention really was patentable, a denial reduces prospective inventors' expectations of receiving patent rights when they deserve them in the future, and the public is harmed as a result of these diminished ex ante incentives—the same I_{FN} term introduced above. The expected cost of denying patent rights can thus be written as $q * I_{FN}$—the probability-discounted harm to future ex ante incentives in the event of a false negative.

Patent rights should be awarded when the expected cost of doing so is less than the expected cost of denying them. Under the terms given above, this is satisfied when $C = (1 - q) * I_{FP} < q * I_{FN}$, which means that a patent should be available if and only if $q > T'$, where

---

    90   As with C and I_{FN}, the objective examiner is not permitted to make any application-specific determinations of I_{FP}. Concerns that certain kinds of patent applications may affect costs or incentives in atypical ways must be channeled through generally applicable patent law. See note 31.

(3) $$T' = \frac{C + I_{FP}}{I_{FN} + I_{FP}}$$

Comparing this minimum probability of patentability T' to the minimum probability of patentability T derived from the one-sided incentives model reveals many similarities but also some important differences. As before, the probability necessary to justify awarding patent protection turns on the relationship between the ex post costs C of granting an additional patent and the public harm I_{FN} from reduced ex ante incentives following an erroneous denial. But the probability necessary to justify awarding patent protection also turns on the public harm I_{FP} from reduced ex ante incentives following an erroneous *grant*. This term is missing from the conventional, one-sided incentives model.

The probability of patentability necessary to justify patent rights under the two-sided incentives model can be simplified significantly with a few assumptions discussed in the next Section. But even the unsimplified form of equation (3) reveals an intuitive principle that has often been overlooked by prior analysis of errors in the patent system: it would be important to examine and deny patent applications even if individual patents imposed no ex post costs at all.

To see this, suppose that, somehow, the patent system was able to offer privately valuable rights at zero public cost (that is, $C = 0$). On these facts, the one-sided incentives model would call for granting a patent anytime the probability of patentability was greater than zero—essentially, giving a patent to anyone who asked for one. But such promiscuity would be folly, because giving everyone a patent regardless of merit would drive the marginal reward for *earning* a patent down to zero too.[91] Doing so would be the patent system equivalent of giving every student in the class a gold star—a poor incentivization technique, even in a classroom with an effectively limitless supply of gold stars.

The two-sided incentives model takes account of these effects, and thus calls for enforcement of the patentability requirements even as the ex post costs of patents disappear. So long as false positives result in *some* harm to future incentives

---

[91]   To be clear, this practice would drive the marginal reward from the *patent system* down to zero. Firms may nonetheless have nonpatent incentives to invest in invention. See Hemel and Ouellette, 92 Tex L Rev at 310 (cited in note 11).

(that is, $I_{FP} > 0$), granting patents to everyone will never be the right answer, even in a world of costless patents. As this model recognizes, withholding rewards when they are undeserved is not simply a matter of economizing ex post costs; it is critical for producing the patent system's public benefits.

D. Reassessing the Balance of Errors

The incentive-reducing effects of false positives do more than change how the patent system ought to err in extreme hypotheticals, such as the zero-cost patent system discussed in the prior Section. In fact, these previously overlooked effects can have a significant impact on the optimal balance of errors across a range of plausible scenarios.

Comparing the minimum probability threshold of the two-sided incentives model T' to that of the one-sided incentives model T, the former is consistently more demanding: a higher probability of patentability will always be necessary to justify patent rights once the incentive harms of false positives are taken into account. This result relies on only a minimal set of assumptions: that both the incentive harm from a false positive and the incentive harm from a false negative are greater than zero, and that the costs and benefits of the patent system are such that it ever makes sense to grant patents.[92] So long as the one-sided incentives model would set the minimum probability of patentability somewhere below 100 percent, recognition of the incentive harms of false positives would push that threshold higher still.

To say exactly how much more demanding the two-sided incentives model will be than the one-sided incentives model,

---

[92] Formally, these assumptions can be written as: (1) $I_{FN} > 0$; (2) $I_{FP} > 0$; and (3) $I_{FN} > C$. The first two assumptions mean that both false positives and false negatives produce *some* harmful effect on future incentives (however small). The third assumption is necessary to exclude the trivial case in which the ex ante incentive harm from denying a deserved patent is smaller than the expected ex post costs of granting a patent—trivial, because in that case both the one-sided and two-sided incentives models would command denying patent rights regardless of the probability of patentability. To see that the two-sided incentives model will generally demand more certainty than the one-sided incentives model, begin with this third assumption that $I_{FN} > C$. Observe that one can manipulate the assumption algebraically so that $I_{FN} * I_{FP} > C * I_{FP}$ and $I_{FN} > (C * I_{FP})/I_{FN}$. Likewise, $C + I_{FP} > C + (C * I_{FP})/I_{FN}$, and then $C + I_{FP} > (C * I_{FN} + C * I_{FP})/I_{FN}$, so that $(C + I_{FP})/(I_{FN} + I_{FP}) > C/I_{FN}$, which is equivalent to $T' > T$. Thus, on these assumptions, the minimum probability of patentability under the two-sided incentives model T' will always be greater than the minimum probability of patentability under the one-sided incentives model T.

one would need to make some assumptions about these values. But across a range of plausible estimates for the incremental costs and benefits of patent protection, the difference is a material one. This can be shown by adopting the baseline case where $I_{FP} = I_{FN}$[93] and plugging a few sample values of C/I_{FN} into the respective models:

TABLE 3. PROBABILITY-OF-PATENTABILITY THRESHOLDS

$\dfrac{C}{I_{FN}}$	Minimum probability of validity required, one-sided incentives model[94] (%)	Minimum probability of validity required, two-sided incentives model[95] (%)
0.125	12.5	56.25
0.25	25.0	62.5
0.375	37.5	68.75
0.5	50.0	75.0
0.625	62.5	81.25
0.75	75.0	87.5
0.875	87.5	93.75

As this table illustrates, recognizing the effect of false positives on marginal rewards necessarily elevates the minimum probability of patentability required to justify granting a patent. For example, when an incremental patent's ex post costs are low compared to the ex ante benefits that can be lost through an error, as in the case in which $C/I_{FN} = 0.25$, the one-sided incentives model counsels promiscuity: patents should be granted if there is a 25 percent chance the requirements have been

---

[93]   The baseline case is the weakest form of the assumption that the incentive harm from a false positive is at least as large as the incentive harm of a false negative. See notes 96–105 and accompanying text (discussing the basis for this assumption). Relaxing this assumption would affect the magnitude of these values, but would not change the more general point that the two-sided incentives model is consistently more demanding.

[94]   These probability thresholds in this column are determined by plugging sample values for C/I_{FN} into equation (1).

[95]   These probability thresholds in this column are determined by plugging sample values for C/I_{FN} into equation (4), which is the simplified form of equation (3) for the case where $I_{FP} = I_{FN}$. See note 96 and accompanying text.

satisfied. On the same facts, the two-sided incentives model would impose a significantly higher threshold to justify a patent: a 62.5 percent probability of patentability. A similar effect occurs in the moderate case, where $C/I_{FN} = 0.5$: while the one-sided incentives model would suggest granting at the 50–50 point, the two-sided incentives model would require at least 75 percent probability of patentability. And even when the ex post costs of an extra patent are high compared to the ex ante benefits that can be lost through an error (that is, $C/I_{FN} = 0.75$), both models suggest heightened scrutiny, but the two-sided model again suggests even more. A 75 percent probability of patentability would be enough for the one-sided model, but 87.5 percent would be required under the two-sided model.

These sample values hint at a larger conclusion of the two-sided incentives model that can be both generalized and formally derived: if the incentive harm of a false positive is greater than or equal to the incentive harm of a false negative, so that $I_{FP} \geq I_{FN}$, then patent rights should be granted only if it is more likely than not that they are deserved.

This can be demonstrated by simply plugging the weakest form of the condition, $I_{FP} = I_{FN}$, into equation (3). In that case, the minimum probability necessary to justify patent rights in the two-sided incentives model T' becomes:[96]

$$(4) \qquad\qquad T' = 0.5 + \frac{C}{2I_{FN}}$$

The minimum probability of patentability to justify patent rights T' is thus at least 50 percent, even when the ex post costs of patent protection C are zero. Once patent rights impose ex post costs (as they surely do), the minimum probability of patentability to justify patent rights necessarily goes up from there.

This conclusion depends, of course, on the condition that the incentive harm of a false positive is greater than or equal to the incentive harm of a false negative, that is, $I_{FP} \geq I_{FN}$. But there

---

[96]   The proof that T' takes this value when $I_{FP} = I_{FN}$ is trivial; equation (3) immediately reduces to equation (4) with this substitution. To confirm that T' goes only up from there as I_{FP} begins to exceed I_{FN}, one can take the derivative of T' with respect to I_{FP}, or simply plug in a test case. For example, when $I_{FP} = 1.5 * I_{FP}$, T' becomes $0.6 + C/(2.5 * I_{FN})$. This is even more demanding than the threshold given by (4), with a minimum probability of patentability of 60 percent when $C = 0$. Thus the minimum probability of patentability goes only up from 50 percent as I_{FP} begins to exceed I_{FN}.

are reasons to think that this will often be the case. First, notice that both I_{FP} and I_{FN} represent a public harm caused by the same basic mechanism. A false positive reduces marginal rewards by increasing prospective inventors' expectation of the probability of receiving a patent when they do not deserve one, while a false negative reduces marginal rewards by reducing prospective inventors' expectation of the probability of receiving a patent when they do deserve one. If each kind of mistake affects prospective inventors' probability estimates by an equivalent amount, then each kind of mistake should have the same effect on marginal rewards and impose the same harm on the public.[97] And, conversely, if one kind of mistake has a *larger* effect on prospective inventors' probability estimates than the other, then that kind of mistake will impose a larger harm on the public. The comparative magnitude of I_{FP} and I_{FN} is therefore really just a question of how false positives and false negatives will affect prospective inventors' expectations of the chances of experiencing that kind of mistake in the future.

