2013

Raising the Stakes in Patent Cases

Anup Malani

Jonathan Masur

Raising the Stakes in Patent Cases

ANUP MALANI & JONATHAN S. MASUR[*]

INTRODUCTION

One of the fundamental goals of a patent system is to encourage the research and development of the most socially valuable inventions—those innovations that will produce the greatest benefits for society at large. If the government could determine in advance which inventions are most socially valuable, it could simply offer direct rewards for their development.[1] The fact that the United States has chosen to employ patents rather than direct rewards to encourage innovation reflects a decision to decentralize the task of picking winners. This policy choice is premised on the notion that, if inventors or the market are in a better position than the government to identify valuable innovations, the government should delegate the task by granting inventors a patent as a reward for innovation.[2] Patents entitle

[*] Lee and Brena Freeman Professor of Law and Deputy Dean, Professor of Law and Herbert and Marjorie Fried Teaching Scholar, University of Chicago Law School, respectively. We thank Omri Ben-Shahar, Colleen Chien, John Golden, Tim Greene, Todd Henderson, Mark Lemley, Saul Levmore, Rob Merges, Ed Morrison, David Schwartz, Nick Stephanopolous, and participants in a faculty workshop at the University of Chicago Law School and the 2012 Intellectual Property Scholars Conference for helpful comments and conversations, and Carl Newman and Dayron Silverio for excellent research assistance. This work was supported by the David and Celia Hilliard Fund, the Microsoft Fund and the Samuel J. Kersten Faculty Fund.

[1] *See* Steven Shavell & Tanguy van Ypersele, *Rewards Versus Intellectual Property Rights*, 44 J.L. & ECON. 525, 534–36 (2001); *see also* Michael Kremer, *Patent Buyouts: A Mechanism for Encouraging Innovation*, 113 Q.J. ECON. 1137, 1137–38 (1998) ("Under symmetric information and full commitment, the first-best solution to underprovision of ideas is subsidizing research, rather than creating a new set of monopoly price distortions through the patent system. However, before research is conducted, the government may not know the costs and expected benefits of research, and may not even be able to conceive of some inventions." (citations omitted)).

[2] Of course there are other reasons one might still prefer a reward system to a patent system. For example, a reward system that released the innovation into the public domain would impose less deadweight loss to welfare. The information advantage of inventors is a necessary but not sufficient condition to preferring patents over rewards.

inventors to monopoly profits from an innovation,[3] and monopoly profits tend to increase as the social value of an innovation increases.[4] Thus, the patent system generally encourages inventors to work on what they believe will be the most valuable inventions.

This rationale for choosing a patent system over a reward system explains why the government may choose to grant patents in the first place. It does not explain, however, why the government sometimes takes patents away after they have been granted. These patent revocations are commonly triggered when the defendant in a patent infringement case successfully challenges the validity of the patent held by the plaintiff.[5] The logic behind patent challenges and revocations is that the government, when implementing a patent system, might accidentally give out patents to entities that did not innovate or did not need a reward in order to innovate.[6] Such invalid patents have no upside: they do not encourage innovation, and they impose deadweight losses on welfare. In short, patent challenges weed out invalid patents.[7]

---

[3] Shavell & van Ypersele, *supra* note 1, at 529.

[4] This claims rests on the absence of any a priori reason why the fraction of social surplus extracted by a monopolist rises or falls with the size of that surplus. *See infra* section I.A.

[5] Patents can also be revoked via administrative proceedings before the Patent and Trademark Office. *See* 35 U.S.C. §§ 301 –329 (2006 & Supp. V 2011).

[6] The likely reason for these mistakes is that the Patent and Trademark Office must make decisions about which applicants deserve patents with very little information about their innovation beyond that which the applicant itself provides. Peer review does not come until later, when profit sharing becomes a motive for an infringer to provide the government with more balanced information about the validity of a patent. *See infra* section I.B.

[7] Scholars have criticized patent law for making mistakes in weeding out socially worthless patents. The gist of the argument is that the criteria patent law employs to judge the validity of, say, utility patents—novelty, non-obviousness, and utility—do not perfectly identify those innovations that improve social welfare, as an economist might define it. Accordingly, the argument goes, these criteria do not successfully induce valuable innovation while deterring socially worthless research**.** *See, e.g.*, FED. TRADE COMM'N, TO PROMOTE INNOVATION: THE PROPER BALANCE OF COMPETITION AND PATENT LAW AND POLICY 6–8 (2003), *available at* http:// www.ftc.gov/os/2003/10/innovationrpt.pdf; ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS: HOW OUR BROKEN PATENT SYSTEM IS ENDANGERING INNOVATION AND PROGRESS, AND WHAT TO DO ABOUT IT 24–35 (2004); Jay Dratler, Jr., *Alice in Wonderland Meets the U.S. Patent System*, 38 AKRON L. REV. 299 (2005). We share these concerns, but we have little to add to them and they do not affect the arguments we make later in the text about the problem with patent challenges and potential reforms to improve those challenges. Therefore, we proceed under the

There are two problems, however, with patent challenges. In some cases they impose costs on valid patents, and in other cases they fail against invalid patents. These flaws stem from two sources. First, infringers occasionally bring challenges even when patents are valid, causing the holders of valid patents to bear litigation costs in responding to a patent challenge. At the same time, alleged infringers may fail to challenge all holders of invalid patents, allowing these patents to continue imposing deadweight loss. Second, occasionally courts may make an error when judging whether a patent is valid or invalid. This may cause the holder of what is truly a valid patent to lose that patent or allow an invalid patent to stand. These failures decrease the *ex ante* returns to any innovation that deserves a valid patent, undermining the incentives at the core of the patent system, and increase the social costs of the patent system.