Much about how inventors form their impressions about the accuracy of the patent system remains unknown, but it seems likely that false positives will tend to have a more dramatic effect on prospective inventors' probability estimates than will false negatives.[98] By their nature, false positives tend to be the more observable form of error.[99] For example, if an application is wrongly denied at the examination stage, competitors and the public may never have any reason to notice.[100] But if an application is

---

[97]   To see this, let ε_0 represent the increase in p_0 following a false positive, and let ε_A represent the reduction in p_A following a false negative. If (as stated in the main text) each form of error affects future estimates of the relevant probability of error by an equivalent amount, then $\varepsilon_A = \varepsilon_0$. Equation (2) shows that marginal rewards are given by $M = S(p_A - p_0)$. Therefore, decreasing p_A by ε_A or increasing p_0 by ε_0 will have an identical effect on marginal rewards—a reduction of $S * \varepsilon_0$. Note, however, that this depends on the assumption that the resulting patent protection will have the same value S to the inventor whether or not the patentability requirements were indeed satisfied, that is, $S_A = S_0$. This assumption seems reasonable as a general matter—given two patents of equal enforceability, why would an inventor care if one was procured undeservedly?—but it's possible it breaks down in the obviousness context. See note 117.

[98]   That is, $\varepsilon_0 > \varepsilon_A$.

[99]   See Wagner, 157 U Pa L Rev at 2141 n 16 (cited in note 6).

[100]   In some cases, denied patent applications may be fully hidden from public view. See 37 CFR § 1.14(a) (describing the limited circumstances in which an unpublished, abandoned application may be opened for inspection). However, because the vast majority of applications are now published eighteen months after filing, rejected patent applications are for the most part available for public inspection, if anyone cares to look. See note 28.

2017]

wrongly *granted*, competitors and the public may see it quite a bit.[101] Similarly, if an issued patent is wrongly ruled invalid, even once, collateral estoppel prevents its owner from asserting it any further, and the patent may fall out of public view.[102] But if an issued patent is wrongly *upheld*, it lives on to appear in headlines and in demand letters. False negatives are prone to being forgotten, while false positives are prone to further publicity. And if the effect of a false positive on prospective inventors' probability estimates is merely equal to the effect of a false negative on prospective inventors' probability estimates, then the incentive harms from each form of error should generally be equal too.[103]

Moreover, false positives in the patent system have an additional downside, one that provides an independent reason for thinking that these errors will tend to do more harm to future incentives than false negatives. As others have observed, in some cases, patents awarded undeservedly will reduce the rewards left over for true inventors, who must now share their royalties with overlapping claimants.[104] This harm is conceptually distinct—it

---

101 See Sawicki, 39 Fla St U L Rev at 767–68 (cited in note 6).

101 See Sawicki, 39 Fla St U L Rev at 767–68 (cited in note 6).

102 See *Blonder-Tongue Laboratories, Inc v University of Illinois Foundation*, 402 US 313, 332–34, 350 (1971); *Mendenhall v Barber-Greene Co*, 26 F3d 1573, 1577–78 (Fed Cir 1994).

103 See note 97 and accompanying text. One potential complication here comes not from patent law but from psychology: factors that are more colorful or distinctive tend to wield a disproportionate influence on human behavior. See Shelley E. Taylor, *The Availability Bias in Social Perception and Interaction*, in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds, *Judgment under Uncertainty: Heuristics and Biases* 190, 192 (Cambridge 1982). These salience effects make it possible that false negatives (perhaps vivid, highly recallable injustices) could have a larger effect on inventors' perceptions of rewards than false positives, despite the latter being more easily observable. See Deborah H. Schenk, *Exploiting the Salience Bias in Designing Taxes,* 28 Yale J Reg 253, 261–63 (2011) (distinguishing salience from observability). But it is also possible that the effect cuts the other way: perhaps it is jaw-dropping jury verdicts on the basis of seemingly obvious inventions or threadbare disclosures that tend to capture prospective inventors' imaginations. Suffice to say, then, that it is not obvious which way salience effects will cut. Future work will be necessary to understand how inventors' perceptions are affected by each form of error. See Part V.A.

104 See, for example, Robert Patrick Merges and John Fitzgerald Duffy, *Patent Law and Policy: Cases and Materials* 609 (Lexis 6th ed 2013); Lisa Larrimore Ouellette, *Pierson, Peer Review, and Patent Law*, 69 Vand L Rev 1825, 1827–28 (2016) ("[W]hen an applicant receives a patent on an uncertain research plan, it not only means that the patent is not serving a useful teaching function—it also limits the patent incentive for others to solve the problems necessary to obtain the completed invention."). An extreme form of this dilution can be found in single-purse contests, in which giving the prize to someone who does not deserve it necessarily means denying the prize to the one who

can occur with or without the marginal reward effect described above[105]—and, standing alone, it would be quite difficult to quantify. But in conjunction with the marginal reward effect described above, it lends additional weight to the proposition that the incentive harm of a false positive will tend to be at least as large as the incentive harm of a false negative.

It is important to note that, while the 50 percent minimum probability-of-patentability threshold depends on the mutually exclusive choice condition and the assumption that false positives impair future incentives at least as much as false negatives,[106] it does *not* depend on any assessments of the public costs and benefits of patent protection in general. Estimates of those costs and benefits will affect how high *above* the 50 percent mark the threshold for awarding patent rights should be set, but they do not implicate the conclusion that patents should not be awarded with any probability less than that. When these stated conditions hold, the two-sided incentives model dramatically limits the range of potential patentability thresholds, suggesting that patents should be granted only when it appears more likely than not that the patentability requirements have been satisfied.

Moreover, as a theoretical point, this conclusion can be generalized and transplanted back to the simple prize context from which it was derived: when seeking to increase private incentives to do X to the exclusion of Y through the promise of a fixed-value prize with positive cost, the offeror should award that prize only in cases in which it is more likely than not that the applicant has indeed done X and not Y. Counterintuitively, this minimal probability threshold applies without regard to either

---

does. In those cases, a false positive comes paired with a false negative, giving participants a direct interest in seeing that the prize is not erroneously bestowed on their competitors. See Burstein and Murray, 29 Harv J L & Tech at 431–32 (cited in note 65).

[105] For example, imagine the patent system has the ability to provide $1 million a year of royalty value by granting rights to exclude in a new technology area. An error-free patent system would award control of the entire royalty stream to the true inventor of the technology. But instead, the patent system erroneously grants redundant rights to exclude to both the true inventor and an impostor. Future incentives to invent are reduced in two ways as a result of this mistake: the absolute reward for invention is cut in half as a result of royalty sharing *and* the marginal reward for good behavior is reduced as a result of an increased expectation of unjustified rewards for noninvention.

[106] That is, that $I_{FP} \geq I_{FN}$.

the value that the offeror places on incentivizing this conduct or the cost of the prize itself.[107]

A number of questions remain to be settled, however, in order to apply this rule in the patent context. Most importantly, it is not clear whether (or when) the goals of the patent system can be stated as "increasing incentives to do X to the exclusion of Y." If this condition fails—if a goal of the patent system is as simple as incentivizing X—then false positives do not reduce marginal rewards, and the more relaxed, one-sided incentives model of error costs is fully complete. As such, the optimal balance of errors in the patent system turns on a question that has previously escaped notice: *Are the patentability requirements seeking to influence a mutually exclusive choice?*

III. Mutually Exclusive Choices in the Patent System

As the prior Part explained, the question of how the patent system ought to err in cases of uncertainty depends significantly on whether the patentability requirements are designed to influence a mutually exclusive choice. This Part explores this condition in more detail.

The mutually exclusive choice condition turns out to be a complex one, for the simple reason that the various patentability requirements are designed to influence private actors at so many distinct decision points. Clearly, one goal of the patent system is to encourage invention over noninvention—a decision explored in detail in Part III.A. But patent law also evinces an interest in encouraging disclosure over nondisclosure, in stable ownership over theft, and in licensing existing technologies over wasteful reinvention. The legal rules affecting these "noninventive" choices are the focus of Parts III.B and III.C. Finally, several patentability requirements are not intended to influence private conduct at all and are, on their own terms, rooted entirely in concerns about ex post costs. These cost-only doctrines are the subject of Part IV.

---

[107] As the cost of the prize rises in relationship to her interest in incentivizing X, the offeror will want to demand even more certainty that the applicant has indeed done X. The only assumption necessary to derive this 50 percent minimum probability threshold is that false positives affect future applicants' perceptions of the false-positive rate as much as false negatives affect future applicants' perceptions of the false-negative rate. This would not be true if, for example, the prize offeror has the ability to secretly deny deserving prize claimants without detection.

A. To Invent or Not to Invent?

Perhaps the most basic decision patent law seeks to influence is the choice between inventing and noninventing. Doctrinally, marginal incentives to invent are primarily channeled through the obviousness requirement—the condition that a patent shall not be granted "if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains."[108] Though the modern Patent Act frames this question in terms of obviousness, the Supreme Court has interpreted this language as a codification of the long-standing rule that it was necessary to make an *invention*—not merely tinker with the skill of an "ordinary mechanic"—in order to receive a patent.[109]

Although it is deeply intuitive (and universally accepted) that patents should not be granted for obvious improvements,[110] scholars disagree as to what, exactly, is at stake in this determination. Traditionally, obviousness was regarded as a cost-saving tool—a way to avoid the ex post costs of patents for technologies that would have been made soon enough with or without the promise of patent protection.[111] But in recent years, scholars have suggested that the invention requirement may do more

---

[108] 35 USC § 103.

[109] See *Graham v John Deere Co*, 383 US 1, 11–17 (1966); *Hotchkiss v Greenwood*, 52 US (11 How) 248, 265–67 (1850). Indeed, in international contexts, the obviousness inquiry is framed as a requirement of an "inventive step." See Agreement on Trade-Related Aspects of Intellectual Property Rights Art 27(1) n 5, 33 ILM 1197, 1208 (1994) ("TRIPS"). The requirement that the inventor has actually made that step (as opposed to hazarding a guess at it) is enforced through the enablement and written-description doctrines. See Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L Rev 123, 127–30, 147–48 (2006). These are discussed in more detail in the next Section.

[110] See, for example, Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 High Tech L J 1, 14 (1992); Timothy R. Holbrook, *Patent Anticipation and Obviousness as Possession*, 65 Emory L J 987, 1027–29 (2016); *Bonito Boats, Inc v Thunder Craft Boats, Inc*, 489 US 141, 156 (1989) (relating obviousness to novelty).

[111] See *Graham*, 383 US at 11 ("The inherent problem [of the obviousness doctrine] was to develop some means of weeding out those inventions which would not be disclosed or devised but for the inducement of a patent."); *Roberts v Sears, Roebuck & Co*, 723 F2d 1324, 1345 (7th Cir 1983) (en banc) (Posner concurring in part and dissenting in part) ("The balance tips against protection when the invention is the sort that was likely to be made, and as soon, even if no one could have patented it."). See also Michael Abramowicz and John F. Duffy, *The Inducement Standard of Patentability*, 120 Yale L J 1590, 1627–31 (2011).

than previously appreciated. This view, which Professors Michael Meurer and Katherine Strandburg call the "carrots and sticks" approach to obviousness,[112] understands the doctrine not simply as a tool for avoiding the costs of unnecessary patents, but also as a mechanism for influencing the selection of research projects in the first place. In their telling, the prohibition on patenting obvious inventions works to nudge researchers away from incremental improvements of moderate social value and in the direction of groundbreaking initiatives having greater social value.[113] Awarding patents only to nonobvious inventions increases the probability that nonobvious inventions will indeed be produced.

Although not stated in these terms, the carrots-and-sticks approach to obviousness relies on a certain mutual exclusivity among projects. Indeed, Meurer and Strandburg introduce their theory by asserting that the question facing the researcher is not, "Shall I produce this invention?" but rather "Which research path shall I pursue?"[114] It is this constraint that leads to the risk that awarding noninventive projects may counterproductively reduce the number of inventive projects.[115] If it were otherwise—if inventors had the ability to pursue every project with positive expected value—there would be no risk of a project of moderate social value "squeezing out" a project of higher social value.[116]

Because the carrots-and-sticks view is predicated on a public interest in influencing a private mutually exclusive choice, the same reasoning suggests that obviousness determinations ought to be balanced using the two-sided incentives model developed above. A granted patent imposes ex post costs C whether or not the underlying invention was obvious. But a patent granted on an obvious invention also reduces marginal rewards to create

---

112 See Meurer and Strandburg, 12 Lewis & Clark L Rev at 549–50 (cited in note 5). See also Lunney, 7 Mich Telecom & Tech L Rev at 412–13 (cited in note 8).

113 Meurer and Strandburg, 12 Lewis & Clark L Rev at 561–62 (cited in note 5). See also Lunney, 7 Mich Telecom & Tech L Rev at 412–13 (cited in note 8).

114 Meurer and Strandburg, 12 Lewis & Clark L Rev at 549 (cited in note 5).

115 See id at 561–62. See also Lunney, 7 Mich Telecom & Tech L Rev at 409–11 (cited in note 8) (noting that, in some cases, a system that rewards noninventive projects could yield the same outcomes as having no patent system at all); Amy Kapczynski and Talha Syed, *The Continuum of Excludability and the Limits of Patents*, 122 Yale L J 1900, 1945–46 (2013) (noting that patent protection can further distort incentives to invest in projects of lower social value).

116 See Lunney, 7 Mich Telecom & Tech L Rev at 405–07 (cited in note 8).

nonobvious inventions in the future, resulting in some loss of public benefits I_{FP}. This incentive harm from false positives is missing from the conventional approach to error costs, and leads to the conclusion that a higher probability of nonobviousness should be required to justify patent rights than was previously appreciated.[117]

But not everyone accepts the carrots-and-sticks understanding of obviousness. In fact, a number of commentators continue to analyze the doctrine exclusively through the lens of ex post costs. According to the dominant "inducement" approach to obviousness, the purpose of the requirement is simply to avoid the costs of patent protection for inventions that would have been created soon enough with or without the promise of a patent. The doctrine exists, in this view, only to "weed[ ] out those inventions which would not be disclosed or devised but for the inducement of a patent."[118] As Professors Michael Abramowicz and John Duffy explain, denying patent protection in cases in which non-patent incentives were already sufficient "costs society nothing . . . and saves society from needlessly suffering the well-known negative consequences of patents."[119]

Implicit in this rejection of the carrots-and-sticks view is a *lack* of mutual exclusivity when it comes to the selection of research projects. The standard inducement view examines the

---

[117] Note, however, that in the case of obviousness, it's possible that the certainty required to justify patent rights could be less than 50 percent, even if the two-sided incentives model applies. As discussed above, the 50 percent minimum depends on the condition that the incentive harm from a false positive is greater than or equal to the incentive harm from a false negative, that is, $I_{FP} \geq I_{FN}$. It may be reasonable to assume that condition as a general matter, as false positives tend to be more observable than false negatives. See notes 99–103 and accompanying text. But obviousness presents a complication, because it is possible (though not inevitable) that the private value S_0 of a patent on an obvious development may tend to be smaller than the private value S_A of a patent on a nonobvious invention, even if both patents are equally likely to be enforceable in court. If this is the case, then a change of equal magnitude to the probability p_A of getting a patent deservedly and the probability p_0 of getting a patent undeservedly would not have the same effect on marginal rewards, potentially causing the $I_{FP} \geq I_{FN}$ condition to fail. When this occurs, the incentive harm of false positives still pushes in the direction of greater scrutiny (assuming a patent on an obvious improvement has *some* value), but the 50 percent minimum probability threshold does not necessarily hold.

[118] *Graham*, 383 US at 11.

[119] Abramowicz and Duffy, 120 Yale L J at 1594 (cited in note 111). See also Merges, 7 High Tech L J at 29 (cited in note 110). On this account, the obviousness doctrine may play a particularly important cost-saving function when an unexpected development makes a new technology suddenly possible. See John F. Duffy, *Rethinking the Prospect Theory of Patents*, 71 U Chi L Rev 439, 505 (2004).

expected costs and benefits of potential inventions in isolation, and without considering the opportunity costs inventors might face by selecting one project over the other.[120] As a result, the inducement approach has the potential to deny patent protection to some of the most socially valuable projects (because, in isolation, they are already privately profitable), while granting patent protection to projects of only moderate social value (because, in isolation, they are not). *If* the choice of research projects is mutually exclusive, such a scheme may not only deny subsidies to the projects most in need of encouragement, but actively steer prospective inventors in precisely the wrong direction.[121]

Unfortunately, little is actually known about whether and when inventors must make mutually exclusives choices about which projects to pursue. The proponents of the carrots-and-sticks view of obviousness simply assume the existence of such a constraint.[122] (For their part, those who do *not* subscribe to the carrots-and-sticks view similarly assume the absence of such a constraint.)[123] And there are reasons to wonder why an inventor—or, at least, a firm that employs inventors—would not be able to pursue two positive-value projects simultaneously. Increasing cost of capital provides at least one theoretical answer,[124] as does the long lead time required to train cutting-edge researchers in some fields.[125] But these theoretical possibilities do not tell us

---

[120] See Abramowicz and Duffy, 120 Yale L J at 1663–67 (cited in note 111) (noting that, under the inducement standard, "whether an invention is obvious depends on how costly an experiment would be and the probability that the experiment would be successful").

[121] See Kapczynski and Syed, 122 Yale L J at 1945–46 (cited in note 115).

[122] See Meurer and Strandburg, 12 Lewis & Clark L Rev at 549 (cited in note 5). This mutual exclusivity condition is similarly an explicit precondition in the analysis of Professor Glynn S. Lunney Jr. See Lunney, 7 Mich Telecom & Tech L Rev at 408 (cited in note 8).

[123] See Abramowicz and Duffy, 120 Yale L J at 1624 (cited in note 111); Lemley, 83 Tex L Rev at 1057 (cited in note 22) ("Economic theory offers no justification for awarding creators anything beyond what is necessary to recover their average total costs.").

[124] See Kapczynski and Syed, 122 Yale L J at 1945 (cited in note 115).

[125] Note that, in the case of obviousness, the question of mutual exclusivity must be assessed at a social level: Does one firm's selection of an obvious project over a nonobvious project preclude *others* from pursuing the nonobvious project in its stead? Even if a particular firm has only so many engineers with so many hours in a day, the choice of research project is not necessarily mutually exclusive if a different firm can pursue the competing project with equivalent speed and effectiveness.

whether capacity constraints on project selection are the exception or the rule, or indeed whether they ever occur at all.[126]

These questions will not be settled here. But the discussion above reveals that they have greater consequence than has been previously appreciated. Without a complete picture of the value of accuracy in the patent system, the conflict between the inducement view and the carrots-and-sticks view has only subtle implications for obviousness doctrine. (Neither the inducement nor the carrots-and-sticks view of obviousness leads to a universally stricter substantive standard than the other—the two schools simply teach that different considerations should be taken into account.)[127] But once the effects of false positives are considered, these two views of the doctrine lead to very different conclusions about how that doctrine should be applied in cases of uncertainty.

B. To Disclose or Not to Disclose?

Another decision patent law seeks to influence is an inventor's choice to either publicly disclose her invention or keep it secret. Although many scholars consider the goal of encouraging disclosure to be of secondary importance, the Supreme Court has repeatedly placed it on an equal plane with the goal

---

[126] Additional complications lurk deeper still. Although the obviousness of a project is, legally, a binary inquiry—an invention is either obvious or it is not—the social value of various projects can surely range more broadly. Work that falls just short of the obviousness standard might nonetheless yield important public benefits, just as work that far exceeds that threshold might yield greater benefits still. If the goal of patent law is to push inventors toward increasingly ambitious projects (and not just to some minimally inventive threshold), then the costs of errors will depend on the magnitude of the deviation from the legal standard rather than just the presence or absence of a mistake. Errors close to the line of patentability might be much less significant than errors on either end of the spectrum. This distinction is not reflected in patent doctrine (and so is beyond the scope of this Article, which takes substantive patent law as a given, see notes 32–34 and accompanying text), but it would likely be an important consideration for any initiative to increase the accuracy of obviousness determinations. See Christopher Buccafusco, et al, *Experimental Tests of Intellectual Property Laws' Creativity Thresholds*, 92 Tex L Rev 1921, 1942–43 (2014) (describing the potential for achievement thresholds to impair performance).