Moreover, these failures are most acute in cases involving the most socially valuable patents and the largest firms. Whereas the patent system seeks to decentralize the choice of innovation, patent litigation also decentralizes the decision to challenge a patent. Specifically, it delegates the decision to private parties, ideally potential entrants into the patent holder's market. However, there may be few firms in a position to challenge a patent and large fixed litigation costs to filing a challenge. Thus challengers tend to target holders of the most profitable and (and often most socially valuable) patents.[8] Smaller patent holders are particularly vulnerable because they cannot afford substantial litigation costs. By implication, challengers tend to avoid taking on larger firms because there is a lower likelihood of succeeding against even an invalid patent held by such firms.[9] This discourages innovation at smaller firms and tolerates socially harmful patents held by larger firms.

At bottom, the problems with patent challenges are primarily attributable to judicial and administrative errors. If the Patent and Trademark Office (PTO) granted only valid patents, or if the courts could be trusted to uphold all valid patents and strike down all invalid ones, our

---

assumption that the law determining which patents are valid operates as a reasonable proxy for which patents (and the inventions they protect) increase social welfare. We shall focus instead on errors in application of that law by courts.

[8] C. Scott Hemphill & Bhaven N. Sampat, *When Do Generics Challenge Drug Patents?*, 8 J. EMPIRICAL LEG. STUD. 613, 632 (2011).

[9] Michaxel J. Meurer, *Inventors, Entrepreneurs, and Intellectual Property Law*, 45 HOUS. L. REV. 1201, 1234 (2008) (detailing numerous advantages of large firms over small firms in the use of intellectual property for profit and in litigation).

system of patent challenges would function almost perfectly. Yet errors are endemic throughout all levels of the process. The PTO issues scores of invalid patents every year,[10] and the federal courts are notoriously inaccurate when adjudicating patent validity.[11]

If judicial and administrative inaccuracy is the disease, then improving that accuracy would seem the most obvious cure. Indeed, proposals to reduce the error rate within the federal courts and the PTO are legion and involve everything from increased funding and technical training to full-scale restructuring of the judicial process.[12] Nonetheless, the patent system's endemic errors and inaccuracies have proven notoriously resilient in the face of ongoing ameliorative efforts.[13] It may be that there are upper limits to the level of precision that generalist judges can bring to a system involving such technically complex subject matters, for instance.[14]

In this Article, we suggest that it might be possible to improve the value of patent challenges even without increasing their accuracy. Put simply, we

---

[10] *E.g.*, JAFFE & LERNER, *supra* note 7, at 32–33 (describing a patent granted based on the "determination that surrounding jam with peanut butter so the bread will not get soggy is a new idea, and one that was not previously obvious to skilled sandwich-makers"); Mark A. Lemley, Essay, *Rational Ignorance at the Patent Office*, 95 NW. L. REV. 1495, 1495 & n.1 (2001) ("Complaints are legion."); Mark Lemley, Doug Lichtman & Bhaven Sampat, *What To Do About Bad Patents?*, REGULATION, Winter 2005–2006, at 10, 10 (2005) ("Bad patents are everywhere."); Doug Lichtman & Mark A. Lemley, *Rethinking Patent Law's Presumption Of Validity*, 60 STAN. L. REV. 45, 47 (2007) (arguing that PTO grants "patents that should never have been issued" because of presumption of validity in judicial review); 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 (1999) (noting that "patents for 'business methods' implemented in software . . . are of extremely poor quality"); Michael J. Meurer, *Patent Examination Priorities*, 51 WM. & MARY L. REV. 675, 676 (2009) (describing how "PTO struggles to improve examination quality").

[11] *See* David L. Schwartz, *Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 MICH. L. REV. 223, 258–59 (2008) (demonstrating that even sophisticated and experienced federal courts struggle with patent cases).

[12] We describe and critique these various proposals in section III.A, *infra*.

[13] *See, e.g.*, NAT'L RES. COUNCIL, A PATENT SYSTEM FOR THE 21ST CENTURY (Stephen A. Merrill et al. eds., 2004) (describing a continuing crisis within the patent system); Peter S. Menell, *The Property Rights Movement's Embrace of Intellectual Property: True Love or Doomed Relationship?*, 34 ECOLOGY L.Q. 713, 737 (2007) (describing "a growing patent crisis"); *see also* Schwartz, *supra* note 11, at 258–59 (demonstrating that judges to not appear to improve as they gain experience with patent cases).

[14] *See infra* section III.A.

propose raising the stakes involved in patent litigation. A patent owner who prevails at trial should collect enhanced rewards, above and beyond the damages the owner would normally be paid in compensation for the infringement. A patent owner whose patent is invalidated at trial should be forced to pay significantly enhanced penalties. At first glance, our proposal might seem entirely counterintuitive. If patent adjudications are riddled with errors, one would think that it would be preferable to *lower* the stakes involved, rather than increase them. Scholars and courts have largely confined themselves to that approach.[15]

Yet contrary to the conventional wisdom, we demonstrate that enhanced rewards and penalties can correct many of the flaws inherent to patent challenges even without affecting the accuracy of the adjudications themselves. They accomplish this by restoring patent holders' *net expected trial outcomes* to appropriate levels.[16] Enhanced rewards would compensate holders of valid, valuable patents for the risks they run at trial. This would incentivize the optimal amount of research and innovation, as well as continued research on the most socially valuable inventions. At the same time, enhanced penalties would reduce or eliminate invalid patent owners' opportunities to earn positive returns at trial, vastly diminishing their incentives to assert their invalid patents in the first place.[17]

[15] *See infra* notes 82–84 and accompanying text.

[16] In a way, our proposal is related to Gary Becker's observation that deterrence is a function of the probability of apprehension times the fine paid upon apprehension. *See* Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 172, 180 (1968). Focusing on improving the accuracy of courts is akin to changing the probability of apprehension, although our proposal to raise stakes is akin to changing the criminal fine. Our proposal is also related, though less directly, to Keith Hylton's observation that the enforcer's dilemma that plaintiffs face can be overcome by type II errors by courts. In other words, plaintiffs can be encouraged to litigate by courts who accidentally rule for them. *See* Keith L. Hylton, *Costly Litigation and Legal Error Under Negligence*, 6 J. L., ECON. & ORG. 433, 434 (1990). We make an opposite point: court (and PTO) errors in favor of patent holders can be overcome by increasing the penalties that invalid patent holders pay when their patent is invalidated.

[17] Our proposal resembles, but is distinct from, the English Rule, which requires that the losing party in a civil litigation compensate the prevailing party for its direct litigation costs. First, our remedy allows an increase in damages that substantially exceed the cost of litigation. Second, we advocate disconnecting the amount that the prevailing party received from the amount that the losing party pays. For example, we advocate giving prevailing patent holders patent extensions, though those patent extensions would not be paid for by infringers. *See infra* section III.D.

The enhanced rewards and penalties we propose would thus allow our imperfect patent system to mimic one in which courts erred less frequently. Patent owners—be they genuine innovators or patent trolls—and their competitors would behave as if they could rely upon the courts to reach the correct outcome in essentially every case. The system would generate substantial benefits to innovation and competition at minimal cost. Where direct efforts to improve judicial accuracy have failed, raising the stakes of patent cases might yet succeed.

The Article proceeds as follows. Part I explains the manner in which patents direct research and innovation toward the most socially valuable inventions and describes the value of patent challenges. Part II analyzes the problems created by patent challenges when courts err in assessing the validity of patents. Part III presents our proposal for enhanced rewards and penalties and offers a theoretical demonstration of its ability to realign research and litigation incentives. Part IV suggests a useful refinement that would involve tailoring the availability of enhanced remedies and penalties to particular industries or technical fields and examines several important issues surrounding the implementation and effects of our approach.

I. THE LOGIC BEHIND PATENTS AND PATENT CHALLENGES

In this Part, we first examine the rationale behind the existence of patents and then the process by which patents are granted and revoked.

A. PATENTS AND PROPORTIONAL REWARDS

The patent system is premised on the idea that an inventor's payoff for innovation should be proportional to the *ex post* social surplus from that innovation.[18] Our evidence is that the payoff to the inventor of possessing a patent is the monopoly profits from having the exclusive right to market her innovation. Monopoly profits are not special in and of themselves. Indeed, monopoly pricing is in general associated with deadweight loss to welfare, which is typically considered a cost of the patent system. However, monopoly profits have the useful feature that they roughly scale with the social surplus from an innovation. In other words, the patent on an

[18] By *ex post* social surplus we mean the consumer plus producer surplus from an invention after it is developed. This surplus excludes the cost of research required to develop the invention.

innovation with twice the social value of another will typically generate twice the monopoly profits of the other. The reason is that, of the factors that determine monopoly profits (the level of demand, the slope of demand, the ability to price discriminate, and competition from other patents), only the level of demand must a priori scale with social value.[19] There is no theoretical reason why the other factors are correlated with the *ex post* social surplus from an invention.

The reason why the patent system seeks to scale rewards with *ex post* social surplus is not primarily that this scaling is, in general, the optimal strategy for encouraging innovation. It is easy enough to see that, for example, if there are diminishing returns to rewards or increasing cost to the use of rewards, then rewards should be roughly concave in the *ex post* surplus from an innovation.[20] Rewards in turn may have diminishing returns because individuals have diminishing marginal utility of income and thus inventors exert less incremental effort as reward rises.[21] And the cost of rewards may be increasing if there are fixed costs to entering a patent race,

---

[19] For a definition of consumer surplus and its relationship to both the demand curve and social welfare, see Daniel T. Slesnick, *Empirical Approaches to the Measurement of Welfare*, 36 J. ECON. LIT. 2108, 2110–13 (1998).

[20] We can demonstrate this with a simple model similar to that employed by Shavell & van Ypersele, *supra* note 1, at 530–32. Suppose the probability of generating an invention is $p(r)$ where r is the reward for an innovation, V is the social value from an invention, $c(r)$ is the cost of providing a reward. Costs might include the costs of a patent race or simply the costs of administering a patent system. The social welfare accounting for the reward is $W = p(r)V - c(r)$. The level of reward that maximizes social welfare satisfies the condition $p'(r)S = c'(r)$, that is, the marginal benefits of rewards must equal their marginal costs. Because rewards r and social value V are complements, i.e., $d^2W/drdV = p'(r) > 0$, and the upper bound on rewards, $r < V$, obviously increases with V, the optimal reward is increasing in the social value of the innovation by Topkis's Monotonicity Theorem. Donald M. Topkis, *Minimizing a Submodular Function on a Lattice*, 26 OPERATIONS RES. 305, 317 (1978). Moreover, it is easily verified that, unless $p(r)$ and $c(r)$ are linear or the upper bound on rewards is binding, the optimal reward is nonlinear.

[21] This concern vanishes if innovations are created by firms which are held by diversified shareholders and thus do not experience diminishing marginal utility of income. Of course, some innovations are made by individuals or privately held firms with limited shareholders. And even in large corporations, agency problems between managers and shareholders can mimic the results from diminishing marginal utility of income. This is most evident when the chief executive is paid a fraction of profits; because the chief executive experiences diminishing returns and controls the corporation, the corporation will behave as if it has diminishing returns. *See also* JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION 42 (1988).

so that multiple inventors only compete when the reward is large enough to cover their fixed costs.