[127] The inducement view keys off two variables, the private costs and the private benefits of a project prior to patent incentives. By contrast, the carrots-and-sticks view focuses on positive externalities—that is, the magnitude of any benefits not captured by the inventor in the absence of patent protection. Thus, under either view, obviousness turns out to be a highly fact-specific and nuanced determination.

of encouraging invention,[128] even suggesting that subsidizing disclosure may be the principal justification for having a patent system at all.[129]

This preference for disclosure is the reason patent law requires applicants to provide a "written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art . . . to make and use the same."[130] Together, these obligations are called the "disclosure" requirements,[131] and the Supreme Court has long held that compliance with them is an essential part of the "quid pro quo" for the inventor's rights to exclude.[132]

On their face, the disclosure requirements are intended to influence a mutually exclusive choice: whether or not to release to the public all the details necessary to practice the newly invented

---

[128] For example, in 1944 the Court explained that "[a]s a reward for inventions and to encourage their disclosure, the United States offers a seventeen-year monopoly to an inventor who refrains from keeping his invention a trade secret." *Universal Oil Products Co v Globe Oil & Refining Co*, 322 US 471, 484 (1944). Similar statements can be found as early as 1832. See *Grant v Raymond,* 31 US (6 Pet) 218, 247 (1832).

[129] See *Kewanee Oil Co v Bicron Corp*, 416 US 470, 481 (1974):

> When a patent is granted and the information contained in it is circulated to the general public and those especially skilled in the trade, such additions to the general store of knowledge are of such importance to the public weal that the Federal Government is willing to pay the high price of 17 years of exclusive use.

See also *Sinclair & Carroll Co v Interchemical Corp*, 325 US 327, 330–31 (1945) ("The primary purpose of our patent system is not reward of the individual but the advancement of the arts and sciences. . . . [I]t is not a certificate of merit, but an incentive to disclose."); *Grant*, 31 US (6 Pet) at 247 (describing disclosure as "the advantage for which the privilege [of patenting] is allowed, and [ ] the foundation of the power to issue the patent"). For a scholarly defense of patent law's disclosure goal, see Jeanne C. Fromer, *Patent Disclosure*, 94 Iowa L Rev 539, 547–51 (2009).

[130] 35 USC § 112(a). Additionally, the inventor is required to set forth what she contemplates as the "best mode" of carrying out the invention. 35 USC § 112(a). In recent years, this latter requirement has been rendered essentially unenforceable. After the AIA, failure to disclose the best mode is no longer grounds to invalidate a patent in litigation. See AIA § 15(a), 125 Stat at 328, 35 USC § 282(b)(3)(A). Moreover, the position of the PTO is that best-mode rejections are rarely proper in ex parte proceedings (such as examination). See Patent and Trademark Office, *Manual of Patent Examining Procedure* § 2165.03 (Nov 2015), archived at http://perma.cc/JQ35-DWBX (detailing the requirements for rejection for failing to disclose the best mode).

[131] See Jason Rantanen, *Patent Law's Disclosure Requirement*, 45 Loyola U Chi L J 369, 370 n 2 (2013).

[132] *Kewanee Oil*, 416 US at 484 (emphasis omitted). See also *Universal Oil*, 322 US at 484.

technology. At the start, disclosure is technically optional, in the sense that participation in the patent system is voluntary. An inventor is free, after all, to forego patent protection and instead take her chances by relying on trade secrecy protection.[133] But having weighed the comparative benefits of secrecy and patent protection and opted for the latter, an inventor is not supposed to be able to keep one foot in the other boat.[134] Enforcing the disclosure requirements thus guards against applicants' "selfish desire to obtain patent protection without making a full disclosure,"[135] furthering patent law's "ultimate goal" of moving "new designs and technologies into the public domain."[136]

Because the disclosure requirements are intended to influence this mutually exclusive choice, the two-sided incentives model is the appropriate framework for enforcing the requirements in situations of uncertainty. As before, a granted patent imposes ex post costs C, whether or not the disclosure requirements have actually been satisfied.[137] But a patent granted despite inadequate disclosure also reduces marginal rewards to make complete disclosures in the future, resulting in some loss of public benefits I_{FP}.[138] This incentive harm from false positives is missing from the conventional approach to error costs, and suggests that greater confidence that an application has satisfied the

---

[133] See *Kewanee Oil*, 416 US at 487–88.

[134] See W. Nicholson Price II, *Expired Patents, Trade Secrets, and Stymied Competition*, 92 Notre Dame L Rev 1611, 1614–17 (2017). This, at least, is the theory. In practice, various loopholes sometimes permit inventors to have some of the benefits of both forms of protection. See id at 1617–18.

[135] *Application of Nelson*, 280 F2d 172, 184 (CCPA 1960), overruled on other grounds by *Application of Kirk*, 376 F2d 936 (CCPA 1967).

[136] *Bonito Boats*, 489 US at 151. See also Fromer, 94 Iowa L Rev at 589–93 (cited in note 129) (calling for increased enforcement of the disclosure requirements).

[137] An additional consideration may further strengthen the conclusion that greater scrutiny should be applied to the disclosure requirements: it seems plausible that patents with inadequate disclosure may on average impose higher ex post costs C_{ND} than a patent with complete disclosure C_D. See Ouellette, 69 Vand L Rev at 1827 (cited in note 104) ("When an applicant presents data showing that an invention works but obfuscates key steps of the method, the public loses out on the teaching function that a clearer protocol would have provided."). This effect would be difficult to quantify, but, to the extent that $C_{ND} > C_D$, it would tend to push the minimum probability of patentability necessary to justify patent rights even higher.

[138] Note that the potential complication that a deserved patent may be inherently more valuable than an undeserved patent, see note 117, does not apply here, because (unlike obviousness) there is no reason to think that a patent with inadequate disclosure will be *less valuable* to an inventor than a patent with adequate disclosure. In fact, it seems plausible that the opposite would be true.

disclosure requirements should be required to justify patent rights.[139]

C.   To Build, Buy, or Steal?

Another decision that patent law seeks to influence is an inventor's choice to either rely on an existing solution or set out to invent something new. Generally, of course, patent rights are reserved for things that did not exist before.[140] But in some situations reinvention is preferable or even necessary, a fact patent law quietly concedes through various limits on the universe of prior knowledge that can disqualify an invention for patent protection.[141]

The requirement that an invention be new in order to get a patent is found in 35 USC § 102, which states that no patent can be granted if the claimed invention was, prior to the patent filing, "patented, described in a printed publication, or in public use, on sale, or otherwise available to the public."[142] If a single qualifying prior reference or activity, anywhere in the world, describes or contains the complete invention, the invention is said to be "anticipated," and protection is denied.[143]

---

[139] Some readers might wonder about the observability of false positives when it comes to the disclosure requirements, because, by its nature, nondisclosure seems like a difficult thing for outsiders to detect. But it must be remembered that patent law's disclosure requirements are for the most part objective, not subjective. The question is not whether the applicant disclosed everything she knew about an invention, but whether she disclosed enough to allow those skilled in the art "to make and use the invention without undue experimentation." *In re Wands*, 858 F2d 731, 737 (Fed Cir 1988). This requirement is tested by reading the patent specification, see id at 735, so the public has the same opportunity to observe false positives on disclosure grounds as it does false positives on other grounds. The one exception is the best-mode requirement, which *does* require an inquiry into the inventor's subjective state of mind. See *Chemcast Corp v Arco Industries Corp*, 913 F2d 923, 925–26 (Fed Cir 1990). But, as noted above, best mode has recently been rendered an essentially unenforceable requirement. See note 130.

[140]   See *Alexander Milburn Co v Davis-Bournonville Co*, 270 US 390, 402 (1926) ("The fundamental rule . . . is that the patentee must be the first inventor.").

[141]   See *Gayler v Wilder*, 51 US (10 How) 477, 496–97 (1850) ("In the case thus provided for, the party who invents is not strictly speaking the first and original inventor. . . . Yet his patent is valid if he discovered it by the efforts of his own genius, and believed himself to be the original inventor.").

[142]   35 USC § 102(a)(1). In some cases, 35 USC § 102(b)(1) carves out activity that would otherwise qualify as prior art under § 102(a)(1).

[143]   See *Atlas Powder Co v Ireco, Inc*, 190 F3d 1342, 1346 (Fed Cir 1999). Prior to statutory amendments enacted in 2011 and effective in 2013, use and sale activities were disqualifying only if they took place in the United States. See 35 USC § 102(a)–(b)

The most basic choice these requirements seek to influence is the decision to seek a patent (or not) on an invention that the applicant copied from someone else. Indeed, when the requirement of absolute novelty admits of an exception, it is always on the condition that the inventor applied for a patent with a good-faith belief in the invention's originality. When the duplication isn't so innocent—when an applicant acquired the technology from the true inventor rather than inventing it herself—patent protection is categorically denied.[144] In fact, seeking patent protection on a technology knowingly taken from another is one of the few transgressions in patent law that can lead not only to denial by the patent office but also to criminal prosecution.[145] The mutually exclusive choice here is straightforward and clear: at a minimum, patent law seeks to deter would-be thieves from claiming ownership of things they know were invented by others.

(2006). Additionally, the prior art must be enabling in order to be anticipatory. See *Amgen Inc v Hoechst Marion Roussel, Inc*, 314 F3d 1313, 1354 (Fed Cir 2003).

[144] See *Gayler*, 51 US at 496–97 ("[H]is patent is valid *if* he . . . believed himself to be the original inventor.") (emphasis added). Until recently, a separate statutory provision prohibited granting a patent in cases in which the invention was taken from another. See 35 USC § 102(f) (2006) ("A person shall be entitled to a patent unless . . . he did not himself invent the subject matter sought to be patented."); Patent Act of 1870 § 61, 16 Stat 198, 208. This long-standing provision was removed by the AIA, but the position of the PTO is that stolen inventions are implicitly barred by 35 USC § 101, which states that "[w]hoever *invents* or discovers . . . may obtain a patent therefor" (emphasis added). See Dennis Crouch, *With 102(f) Eliminated, Is Inventorship Now Codified in 35 U.S.C. 101? Maybe, but Not Restrictions on Patenting Obvious Variants of Derived Information* (Patently, Oct 4, 2012), archived at http://perma.cc/ETZ3-5M64 (discussing the AIA's changes to the inventorship requirement).