Rather, the main reason why the patent system provides rewards that are proportional to *ex post* social surplus from an innovation is that the government does not know which innovations actually enhance social welfare. The system functions under the assumption that potential inventors and the market have better knowledge about the value of their invention. To encourage inventors to exert most of their effort on developing innovations that are socially productive, the system uses the incentive of a reward that scales with *ex post* social surplus.[22] In other words, the proportional reward from patents is the solution to a principal–agent problem in which the principal is the government and the agent is an inventor. The agent has private information on which project yields the greatest surplus. The government incentivizes the agent to choose the project that is most valuable by giving her a fixed portion of surplus from the project she chooses (and completes).

To illustrate this point, consider a principal–agent model where the agent may work on either of two projects, *0* or *1*. Due to time constraints, the agent cannot work on both. The cost to the agent of working on either project is the same, c, and her reservation wage is normalized to 0. If executed, projects have payoffs of $\theta_0 > 0$ and $\theta_1 = \theta_0 + \epsilon$, respectively, where ϵ takes a value of 1 or -1 with equal probability. Suppose that both principal and agent know the payoff to project *0*, but only the agent knows the value of ϵ before any project is undertaken, that is, θ_0 is common knowledge but the agent has private information on ϵ.[23] The payoff to the principal is $\theta_0 + \mathbb{I}_1\epsilon - w(\theta_0, \mathbb{I}_1)$, where $\mathbb{I}_1$ is an indicator for whether the agent chose project *1* and w is a wage that may depend on information available to the principal, namely, the value of project *0* and whether the agent works on project *0* or *1*. We assume a risk-neutral agent who obtains a payoff of $w(\theta_0, \mathbb{I}_1) - c$ if she works on either project and 0, her reservation wage, if she does not. It is easy to verify that the principal's optimal strategy is to sell the choice between projects (as well as the return to the projects) to the agent for a cost of θ_0 and the agent will accept because $E(\epsilon) \geq R = 0$. This equilibrium also coincides with the first best because

[22] The reward also incorporates the cost of research and development.

[23] The principal may not know the payoff to project *1* either because the payoff goes to some other third party the principal cares about or because it is realized well after a wage must be paid to the agent.

the agent is risk neutral.[24] For our purposes, the result shows that when the agent has private information on the value of projects, she should be incentivized to choose the right one by giving a wage equal to payoff from the projects, even if the cost of research and development are the same for both projects.[25]

B. PATENTABILITY STANDARDS, PATENT APPLICATIONS, AND PATENT CHALLENGES

The above rationale for the patent system assumes, first, that innovations do not occur without rewards and, second, that patents end up in the hands of people who develop innovations. Neither assumption is appropriate in all cases. Some innovations emerge without explicit rewards, or at least without rewards from the government.[26] Prominent examples include academic medical research,[27] freeware software, [28] and fashion

[24] If the agent is risk averse and the principal did not observe ϵ, it would still be the optimal strategy for the principal to offer to sell the choice and payoffs to the agent for θ_0. However, because the agent suffers a utility loss from the random variable ϵ, this strategy is not first best. The principal will not sell for less than θ_0 to provide the agent with some compensating insurance because the principal would do better by simply offering the agent a small positive wage (lower than the contemplated price discount) to work on project *0* and no wage to work on project *1*.

[25] If the agent were risk averse and the principal received a noisy but informative signal about ϵ, the optimal contract would be proportional to (monotonic in) the signal, and thus to ϵ, which proxies for social surplus. There is no a priori reason why the contract would be concave or convex in that signal.

[26] *See, e.g.*, Mark A. Lemley & Ziv Shafir, *Who Chooses Open-Source Software?*, 78 U. CHI. L. REV. 139 (2011) (describing the intersection of academic research and open source software); Ira V. Heffan, Note, *Copyleft: Licensing Collaborative Works in the Digital Age*, 49 STAN. L. REV. 1487, 1504–07 (1997) (discussing the history of freeware and its philosophical opposition to traditional intellectual property rights).

[27] Council on Ethical & Judicial Affairs, American Medical Association, *Ethical Issues in the Patenting of Medical Procedures*, 53 FOOD & DRUG L.J. 341, 349 (1998) (rejecting “the argument that the patenting of medical processes is necessary to enable and promote [medical] procedural advances” and arguing that decades of medical advancement occurred “despite the absence of medical process patents”); Wendy W. Yang, Note, *Patent Policy and Medical Procedure Patents: The Case for Statutory Exclusion from Patentability*, 1 B.U. J. SCI. & TECH. L. 5 par. 17 (1995) (arguing that the inherent incentives of the scientific community can supplant “economic incentive[s] provided by patent monopolies”).

[28] *See* Heffan, *supra* note 26, at 1504–07.

innovations.[29] Moreover, parties who have developed an innovation may not be able to demonstrate that they did so and parties who did not develop an innovation have an incentive to claim they did to obtain market power. If parties innovate in the midst of competition, they may accidentally release the innovation in the public domain before filing the paperwork required to secure patent rights over the innovation.[30] On the flip side, there are frequent complaints about "patent trolls" or "non-practicing entities" (NPEs) who either patent ideas that require little research or purchase patents based on others' research, then do not make any risky investment to develop those patented ideas.[31] Instead, critics contend, an NPE waits until some other party takes the expense and risk to commercialize these ideas and, if the other party is successful, files an infringement suit to extract a portion of the latter's profits.[32]

---

[29] Kal Raustiala & Christopher Sprigman, *The Piracy Paradox: Innovation and Intellectual Property in Fashion Design*, 92 VA. L. REV. 1687, 1689 (2006).

[30] *See* 35 U.S.C. § 102(b) (2006 & Supp. V 2011) (patent is invalid if the invention was published or in public use more than one year before the patent application was filed). In March 2012, the provisions of the America Invents Act amending § 102 section come into effect, generally replacing the one-year provision with a bar of public disclosure or use at any time before the patent's effective filing date, subject to certain exceptions. *See* Pub. L. No. 112-29, §3(b)(1), 125 Stat. 284, 285 (2011).

In the pharmaceutical industry, the problem of innovators being unable to secure patent rights is particularly acute. Pharmaceutical development involves both discovery of a molecule and demonstration that it is effective at treating humans in clinical trial. Patents are granted, however, after discovery and before the trials. Moreover, a single molecule may have multiple medical applications, not all evident when the molecule was discovery. If the idea for a particular application lags substantially behind the discovery, the molecule may enter the public domain (become generic) before the particular application is demonstrated. In other words, the innovative but belated application cannot be protected by patent rights. Benjamin N. Roin, *Unpatentable Drugs and the Standards of Patentability*, 87 TEX. L. REV. 503, 519–21 (2009).

[31] *See, e.g.*, John R. Allison et. al., *Patent Quality and Settlement Among Repeat Patent Litigants*, 99 GEO. L.J. 677, 691–95 (2011) (detailing the practices of "patent trolls" and their litigation habits).

[32] *See* Walter O. Alomar-Jiménez, *Harmonizing Ebay*, 1 U. P.R. BUS. L.J. 17, 24 (2010) (explaining that the modus operandis of patent trolls is "not invest[ing] in research and development (R&D) to create their inventions," purchasing patents "cheaply," monitoring the "technology field of his acquired patents," suing "defendant-infringer[s]," and "demanding a licensing fee"); Christopher A. Harkins, *Fending Off Paper Patents and Patent Trolls: A Novel "Cold Fusion" Defense Because Changing Times Demand It*, 17 ALB. L.J. SCI. & TECH. 407, 407 (2007) ("Instead of commercializing products, patent trolls buy up patents (oftentimes older paper patents), wait for the technology and industry to grow up around the patent, and then use the patent as a holdup device for extorting money

A well-functioning patent system must have a way to ensure that patents are not granted when they are either unnecessary or undeserved. The U.S. patent system, like many others around the world,[33] solves these problems in two complementary ways. First, it sets up criteria to judge when an innovation deserves a patent. Specifically, the creation must be novel, non-obvious, and have some utility.[34] These doctrines, particularly the requirements of novelty and non-obviousness,[35] are meant to determine whether an invention has actually contributed any new knowledge to the world.[36] The patent system then relies upon inventors' incentives in the marketplace to ensure that the invention is socially valuable. If the invention has no value, there will be no market for it and no reason to invest resources in creating it in the first place. If the invention is valuable *and* non-obvious, then the inventor has presumably contributed some valuable knowledge, and with it some social surplus. Patent law's doctrines thus

---

from would-be defendants wishing to avoid the exorbitant costs of defending against an overreaching broadly claimed invention.").

[33] For a description of the European patent system, see generally Patrick Coyle, Note, *Uniform Patent Litigation in the European Union: An Analysis of the Viability of Recent Proposals Aimed at Unifying the European Patent Litigation System*, 11 WASH. U. GLOB. STUD. L. REV. 171 (2012). For comparisons of the United States, European, and Japanese patent systems, see generally John R. Thomas, *Litigation Beyond the Technological Frontier: Comparative Approaches to Multinational Patent Enforcement*, 27 LAW & POL'Y INT'L BUS. 277, 282–88 (1996).

[34] *See* 35 U.S.C. §§ 101 & 103 (2006 & Supp. V 2011).

[35] Utility only plays a meaningful role at the patent-granting stage for biotechnology and chemistry patents. *See* Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1644–46 (2003) ("In the last several decades . . . the utility requirement has lost much of its force<no space>. . . .The only exceptions are biology and chemistry."). Even then, it mainly serves to prevent a firm from patenting a compound (or genetic sequence) at too early a stage. The judgment is that it would be a mistake to allow one firm to lock up a compound before they have any real use for it, removing it from the public domain as a subject for study. But even here the utility hurdle is not all that high. Demonstrated in vitro effects are enough to overcome it. In vivo effects on mice are also enough. *See also In re* Brana, 51 F.3d 1560, 1562–63 (Fed. Cir. 1995). **(**"Even chemical similarity to other effective compounds is enough.").

[36] *See* U.S. CONST. art. I, § 8, cl. 8 (granting Congress the power to "promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries"**).**

provide reasonable standards for judging when a patent is unnecessary or undeserved—at least when they function correctly.[37]

Second, the patent system applies these criteria at two different points during the lifecycle of an innovation.[38] Before a product is commercialized,[39] an inventor may apply for a patent with the Patent and Trademark Office (PTO). The office has been criticized, however, for granting too many patent applications.[40] One reason is limited resources. The PTO's review is relatively cursory because it receives a very large number of applications but is short-staffed.[41] In addition, the PTO only has the information provided by the patent applicant and whatever limited information the patent examiner is able to discover on her own.[42] Another reason is poor incentives. PTO examiners lack the incentives to conduct extensive searches for prior art, and their searches are notoriously less

---

[37] Of course some scholars question whether the standards for patentability perfectly correlate with the necessity and deservedness of patents. We addressed this point in note 7, *supra*.

[38] *See* Andres Sawicki, *Better Mistakes in Patent Law*, 39 FLA. ST. U. L. REV. 735, 746–49 (2012).