[145] An inventor seeking patent protection has been required to submit a sworn statement that he or she "believes himself or herself to be the original inventor . . . of a claimed invention." 35 USC § 115(b)(2). An inventor who submits this oath with actual knowledge that the invention was taken from another would be making a knowingly false statement to a federal agency, raising the specter of criminal fines or up to five years in prison. See 35 USC § 115(i); 18 USC § 1001. See also Irving Kayton, John F. Lynch, and Richard H. Stern, *Fraud in Patent Procurement: Genuine and Sham Charges*, 43 Geo Wash L Rev 1, 79–80 (1974). Though these prosecutions are exceedingly rare as a practical matter, see Oskar Liivak, *Overclaiming Is Criminal* *20–22 (Cornell Law School Research Paper No 16-35, 2016), archived at http://perma.cc/JV3R-NZDE, the inclusion of a requirement to make a "true inventor" or "original inventor" statement under penalty of criminal sanctions evinces a clear intent to dissuade inventors from knowingly filing for undeserved rights. In addition, an applicant who files a claim she knows to be anticipated may later face antitrust liability, see *Walker Process Equipment, Inc v Food Machinery & Chemical Corp*, 382 US 172, 174 (1965), and sanctions for inequitable conduct, see *Ohio Willow Wood Co v Alps South, LLC*, 813 F3d 1350, 1357, 1360 (Fed Cir 2016).

But the novelty rules do much more. In fact, typically an applicant's actual knowledge is irrelevant to the novelty inquiry, for the simple reason that she is "presumed to know" *all* of the prior art anyway.[146] This presumptive knowledge will not trigger the patent system's more drastic penalties,[147] but it will work to deny protection to an applicant who believes, wrongfully but faithfully, that she was the first to invent a particular technology.[148]

A rule that holds inventors responsible for prior art they did not know about cannot be explained as a mechanism for discouraging theft, because it applies equally in cases involving no theft at all. Instead, this legal fiction appears to be rooted in concerns about the cost of reinvention. As Professors Robert Merges and John Duffy have explained, in cases in which a reasonable amount of searching would have revealed a known solution, it is preferable for the prospective inventor to search and find that solution rather than waste time and money re-creating it.[149] Denying patent protection based on what an inventor could have reasonably discovered encourages prospective inventors to search existing solutions before setting out to make something new, reducing the risk of wasteful reinvention in the future.[150]

On this understanding, the novelty requirements are designed not only to deter theft but also to influence an additional mutually exclusive choice: whether or not to perform a reasonably diligent prior art search before attempting to create

---

[146] See *Evans v Eaton*, 16 US (3 Wheat) 454, 496 (1818). Judge Learned Hand was characteristically forthright on the implausibility of this presumption, explaining, "[W]e must suppose the inventor to be endowed, as in fact no inventor ever is endowed; we are to impute to him knowledge of all that is not only in his immediate field, but in all fields nearly akin to that field." *International Cellucotton Products Co v Sterilek Co*, 94 F2d 10, 13 (2d Cir 1938).

[147] See *Kimberly-Clark Corp v Johnson & Johnson*, 745 F2d 1437, 1450 (Fed Cir 1984).

[148] More recently, the legal fiction of presumed knowledge has been replaced by a rule that the inventor's knowledge is simply irrelevant to novelty. See id at 1454.

[149] See Merges and Duffy, *Patent Law and Policy* at 401–02 (cited in note 104). See also Donald S. Chisum, 1 *Chisum on Patents: A Treatise on the Law of Patentability, Validity, and Infringement* § 3.01 (Matthew Bender 2016); William M. Landes and Richard A. Posner, *The Economic Structure of Intellectual Property Law* 303 (Harvard 2003).

[150] See Merges and Duffy, *Patent Law and Policy* at 401–02 (cited in note 104). See also Louis Kaplow, *The Value of Accuracy in Adjudication: An Economic Analysis*, 23 J Legal Stud 307, 316–17 (1994) (observing that accuracy ex post can encourage participants to inform themselves ex ante). As in other areas of law, knowledge is presumed to create incentives to acquire actual knowledge. See Richard R. Powell, 14 *Powell on Real Property* § 82.02[1][d][iii] (Lexis 2000) (discussing constructive notice in title registration regimes).

something new. Sometimes, of course, an inventor who fails to search will emerge unscathed—the novelty rules bite only if anticipating prior art turns out to exist.[151] But in cases in which a reasonably diligent search would have turned up anticipating prior art, enforcement of the novelty requirement increases marginal incentives to perform such a search in the future.[152] When the inventor knew or should have known about the prior art in question, the two-sided incentives framework is the appropriate model for balancing errors in the adjudication of the novelty requirement.

* * *

Though the matter is complicated, it appears that at least several of the private decisions that patent law seeks to influence have the potential to be mutually exclusive. The two-sided incentives model is thus appropriate, at least in some cases, in the application of 35 USC § 112's disclosure requirements, 35 USC § 102's novelty requirements, and (arguably) 35 USC § 103's nonobviousness requirement.

To be clear, all of these doctrines may *also* have important roles to play in limiting the ex post costs of the patent system. Any legal basis for denying patent protection has the potential to make the patent system cheaper by avoiding deadweight losses and transaction costs.[153] But the presence of mutually exclusive

---

[151] In fact, the relationship between an inventor's decision to search and the consequences for her patent rights is indirect in two ways. First, patent law does not formally require applicants to search the prior art at all. See *Nordberg, Inc v Telesmith, Inc*, 82 F3d 394, 397 (Fed Cir 1996). The failure to do so is relevant only if some anticipating prior art actually existed. Second, even in cases in which an inventor *did* search, it is possible that her work may nonetheless be anticipated by prior art that was actually impossible for her to find. See Part IV.B.

[152] If an inventor *did* search and in fact found the existing solution, she becomes an inventor with actual knowledge of anticipating prior art, and the novelty requirements are intended to dissuade her from seeking an invention she knows she does not deserve. See notes 144–45 and accompanying text. Admittedly, some readers might dispute Merges and Duffy's "incentives to search" explanation for the prior art rules, and instead understand these rules as being rooted exclusively in concerns about ex post costs. If this alternative position were adopted, only cases of actual knowledge would trigger the more demanding, two-sided incentives model, and all other novelty disputes would call for the traditional one-sided model. See Part IV.B.

[153] In addition to these generic benefits from patent denial, there may also be some ex post cost savings that are doctrine specific. For example, individual denials on novelty grounds may protect those who have relied on the public domain. Likewise, enforcing the nonobviousness requirement can protect against patent clutter, preventing the transaction

choices gives the adjudication of these doctrines additional significance. Whatever degree of scrutiny concerns about ex post costs might already justify, the incentive harm of false positives pushes the optimal balance of errors toward additional skepticism of patent claims.

These doctrines are central to patent law—indeed, they are some of the most commonly litigated grounds of invalidity.[154] But, critically, they are not *all* of patent law. There are other patentability requirements as well, and they do not map so neatly onto mutually exclusive choices confronted by prospective inventors. The next Part introduces several requirements of patent law that seem not to be directed at inventor incentives at all.

IV. COST-ONLY DOCTRINES

The prior Part introduced a number of private decisions that patent law seeks to influence, showing how some of the most prominent patentability requirements appear to be directed at mutually exclusive choices. This Part explores the negative spaces of that analysis: the patentability requirements that, on their own terms, do not claim any goal of influencing future inventor conduct. For these doctrines, the mutually exclusive choice condition is lacking, thus making the traditional, one-sided incentives model the appropriate tool for balancing errors in adjudication.

A. Patentable Subject Matter

One example of a patent doctrine that is explicitly *not* rooted in a goal of influencing inventor conduct is the patentable subject matter requirement. The patentable subject matter requirement is rooted in 35 USC § 101, which states that "[w]hoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and

---

costs of the patent system from compounding through a profusion of patents on small, unimportant improvements. See Jeanne C. Fromer, *The Layers of Obviousness in Patent Law*, 22 Harv J L & Tech 75, 81 (2008). See also Merges and Duffy, *Patent Law and Policy* at 609 (cited in note 104).

[154] See John R. Allison, Mark A. Lemley, and David L. Schwartz, *Understanding the Realities of Modern Patent Litigation*, 92 Tex L Rev 1769, 1789 (2014) (noting that enablement, novelty, and obviousness accounted for 68 percent of summary judgment observations).

useful improvement thereof, may obtain a patent therefor."[155] These categories of prima facie eligibility—process, machine, manufacture, or composition of matter—have long been interpreted broadly.[156] But they have also been subjected to several judicially created exceptions, and it is here that most patentable subject matter controversies take place. Though these limitations are not found in the text of 35 USC § 101, the Supreme Court has denied protection to "'laws of nature, physical phenomena, and abstract ideas.'"[157] In the last decade, these judicially created exceptions have been the subject of four Supreme Court opinions,[158] and they remain an exceedingly active topic of litigation.[159]

Why are laws of nature, physical phenomena, and abstract ideas categorically excluded from patent protection? Because, the Court has explained, "Phenomena of nature, . . . mental processes, and abstract intellectual concepts are . . . the basic tools of scientific and technological work."[160] Exactly because they are so critical for future invention, there is a concern that "monopolization of those tools through the grant of a patent might tend to impede innovation more than it would tend to promote it."[161] In short, allowing such patents would "risk disproportionately tying up the use of the underlying natural laws, inhibiting their use in the making of further discoveries."[162]

At heart, then, the reason for excluding these kinds of inventions is out of concern for the ex post costs of affording them patent protection.[163] Commentators largely agree that patents in

---

[155] 35 USC § 101.

[156] See *Diamond v Chakrabarty*, 447 US 303, 308–09 (1980).

[157] *Bilski v Kappos*, 561 US 593, 601–02 (2010), quoting *Chakrabarty*, 447 US at 309.

[158] See *Alice Corp Pty Ltd v CLS Bank International*, 134 S Ct 2347, 2354 (2014); *Association for Molecular Pathology v Myriad Genetics, Inc*, 133 S Ct 2107, 2116 (2013); *Mayo Collaborative Services v Prometheus Laboratories, Inc*, 566 US 66, 70–71 (2012); *Bilski*, 561 US at 609–12. See also *Laboratory Corp of America Holdings v Metabolite Laboratories, Inc*, 548 US 124, 125–28 (2006) (Breyer dissenting from denial of certiorari) (outlining the rationale behind the "laws of nature" exception).

[159] See Robert R. Sachs, *Two Years after* Alice*: A Survey of the Impact of a "Minor Case" (Part 1)* (Bilski Blog, June 16, 2016), archived at http://perma.cc/V8FC-GXAH.

[160] *Gottschalk v Benson*, 409 US 63, 67 (1972).

[161] *Mayo*, 566 US at 71.

[162] Id at 73.