[39] Inventors file patent applications prior to commercialization for two basic reasons. First, once a product is commercialized, the PTO might find that it is no longer novel. *See* Richard H. Stern, *The Bundle of Rights Suited to New Technology*, 47 U. PITT. L. REV. 1229, 1249 (1986) ("The subject matter of a patent must be novel, [that is] . . . not disclosed in any prior publication or commercial product . . . ."). Thus commercialization may preclude a successful patent application. Noel Courage, *Sharing the (Genetic) Wealth*, 35 TEX. INT'L L.J. 123, 123–24 (2000) (reviewing KRISHNA R. DRONAMRAJU, BIOLOGICAL AND SOCIAL ISSUES IN BIOTECHNOLOGY SHARING (1998)) ("[S]cientifically novel inventions are not patentable [when] . . . the invention was publicly disclosed prior to filing a patent application. Inventions lacking legal novelty are barred from patent protection . . . ." (footnote omitted)). Second, without a patent, the inventor will face competition during the commercialization process. This may reduce the returns to commercialization. *See* Michael Abramowicz & John F. Duffy, *The Inducement Standard of Patentability*, 120 YALE L.J. 1590, 1648 (2011) ("A patent on a nascent technology . . . can give the patentee 'an incentive to make investments to maximize the value of the patent without fear that the fruits of the investment will produce unpatentable information appropriable by competitors' . . . ." (quoting Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 276 (1977)); F. Scott Kieff, *Property Rights and Property Rules for Commercializing Inventions*, 85 MINN. L. REV. 697, 709–10 (2001) (describing the effects of competition in the commercialization process and noting the additional costs borne by the innovating party due to competition).

[40] *See supra* note 10. However, it is possibly that this lax screening is socially optimal. *See infra* Part III; *see also* Lemley, *supra* note 10, at 1495 n.1.

[41] Lemley, *supra* note 10, at 1499–1500.

[42] *Id.* at 1500.

complete and successful than the searches performed by opposing parties in the course of litigation.[43]

Moreover, the PTO has a stronger incentive to accept applications than reject them. If the agency accepts the patent application, it is unlikely any party will directly complain about the PTO's decision. The work typically has not been marketed and competitors have not emerged. By the time a competitor does emerge, the validity of the patent will have shifted from the jurisdiction of the PTO to that of the federal courts (typically the Federal Circuit), where the matter will likely arise as an infringement action against the competitor. However, if the patent application is rejected, the applicant has an incentive immediately to appeal the PTO's decision. The PTO, seeking to avoid the cost of appeals and the shame of reversal, errs on the side of granting applications.[44]

Patentability criteria (novelty, value, non-obviousness) may be applied a second time after the patent has been granted. A typical case is where a competitor emerges with a product similar to that described in a patent and the patent holder files a lawsuit alleging patent infringement. In order for a patent to have value when asserted against a competitor, it must of course be both valid and infringed. Therefore, as a defense, the competitor may assert that the plaintiff's patent is, in fact, invalid.[45] If the court agrees, the

[43] Merges, *supra* note 10, at 603 (describing patent examiner incentives); Kristen Dietly, Note, *Lightening the Load: Whether the Burden of Proof for Overcoming A Patent's Presumption of Validity Should Be Lowered*, 78 FORDHAM L. REV. 2615, 2655–56 (2010) (discussing the weaknesses of PTO examinations and arguing against the presumption of deference to the PTO).

[44] Jonathan Masur, *Patent Inflation*, 121 YALE L.J. 470, 474 (2011).

[45] As we will reiterate below, we need not define a patent challenge to require an assertion that a patent is invalid. When an infringement suit is filed, the alleged infringer will simultaneously assert that the patent is invalid and, if valid, not infringed. A patent challenge can equally take the form of an argument that the competing product does not infringe the patent. A finding of non-infringement, however, may not always be as damaging to patent holders as a finding of invalidity. After all, the patent holder can always assert the patent against some other party. But in many cases the two have the same functional effect and the same stakes. For instance, "patent trolls" often sue multiple defendants—any firm that might be infringing their patents—simultaneously. Matthew Fawcett & Jeremiah Chan, *March of the Trolls: Footsteps Getting Louder*, 13 INTELL. PROP. L. BULL. 1, 18–19 (2008); *but see* 35 U.S.C. § 299 (2006 & Supp. V 2011) (curbing this practice for lawsuits adjudicated after the passage of the America Invents Act). Part of the reason patent holders bring suit against every conceivable infringer simultaneously is that sequential lawsuits raise the probability of the patent being invalidated in one suit and thus rendered unusable in future lawsuits. In such a suit, a general finding of non-

plaintiff's patent is effectively revoked. This is the canonical example of what we call a "patent challenge" by a private party, in this case a competitor.[46]

This is not the only way a patent challenge can play out. A firm that wishes to challenge a patent can pursue a number of different options, including filing a declaratory judgment action before getting sued for infringement. Alternatively, in the pharmaceutical sector, the challenging firm need not even market a product in order to infringe on a patent. All that is required is the filing of an Abbreviated New Drug Application (ANDA) to obtain Food and Drug Administration approval to market a generic version of a previously approved "branded" drug. Under the Hatch–Waxman Act, the application alone constitutes grounds for the maker of the branded drug to assert a patent infringement claim.[47]

The primary rationale for revisiting a patent grant in this manner is that, for the reasons described above, the PTO grants many patents that are either unnecessary or undeserved. This rationale does not, however, explain why the PTO delegates the authority to trigger a patent challenge to private

---

infringement has the same effect as a finding of invalidity. In addition, a court's interpretation of a patent's claims is often simultaneously determinative of both validity and infringement. The two doctrines thus function frequently as substitutes: if the court interprets the claims broadly, the patent is invalid, and if it interprets them narrowly the patent is not infringed. Accordingly, we will treat these two doctrines largely as substitutes for purposes of the discussion that follows.

[46] We hasten to add that we mean no normative judgment in describing these suits as "challenges." Firms assert a great number of invalid patents every year; they also very commonly attempt to interpret their own patents overly broadly in order to capture as much productive economic conduct as possible. Dan L. Burk & Mark A. Lemley, *Fence Posts or Sign Posts? Rethinking Patent Claim Construction*, 157 U. PA. L. REV. 1743, 1765–70 (2009); Ian Lampl, *Establishing Rules for Resolving* Markman *Failures*, 72 U. CHI. L. REV. 1025, 1038–39 (2005). In these instances, patent challenges are highly socially valuable.