[163] See id at 88 (comparing costs of a patent grant to inventor's contribution). Similar judicial concerns date back to the nineteenth century. See Christopher Beauchamp, *Patenting Nature: A Problem of History*, 16 Stan Tech L Rev 257, 270–71 (2013).

the "exceptional" categories would be staggeringly expensive.[164] Given these substantial costs, the benefits are simply too small to justify granting patent protection, even when the other requirements of patentability have been met.[165]

It is important to notice what this explanation for the patentable subject matter exclusions *doesn't* say. The Court has never suggested that research into these topics lacks benefit—if anything, the justices have stressed the great social value that these kinds of projects can produce.[166] No commentator, apparently, has suggested that the laws of nature, physical phenomena, or abstract ideas are disfavored research topics.[167] The goal of denying patent protection is not to steer inventors to better uses of their time. The problem is simply that it would be prohibitively expensive to afford patent protection to certain foundational technologies.[168]

---

[164] See, for example, Landes and Posner, *The Economic Structure of Intellectual Property Law* at 305–06 (cited in note 149); Golden, 89 Tex L Rev at 1070–74 (cited in note 4); Allen K. Yu, *Within Subject Matter Eligibility—A Disease and a Cure*, 84 S Cal L Rev 387, 428–29 (2011). See also David S. Olson, *Taking the Utilitarian Basis for Patent Law Seriously: The Case for Restricting Patentable Subject Matter*, 82 Temple L Rev 181, 203 (2009) (questioning whether patent rights are necessary to produce innovation in certain areas).

[165] See Landes and Posner, *The Economic Structure of Intellectual Property Law* at 305–06 (cited in note 149); Golden, 89 Tex L Rev at 1070–74 (cited in note 4).

[166] See *Mayo*, 566 US at 85–87. See also *Laboratory Corp*, 548 US at 126 (Breyer dissenting from denial of certiorari).

[167] See Mark A. Lemley, et al, *Life after* Bilski, 63 Stan L Rev 1315, 1329 (2011); Meurer and Strandburg, 12 Lewis & Clark L Rev at 577 (cited in note 5) (contrasting the patentable subject matter requirement with the obviousness requirement). Some have argued that an additional reason not to grant patents on fundamental ideas is that the patent system would not be very effective at rewarding such early-stage inventions anyway. See, for example, Landes and Posner, *The Economic Structure of Intellectual Property Law* at 307 (cited in note 149) (explaining that rewards offered for technologies that are far from commercial development will be small); *Bilski*, 561 US at 651–52 (Stevens concurring in the judgment). But this does not imply that the projects themselves lack value, only that the patent system will not be an effective tool to reward them. Another view is that the patentable subject matter requirement screens out inventions that other patent doctrines would struggle to distinguish. See Golden, 89 Tex L Rev at 1067–68 (cited in note 4). Like the other explanations, this view of the requirement discerns no goal of influencing inventors' future selection of research projects. To be clear, however, none of this rules out the possibility that someone might (in the future) develop a convincing theory of patentable subject matter that implicates a mutually exclusive choice. Such a choice is simply missing from the accounts prominently presented to date.

[168] Historically, § 101 (and its predecessor provision) *was* used to discourage private research in certain fields, through the exclusion of inventions that were "frivolous or injurious to the well-being, good policy, or sound morals of society." See *Lowell v Lewis*, 15 F Cases 1018, 1019 (CC D Mass 1817). Justice Joseph Story famously imagined a

Because the contemporary patentable subject matter requirement is rooted only in concerns for ex post costs, the consequences for future incentives in these determinations are distinctly one-sided. Erroneous denials reduce incentives to invent, disclose, and so on. But erroneous grants cause no corollary incentive harm, because the patent system claims no goal of steering inventors toward patentable subject matter in the first place. The downside of an erroneous patent grant on patentable subject grounds is only the ex post costs resulting from that particular grant. The traditional, one-sided incentives model thus remains the correct approach for balancing errors in determination of the patentable subject matter exclusions.

The appropriateness of the one-sided incentives model here can be reinforced by returning to the hypothetical cost-free patent system discussed above.[169] If, somehow, patents imposed no ex post costs (that is, $C = 0$), the one-sided incentives model would counsel granting patents to anyone who asks for one, regardless of validity. And, in the case of the patentable subject matter exclusions, this conclusion would be entirely correct. If patents on laws of nature, physical phenomena, and abstract ideas came at no cost to the public at all, the stated rationale for excluding these inventions would be rendered hollow.

B.   Obscure Prior Art

While the patentable subject matter requirement is categorically and exclusively rooted in ex post costs, there are other doctrines that may be explained by ex post costs in some circumstances and ex ante incentives in other circumstances. This Section revisits the novelty requirement as one example of a doctrine with blended purposes that can change depending on the facts of the case.

---

number of colorful examples of inventions that ought to be denied patent protection on these grounds, such as those designed "to poison people, or to promote debauchery, or to facilitate private assassination." Id. The Federal Circuit has disavowed this doctrine, see *Juicy Whip, Inc v Orange Bang, Inc*, 185 F3d 1364, 1366–67 (Fed Cir 1999), leaving the judgment to Congress as to which kinds of research should be excluded from the patent system. See, for example, 42 USC § 2181(a) (excluding nuclear weapons from patentable subject matter); AIA § 33(a), 125 Stat at 340 (excluding claims directed to human organisms from patentability).

[169] See note 91 and accompanying text.

As discussed in Part III.C, some applications of the novelty requirement are clearly intended to influence a mutually exclusive choice. When an inventor has actual knowledge that her claimed invention was previously invented by someone else, patent law seeks to deter her from claiming the work of another. Beyond that, when a prospective inventor could have known that a problem had already been solved, the novelty rules are designed to make it more attractive to search for and find the existing solution rather than wastefully reinvent it. In either case, denying patent protection shapes future inventors' incentives in ways patent law cares about.[170]

But the novelty requirement sweeps more broadly than that, working to deny patents to inventions that could have been found only in "obscure" prior art—references or activities unknown to the inventor and that a reasonably diligent search would *not* have uncovered. In cases like these, the novelty requirement appears rooted in costs alone, with no apparent goal of influencing prospective inventors' decisions.

Obscure prior art can arise in two ways. First, some activities and documents can count as prior art even though they would have been extremely expensive to uncover at the time of the invention. For example, the Federal Circuit has held that a single copy of a doctoral thesis sitting in a foreign library can anticipate an invention, so long as the library maintains a subject matter index.[171] In another case, the Federal Circuit concluded that a Usenet post that was not text searchable at the time was nonetheless prior art because Usenet groups themselves were hierarchically organized by topic—in the rough sense that someone interested in learning how to write web code could avoid reading Usenet posts about Wiccan practice or stories of surviving motorcycle accidents.[172]

---

[170] See Part III.C.

[171] See *In re Hall*, 781 F2d 897, 899–900 (Fed Cir 1986).

[172] See *Suffolk Technologies, LLC v AOL Inc*, 752 F3d 1358, 1365 (Fed Cir 2014). Another famous example involved a drawing that was originally filed as part of a Canadian patent application but canceled during prosecution and thus omitted from the issued patent—and that could only be discovered by traveling, in person, to the Canadian Patent Office in Hull, Quebec. The Federal Circuit reasoned that the existence of the published patent would have provided a "roadmap" to make exactly such a trip to see whether any material had been canceled during prosecution. See *Bruckelmyer v Ground Heaters, Inc*, 445 F3d 1374, 1378–79 (Fed Cir 2006).

It seems unlikely that even a reasonably diligent search would have led the prospective inventor to references like these. Of course, reasonable minds might disagree about what, exactly, a reasonably diligent search will entail. But they cannot dispute that the standard implies *some* outer boundary—"reasonable diligence" is not "infinite diligence." As proponents of the incentives-to-search theory have explained, when prior disclosure of the solution was so obscure that the expected costs of finding it exceed the cost of creating it independently, there is no reason to penalize the inventor for coming up with the same thing on her own.[173] From this perspective, outcomes like the ones described in the prior paragraph are tough to justify: it is simply implausible that actual flesh-and-blood inventors are expected to embark on a world tour of foreign libraries and read every Usenet post ever written in the field before doing any inventing of their own.[174]

The second category of obscure prior art does not require any speculation about the metes and bounds of reasonable diligence. Some activities and documents can count as prior art even though they would have been actually impossible for the inventor to find before inventing. Pending patent applications are a classic example. Until they are published or granted by the PTO, patent applications are considered confidential and not available to the public.[175] But after publication or issuance, an application becomes prior art effective as of the day it was filed, meaning it can defeat the novelty of an invention made while the application was pending.[176] As a result, even an inventor who diligently searches the prior art before inventing cannot

---

[173] See Merges and Duffy, *Patent Law and Policy* at 402 (cited in note 104).

[174] Some of these cases suggest that the Federal Circuit has not fully embraced the incentives-to-search theory. Despite repeated references to a standard of "persons interested and ordinarily skilled in the subject matter or art exercising reasonable diligence," see *Kyocera Wireless Corp v International Trade Commission*, 545 F3d 1340, 1350 (Fed Cir 2008) (citations omitted), the Federal Circuit has not stepped through the kind of cost-benefit analysis usually expected of the "reasonable person" in other contexts. These cases turn on objective indicia of accessibility, not a comparison of the expected costs of search versus reinvention. See Margo A. Bagley, *Patently Unconstitutional: The Geographical Limitation on Prior Art in a Small World*, 87 Minn L Rev 679, 711 (2003) (noting that the test seems to be one of "constructive accessibility") (quotation marks omitted). Nonetheless, the outcomes of these cases can perhaps be defended on the basis of ex post costs. See notes 180–83 and accompanying text.

[175] See 35 USC § 122.