[47] 35 U.S.C. §271(e)(2) (2006 & Supp. V 2011). Under the Hatch–Waxman Act, the generic drug maker must demonstrate that its drug is bioequivalent to the branded drug and certify that it does not infringe on the branded drug's patent before it can obtain the approval of the Food and Drug Administration to market its product through the ANDA process. As a reward for encouraging generic drug entry, the statute gives the first generic maker to file for entry into a market 180 days during which it exclusively may compete against the branded drug. This incentive encourages generic companies to file for entry before the branded drug's patent naturally expires (twenty years after it is granted). Once the generic files an ANDA, the branded company has forty-five days to file an infringement suit. *See* 21 U.S.C. § 355(j)(5)(B) (2006); C. Scott Hemphill, *Paying for Delay: Pharmaceutical Patent Settlement as a Regulatory Design Problem*, 81 N.Y.U. L. REV. 1553, 1560–61 (2006) (summarizing generic entry under Hatch–Waxman).

parties rather than simply revisiting its own decision after some time.[48] In some sense the delegation is consistent with the move by the patent system to decentralize the decision of innovations by allowing inventors to apply for patents rather than offering rewards for innovations chosen by the government. But decentralization by itself is not a virtue.

A better justification is that allowing another private party to challenge a patent addresses the ex parte nature of the PTO approval process.[49] Private parties often have better information about which patents are invalid and stronger incentives to search for relevant information and litigate vigorously. Accordingly, a patent challenge is typically brought by a private party that wishes to market a good similar to that described in the patent. Such a party will only exist if the patented work has positive economic value and will litigate only if the patent imposes an economic cost on them. This reduces the risk that court resources will be wasted on screening zero-cost patents. These arguments are similar to the theoretical arguments given for case-or-controversy and standing requirements that define who may litigate cases generally.[50]

II. THE FLAWS OF PATENT CHALLENGES

Although patent challenges may be necessary to weed out some invalid patents, they have two important flaws. First, they sometimes result in valid patents being invalidated ("false negatives"), thereby discouraging innovation. Second, they sometimes fail against even invalid patents ("false positives"), thereby allowing such patents to continue imposing costs on innovators and consumers. We address these false negative and false positive problems in turn.

We pause to note that here and elsewhere, we use the words "valid" and "invalid" to mean "valid and infringed" and "invalid or not infringed,"

---

[48] The PTO could also restrict its review to patents with positive economic value by only reviewing patents that pay their maintenance fee. *See* 35 U.S.C. § 41(b) & (c) (2006). If a patent has zero value, it is unlikely that the holder will pay the maintenance fee.

[49] For a discussion of the *ex parte* nature of patent examination, see Lemley, *supra* note 10, at 1524–25.

[50] *See* Christian B. Sundquist, *The First Principles of Standing: Privilege, System Justification, and the Predictable Incoherence of Article III*, 1 COLUM. J. RACE & L. 119, 131 (2011) ("One classic defense of standing doctrine relies on the . . . argument that the standing requirements are necessary to ensure that the judicial process is controlled by plaintiffs with a sufficient stake in the litigation.").

respectively. Both infringement and validity are, of course, necessary requirements before a patent holder is entitled to damages or a licensing fee. We describe these patents as valid and invalid largely as a matter of shorthand, but also to highlight the importance of the validity decision to the value of the patent and the rights of the patent owner going forward. A judgment of non-infringement may frustrate a patent owner in a single case but the patent might still be worthwhile against other infringers, whereas a judgment of invalidity ends the patent's useful life.[51] Likewise, a patent challenger who obtains a judgment of non-infringement may or may not aid other potential challengers, while one who succeeds in invalidating a patent has provided a public good that advantages similarly situated parties.

A. CHALLENGES AGAINST VALID PATENTS

When the owners of valid patents are forced into court, the consequences can be severe. At minimum, such parties will be forced to pay litigation costs to defend against validity challenges. More significantly, federal courts may mistakenly invalidate truly valid patents. Not only are the costs of litigation and the risks of improper invalidation significant, they can also exert a differential impact on some of the most important and vulnerable patent holders. First, patent challengers tend to target the most profitable patents, imposing costs disproportionately on the most socially valuable innovations. Second, patent challengers also tend to target the smallest patent holders to maximize their chance of victory. These tendencies diminish the fraction of social surplus from an innovation that the patent holder captures, especially for the most valuable innovations and the smallest innovators. This outcome is inconsistent with the basic premise of the patent system, which is designed to allow inventors to capture a greater percentage of the profits from their inventions as a means of inducing innovation. It thus tends to undermine the value of the system.

1. Mistaken Challenges and Invalidation

Delegation of patent challenges to private parties has some benefits: better knowledge about the costs of bad patents, better incentives to produce

[51] *See Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found.*, 402 U.S. 313 (1971) (holding that once a court has declared a patent invalid, the patent holder is estopped from asserting that the patent is valid as against parties not involved in the original lawsuit).

information on patents, and economizing on court resources. But even well-intentioned private parties are not perfect. They may accidentally challenge patent holders that have truly valid patents. A rough indicator of this—if one assumes that courts make no errors—is that courts validate roughly 55% of patents that are challenged.[52] On its face this suggests that the holders of over half of challenged patents must pay litigation costs even though the patents are valid, eating into the deserved payoff.[53]

Of course, courts may accidentally validate truly invalid patents, implying that the 55% validation rate is an overestimate of the errors that challengers make when initiating suit. But, by the same token, courts may accidentally invalidate truly valid patents, suggesting that the 45% invalidation rate may include cases where valid patent holders were both incorrectly targeted by challengers and were incorrectly found to hold invalid patents by courts.[54] These valid patent holders do not simply pay litigation costs; they also lose all future value from their patents. Thus, incorrect court decisions impose even larger costs than correct court decisions from the perspective of valid patent holders.