[176] See 35 USC § 102(a)(2); *Baxter International, Inc v COBE Laboratories, Inc*, 88 F3d 1054, 1062 (Fed Cir 1996) (Newman dissenting) (explaining a similar pre-AIA provision codified at 35 USC § 102(e)).

rule out the possibility that a secret pending application will later emerge and preclude patentability.[177]

Withholding patents in cases like these—when extreme cost or actual impossibility would prevent even a reasonably diligent search from uncovering the anticipating prior art—does nothing to increase incentives to perform a reasonably diligent search in the future. Whether or not the inventor performs a reasonably diligent search will make no difference in her result: if she is unlucky enough to invent something that happens to be anticipated by obscure prior art, her patent is invalid.[178] Denials like these do not guide incentives at the juncture of a mutually exclusive choice, because the outcomes of the cases do not predictably turn on the inventor's decision at all.[179]

The rationale for denying patents in light of obscure prior art appears to be rooted exclusively in concerns about ex post costs. As the Supreme Court has explained, "[T]he stringent requirements for patent protection seek to assure that ideas in the public domain remain there for the free use of the public."[180] Removing something from the public domain risks imposing significant ex post costs—not only on the parties who may have adopted that specific technology, but also on the public more broadly, by undermining the reliability of the public domain going forward.[181] The Supreme Court considers this concern so trenchant that it has, in dicta at least, elevated novelty to the level of a constitutional requirement, declaring that "Congress may not authorize the issuance of patents whose effects are to

---

[177] The chance for unsearchable prior art to defeat patentability was further expanded by the recent move to a first-to-file system. See AIA § 3, 125 Stat at 285. Because the patent goes to the first person to get a proper application on file, a prospective inventor is now vulnerable not only to applications that might already be pending, but also to any *future* applications that might be filed before the prospective inventor is able to file her own application. Obviously, no amount of searching can reveal the existence of competing inventions that have not yet been made.

[178] See *Bruckelmyer*, 445 F3d at 1378.

[179] See Kaplow, 23 J Legal Stud at 313–14, 332 (cited in note 150) (observing that greater accuracy in adjudication can be a waste of resources if actors lack the same information at the moment of their decision-making). In some range, the risk of obscure prior art might cause prospective inventors to *over*invest in search, but, as the cost of discovering existing solutions approaches the cost of reinvention, such inventors will be better off simply doing the development work and taking the risk of obscure prior art. For criticism of the novelty requirement in cases of obscure prior art, see Alan Devlin, *Revisiting the Presumption of Patent Validity*, 37 Sw U L Rev 323, 344–45 (2008).

[180] *Aronson v Quick Point Pencil Co*, 440 US 257, 262 (1979).

[181] See Robert P. Merges, *Justifying Intellectual Property* 141–43 (Harvard 2011).

remove existent knowledge from the public domain, or to restrict free access to materials already available."[182]

This concern about ex post costs is not limited to cases of obscure prior art. Indeed, the novelty requirement serves an important role in limiting the costs of the patent in cases of obscure and nonobscure prior art alike.[183] The difference is that, in the case of the latter, the novelty requirement *also* plays a role in shaping inventors' incentives to conduct a reasonably diligent search before inventing. As a result, errors in the application of the novelty requirement in cases involving obscure prior art should be balanced under the one-sided incentives model, while the more demanding, two-sided incentives model is the appropriate framework for cases involving prior art the inventor knew or should have known about.[184]

* * *

As the prior discussion shows, there is no simple answer to the question *do the patentability requirements seek to influence a mutually exclusive choice?* For the disclosure requirements, the answer appears to be a categorical "yes"—disclosure is a choice patent law cares about, and the decision is intended to be a mutually exclusive one for applicants. By contrast, the patentable subject matter exclusions appear *not* to be targeted at a mutually exclusive choice. Patent protection is withheld for inventions

---

[182] *Graham v John Deere Co of Kansas City*, 383 US 1, 6 (1966). Note, however, that on a number of occasions Congress has granted patent protection after inventions entered the public domain, though its power to do this is questionable after *Graham*. See *Eldred v Ashcroft*, 537 US 186, 233–35 (2003) (Stevens dissenting) (describing private bills to extend patent terms passed between 1790 and 1875).

[183] See *Kimberly-Clark Corp v Johnson & Johnson*, 745 F2d 1437, 1453 (Fed Cir 1984) ("[T]he real reason for the denial of patent rights [in cases lacking novelty] is the basic principle . . . that no patent should be granted which withdraws from the public domain technology already available to the public."). See also Craig Allen Nard, *Legal Fictions and the Role of Information in Patent Law*, 69 Vand L Rev 1517, 1528–29 (2016) (explaining the legal fiction of presumptive knowledge in terms of protecting the public domain and administrability).

[184] To the extent obviousness *ever* implicates a mutually exclusive choice, see Part III.A, it is almost certainly vulnerable to this case-specific complication as well. In some cases, the prior art rendering a solution obvious did not yet exist at the time of invention or would have been extremely expensive for the inventor to uncover. Outcomes like these are difficult to justify on the basis of ex ante incentives, suggesting that at least in some cases the obviousness requirement appears principally concerned with mitigating ex post costs.

in these categories out of concerns of cost, not because these projects are any less valuable.

Other doctrines are harder to categorize. For the novelty requirement, the answer seems to depend on the nature of the prior art at issue. When novelty is precluded by prior art the inventor knew about or could have reasonably found, the doctrine seeks to influence a number of mutually exclusive decisions involving fraudulent patenting, reinvention, and search. But when obscure prior art is at issue, the intent to influence a mutually exclusive choice is absent, suggesting these denials are rooted only in costs. And, when it comes to perhaps the most important choice patent law seeks to influence—the decision to invent—the presence or absence of a mutually exclusive choice is simply unclear.

This analysis is not intended to be conclusive—reasonable minds could disagree about the purposes of these doctrines or suggest *other* mutually exclusive choices that have been overlooked here. But this discussion does illustrate the complexity of the mutually exclusive choice condition. Evaluating that condition requires discerning the purpose of individual patent rules at a granular level and identifying situations in which that purpose may or may not be frustrated as a result of spurious rewards. As the novelty rules illustrate, a doctrine can be rooted in concerns about ex post costs in some cases and rooted in concerns about ex ante incentives in other cases. As the obviousness requirement illustrates, for some doctrines our existing understandings of the patent system and the nature of invention may be inadequate to answer the question.

This distinction might seem a fine one, but it bears great consequence when it comes to the cost of erring in cases of uncertainty. When a doctrine is rooted in concerns about ex post costs, an undeserved patent imposes those ex post costs unnecessarily. But when a doctrine is rooted in concerns about ex ante incentives, an undeserved patent imposes ex post costs *and* reduces ex ante incentives. This is the "double harm" of false positives under the two-sided incentives model. The difference has far-reaching consequences, both for the optimal balance of errors at a systemic level, and for the value of rooting out erroneous patent grants after the fact.

V. IMPLICATIONS AND FUTURE WORK

A. Erring against Patentability

As discussed in Part I, the patent system presently errs in favor of patent rights at nearly every turn. By now, the sources of this bias are well known. Patent applications are reviewed with a presumption of patentability; patent examiners are short on time and lack access to certain categories of prior art; the PTO itself has an institutional interest in favoring allowance; granted patents enjoy a presumption of validity; many such patents are never challenged or litigated to judgment.[185] These procedural rules and structural features cause the patent system to err in favor of affording protection in objectively close cases.

A number of scholars have proposed reforms to rebalance the errors of the patent system to be less friendly toward patent rights in cases of uncertainty. For example, Professor Sean Seymore has suggested that, at the application stage, those seeking patents should be charged with the burden of persuasion when it comes to their entitlement to the rights they seek.[186] Under Seymore's proposal, the PTO would still be charged with the burden of production—that is, bringing forth evidence as to why a claim is unpatentable. Once such evidence is presented, the applicant would then be required to carry the burden of persuasion.[187] The goal of this reform would be to subject applications to additional scrutiny and reduce the number of "low-quality" (that is, false positive) patents issued by the PTO.[188]

Others have suggested changes to make it easier to correct false positives at the enforcement stage. A particularly common target is the presumption of validity, which requires a challenger to show a patent's invalidity by "clear and convincing" evidence. For example, Professors Doug Lichtman and Mark Lemley have

---

[185] See Part I.B.

[186] See Seymore, 97 Minn L Rev at 1023–31 (cited in note 43). See also Lemley, 95 Nw U L Rev at 1524 (cited in note 6).

[187] Seymore, 97 Minn L Rev at 1023–31 (cited in note 43).

[188] Id at 1040–41. Along slightly different lines, Professor Merges has suggested that examiners should be given greater credit for rejecting applications, to correct the pro-issuance bias that exists at the level of individual examiners. See Merges, 14 Berkeley Tech L J at 607–09 (cited in note 1). And Professor Jonathan Masur has suggested that third-party appeal mechanisms could correct the pro-issuance bias that exists at the level of the agency itself. See Masur, 121 Yale L J at 522 (cited in note 1).

suggested that standard be lowered, except in cases in which an applicant previously opted into a more rigorous, "gold-plated" review process in the PTO.[189] Judge William Alsup has proposed a statutory reform in which the strength of the presumption would vary depending on whether the examiner had the opportunity to consider the argument for invalidity in the first instance.[190] All of these proposals are rooted in the claim that the existing presumption provides too much shelter to weak patents—that is, that the system errs too heavily in favor of tolerating false positives to avoid false negatives.[191]

The two-sided incentives model developed above has the potential to lend prescriptive weight and increased specificity to these reforms. At a minimum, recognizing an additional, previously unappreciated harm from erroneous patent grants suggests that the probability of patentability necessary to justify a patent should be higher, all else equal.[192] But, more specifically, under certain conditions the incentive effects of false positives can narrow the range of plausible probability thresholds, suggesting that the minimum probability of patentability to justify patent rights starts at 50 percent and goes up from there.[193] In this way, the incentive harm caused by false positives can provide a theoretical path to showing that the current bias in favor of patent rights is inappropriate, and without relying on any assumptions about the costs and benefits of patent protection in general.

There are, however, several questions that will require future investigation to make these conclusions robust. First, as noted above, the 50 percent minimum probability threshold depends on the assumption that the incentive harm from a false positive is at least as great as the incentive harm from a false negative. There are reasons to think that will often be the case,[194] but, ultimately, the question is an empirical one. If, for example, managers of technology firms are acutely aware of

---

[189] See Lichtman and Lemley, 60 Stan L Rev at 59–62 (cited in note 1). See also Devlin, 37 Sw U L Rev at 352–58 (cited in note 179) (making a similar proposal).

[190] Alsup, 24 Berkeley Tech L J at 1649 (cited in note 33).

[191] See, for example, Lichtman and Lemley, 60 Stan L Rev at 47 (cited in note 1); Alsup, 24 Berkeley Tech L J at 1650 (cited in note 33); Masur, 121 Yale L J at 473 (cited in note 1); Merges, 14 Berkeley Tech L J at 589 (cited in note 1).

[192] See note 92 and accompanying text.

[193] See note 96 and accompanying text.

[194] See notes 98–105 and accompanying text.

improper denials but oblivious to the phenomenon of undeserved grants, that would seriously undermine the claim that false positives affect incentives in this way. Future work will be necessary to better understand how errors affect prospective inventors' perceptions of the marginal rewards offered by the patent system.

A second matter requiring further study is the mutual exclusivity condition. While there are some patentability determinations that appear to be squarely directed at mutually exclusive decisions (such as those involving the disclosure requirements) and some that just as clearly appear not to be (such as those involving the patentable subject matter exclusions), there are others for which the answer is more complicated. As shown in the analysis above, assessing this condition often forces difficult questions about the purposes of the doctrines themselves and the constraints faced by prospective inventors. As the significance of this condition was not previously appreciated, it seems likely that additional analysis (and debate) will be necessary before there is a firm consensus as to which patent doctrines are intended to influence decisions at all and under what conditions those decisions are likely to be mutually exclusive.

Third, there is an additional complication not yet mentioned. It is possible that in some cases requiring multiple patentability determinations, false negatives could be conjunctive, in the sense that an inventor's incentives to do one thing might be affected by a risk that the inventor will be accused of failing to do other things. For example, it is clear than an inventor's incentives to comply with the disclosure requirements will depend on the probability that she will get a patent even if she does not make a complete disclosure (a single potential false positive). But in the other direction, her incentives will also depend on the probability that, even if she does make a complete disclosure, she will nonetheless be denied a patent for failing to satisfy the disclosure requirements (one potential false negative) *or* for failing to satisfy the novelty requirements (another potential false negative) *or* for failing to satisfy the nonobviousness requirements (still another potential false negative), and so on. An error in the application of the patentable subject matter requirements (which, again, are *not* intended to influence a mutually exclusive choice) could inadvertently affect marginal

rewards to fulfill the disclosure requirements (which, again, *are* intended to influence a mutually exclusive choice).

This complication could be avoided if there were a single, agreed goal of the patent system, or if it were understood that one goal so dominated the others that marginal rewards need be attended along only a single frontier. But so long as the patent system seeks to influence multiple decisions simultaneously, one must confront the fact that balancing the errors in the enforcement of any one rule will affect marginal rewards to comply with other rules. In the end, balancing the errors in the patent system may require a policymaker to confront difficult questions about the comparative importance of the various goals of that system.[195]

B.   The Public Benefits of Patent Challenges

Aside from the question of how the patent system should err in cases of uncertainty, these observations also have implications for when and how much scrutiny should be applied to claims of patent rights. This Section briefly explores how the incentives-reducing harm of false positives can affect the value of testing a patent's validity after it has been granted.

Although the discussion above is framed around an abstract "patent system," in practice, decisions about patentability are made at a series of steps across a patent's life. The PTO first examines an application to determine if a patent should be granted at all. This initial examination step is hurried, ex parte, and comparatively cursory.[196] *If* a patent is granted, it may then be subjected to adversarial proceedings of a significantly more searching character, including litigation in federal district court and administrative challenges before the Patent Trial and Appeal Board.[197] These postgrant challenge mechanisms have the benefits of better information and longer deliberation, but

---

[195] See Yelderman, 96 BU L Rev at 1598–1614 (cited in note 13). Note, however, that for some patent doctrines, the potential for conflict may be smaller than it first appears. For example, if the disclosure requirements play a role both in creating incentives to disclose and in creating incentives to invent, see note 109, then strictly enforcing the disclosure requirements may advance both goals simultaneously.

[196] See Frakes and Wasserman, 99 Rev Econ & Stat at *8–9 (cited in note 44); Dreyfuss, 12 Lewis & Clark L Rev at 434 (cited in note 5); Seymore, 49 UC Davis L Rev at 995–96 (cited in note 33).

[197] See 35 USC §§ 282, 311–19, 321–29.

they are expensive (litigating a case through trial typically costs $1 million or more)[198] and are in various ways restrained from correcting certain categories of mistakes that might have been made by the PTO.[199]

Despite the costs and limitations of patent challenges, courts and commentators have long endorsed policies to encourage more of them. For example, to gin up more patent litigation, courts have voided otherwise-enforceable contract terms,[200] overturned long-standing equitable doctrines,[201] and suggested that the act of settling a patent case can itself trigger antitrust liability.[202] For their part, commentators have proposed a number of mechanisms to increase the rate at which issued patents are challenged in postgrant proceedings—things like bounties, expanded standing rules, and one-way fee shifting.[203]

This Article's revised account of error costs in the patent system can potentially provide new support for these pro-challenge policies.[204] In addition to the traditionally cited benefit

---

[198] See *2015 Report of the Economic Survey* I-110 to -112 (American Intellectual Property Law Association 2015). The administrative challenge proceedings are significantly less expensive—the average reported attorneys' fees through the hearing stage are around $330,000. Id at I-139.

[199] District-court litigation cannot reverse grants that are objectively close to the line of patentability because of the presumption of validity, which requires "clear and convincing evidence" of a patent's invalidity. See *Microsoft Corp v i4i Limited Partnership*, 564 US 91, 95 (2011). That heightened presumption does not apply in the PTO's inter partes review process, but those proceedings are limited to considering arguments for anticipation and obviousness in light of printed prior art. See 35 USC § 311(b); *Cuozzo Speed Technologies, LLC v Lee*, 136 S Ct 2131, 2144 (2016); Dolin, 56 BC L Rev at 919–20 (cited in note 50). A third challenge mechanism, postgrant review, permits the challenger to make any invalidity arguments based on any kind of prior art, but these challenges are only available very early in a patent's life. See 35 USC § 321(b)–(c); Dolin, 56 BC L Rev at 914–19 (cited in note 50).

[200] See *Lear, Inc v Adkins*, 395 US 653, 674–75 (1969); *Warner-Jenkinson Co v Allied Chemical Corp*, 567 F2d 184, 188 (2d Cir 1977); *Atlas Chemical Industries, Inc v Moraine Products*, 509 F2d 1, 6 (6th Cir 1974); *Panther Pumps & Equipment Co v Hydrocraft, Inc*, 468 F2d 225, 230–32 (7th Cir 1972); *Massillon-Cleveland-Akron Sign Co v Golden State Advertising Co*, 444 F2d 425, 427 (9th Cir 1971).

[201] See *Lear*, 395 US at 674–75.

[202] See *Federal Trade Commission v Actavis, Inc*, 133 S Ct 2223, 2234–37 (2013).

[203] See, for example, Thomas, 2001 U Ill L Rev at 340–42 (cited in note 1); Farrell and Merges, 19 Berkeley Tech L J at 968–69 (cited in note 1); Lemley and Shapiro, 19 J Econ Persp at 88–89 (cited in note 5); Michael Risch, *Patent Challenges and Royalty Inflation*, 85 Ind L J 1003, 1022 (2010); Malani and Masur, 101 Georgetown L J at 672–73 (cited in note 20); Roger Allan Ford, *Patent Invalidity versus Noninfringement*, 99 Cornell L Rev 71, 123–27 (2013); Burstein, 83 Geo Wash L Rev at 542–48 (cited in note 1).

[204] Traditionally, the most common explanation for why patent cases justify a departure from general pro-settlement principles has been that patent litigation has the

of saving ex post costs, patent challenges could yield public bene-
fits by increasing the accuracy of the patent system. Specifically,
revoking an undeserved patent through a postgrant process can
reduce future prospective inventors' expectations that they will
someday receive a patent undeservedly. This reduction in the
perceived false-positive rate can in turn increase marginal
rewards for inventing and disclosing in the future. The resulting
public benefits can exist independently of any benefit obtained
by reducing ex post costs—an important feature, given that ex
post costs savings are not necessarily available in every patent
challenge.[205]

But specifying this potential benefit also reveals its limits.
Just as not every case is a viable candidate to reduce ex post
costs, not every erroneous patent grant will have a deleterious
effect on ex ante incentives. The reason a grant was erroneous
matters significantly. In cases involving a patent doctrine in-
tended to influence a mutually exclusive choice (such as the dis-
closure requirement), the possibility of obtaining patent protec-
tion undeservedly reduces the marginal reward available for
doing the things that doctrine seeks to encourage. Revoking
patents that were erroneously granted on those grounds might
well increase inventors' incentives to comply in the future.[206] But
in cases *without* a mutually exclusive choice at stake, this false-
positive effect is missing. Challenges rooted in cost-only doc-
trines (such as patentable subject matter or anticipation by
obscure prior art) can benefit the public only if they indeed save
ex post costs.

In sum, the public interest in an accurate patent system
might be able to justify policies to encourage certain patent chal-
lenges, but it can only go so far. Contrary to long-standing as-
sumption, the benefits of patent challenges are neither univer-
sally available nor homogenous. A discrete set of challenges may
have the potential to save ex post costs, while another group
may hold the potential to improve ex ante incentives. And some

---

ability to increase competition, such that it can free the public from the demands of a
"patent monopolist." See Stephen Yelderman, *Do Patent Challenges Increase Competi-
tion?*, 83 U Chi L Rev 1943, 1951–52 (2016). A few commentators have alluded to the
benefits of increased accuracy in general terms. See note 13.

[205] See Yelderman, 83 U Chi L Rev at 1994–95 (cited in note 204).

[206] Note that even in cases in which the public benefits of increased accuracy appear
to be available, their magnitude may be quite vexing to estimate. See id at 1956–57.

challenges may be incapable of providing either benefit at all. Future work will be necessary to better understand how these two potential categories of benefits interact, and whether (and when) these benefits exceed the costs of the legal process necessary to bring them about.

CONCLUSION

It is very much the conventional wisdom that the patent system makes too many mistakes in a pro-patent direction, and that the public has an interest in seeing those mistakes corrected through postgrant litigation. To a great extent, this Article specifies and confirms a set of intuitions that have long been held by courts and commentators. There *is* a public interest in the allocation of the patent system's benefits, and the value of accuracy here is even greater than has been previously appreciated.

However, this Article has also unearthed a number of previously unobserved complications. First and foremost, the importance of accuracy depends significantly on the nature of the potential error in question. An erroneous grant harms ex ante incentives *only* if it involves a patentability doctrine designed to influence a mutually exclusive decision. Future work will be necessary to fully explore whether (and when) the patentability requirements are intended to influence future conduct at all, as well as whether (and when) that conduct is likely to involve mutually exclusive choices.