Why do we suspect that courts might mistakenly invalidate patents when the popular sentiment in recent scholarship is to bemoan the patent system's lax standards for patenting?[55] Many of the arguments for why courts may accidentally validate invalid patents are also arguments for why they might invalidate valid patents.[56] For example, several scholars criticize

---

[52] John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 205 (1998). At the same time, only 25% of patents in lawsuits are found valid and infringed, which indicates that a somewhat higher percentage of patent challenges have merit. Mark A. Lemley, *The Fractioning of Patent Law*, *in* INTELLECTUAL PROPERTY AND THE COMMON LAW (Shyamkrishna Balganesh, ed., forthcoming 2013) (manuscript at 2–3), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1895681.

[53] The low number of patent trials (roughly one hundred per year) may mask a high number of valid patents that paid a cost due to challengers' targeting errors. *See* Lemley, *supra* note 10, at 1501. Approximately 85% of all patent cases settle before trial. Jay P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 271 (2006).

[54] If courts made no mistakes and patent challengers had perfect foresight, no holder of a valid patent would ever be sued. We can reject this scenario, however, because it cannot explain why 55% of patents are validated.

[55] *See supra* note 10.

[56] Another reason is that patent law's standards for judging whether innovations deserve patents—novelty, non-obviousness, patentability—and the doctrines that

the Federal Circuit, which handles the bulk of appeals in patent litigation cases, for promulgating weak standards for patentability.[57] Others have noted the wide variation in validation rates across industries,[58] circuits,[59] and, within jurisdictions, by whether cases are tried to the bench or to a jury.[60] This may reflect not just randomness in litigation but variation in underlying patentability standards.[61] Most concerning may be that prior appellate reversals and job tenure do not improve the performance of district courts judges in patent cases as measured by subsequent appellate reversals.[62] Thus, judges do not appear to learn how to better apply the law. Each of these flaws could translate to mistakes with either invalid or valid patents.

Whatever the cause, imposing the risk of litigation and mistaken invalidation on valid patents reduces the fraction of social surplus that an inventor obtains through a patent. This results in a weakening of the

---

complement them may not be the best correlates of whether a patent is necessary to obtain an innovation and whether that innovation has positive social value. See our caveat in note 7, *supra*.

[57] *See, e.g.*, Mark D. Janis, *On Courts Herding Cats: Contending with the "Written Description" Requirement (and Other Unruly Patent Disclosure Doctrines)*, 2 WASH. U. J.L. & POL'Y 55, 62–69 (2000) (criticizing the formulation of the written description requirement); R. Polk Wagner & Lee Petherbridge, *Is the Federal Circuit Succeeding? An Empirical Assessment of Judicial Performance*, 152 U. PA. L. REV. 1105, 1111–12 (2004) (finding that the Federal Circuit has been only mildly successful in promulgating a coherent and predictable doctrine of claim construction); David O. Taylor, *Clear but Unconvincing: The Federal Circuit's Invalidity Standard*, 21 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 293, 295–96 (2011).

[58] Allison & Lemley, *supra* note 52, at 221–23 (finding variation in validation rates across industries). Incidentally, litigation rates also vary by industry. John R. Allison et al., *Valuable Patents*, 92 GEO. L.J. 435, 477–78 (2003).

[59] Allison & Lemley, *supra* note 52, at 192 (citing Gloria K. Koenig, PATENT INVALIDITY: A STATISTICAL AND SUBSTANTIVE ANALYSIS (rev. ed. 1980)).

[60] *Id.* at 212 (finding that juries are more likely to validate patents).

[61] Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific*, 17 BERKELEY TECH. L.J. 1155, 1185 (2002) (suggesting the non-obviousness standard might be stated in a manner that is harder for software patents to meet than for biotech patents to meet).

[62] Schwartz, *supra* note 11, at 258–59 (finding neither evidence that district court judges learn from prior appeals of their rulings nor a significant relationship between judicial experience and performance). Of course, it is possible that appellate review in the Federal Circuit is effectively random.

incentive the patent system employs to get inventors to work on the most socially valuable patents.[63]

2. Disproportionate Impacts on the Most Valuable Patents

One of the theoretical benefits of delegating patent challenges is that it economizes on court resources. A challenger should not challenge a patent with zero economic value because the market opened by the challenge is unlikely to have value to the challenger and litigation has positive costs. Unless litigation costs are positive but very small, however, this economizing can go too far. If litigation has high marginal costs or large fixed costs, patent challengers will only go after patents that are sufficiently profitable to cover their litigation costs. If it costs $10 million for a firm to pursue a patent challenge,[64] then its profits after successful litigation have to be at least $10 million to warrant the challenge. Because profits after the introduction of a competitor are lower than profits under a monopoly, the patent holder (which had a monopoly) must lose more than $10 million in payoff from the patent.

The problem is compounded if there are fewer competitors in a position to challenge a patent holder than there are patents. In that case, the challengers, if behaving optimally, will go after the most valuable patents, not just the patents that—once invalidated—offer a payoff sufficient to cover the cost of litigation. To illustrate, suppose that there are two patents that could be challenged, one that provides $30 million in profits for its patent holder and the other that provides $60 million in profits to its patent holder, but only one firm that has the ability to challenge these two patents. Suppose also that market-wide profits after entry of a competitor are two-thirds the previous profits of the patent holder. This implies that successfully challenging the $30 million patent yields revenue of $10

[63] *But see* Sawicki, *supra* note 38, at 766 (citing Jonathan M. Barnett, *Private Protection of Patentable Goods*, 25 CARDOZO L. REV. 1251, 1251–52 (2004)) (noting that the disincentive effects of mistaken invalidations depends on the efficacy of non-patent mechanisms—such as trade secrets—that inventors can use to appropriate the social surplus from their inventions).

[64] This is not an atypical expenditure. *See* Allison & Lemley, *supra* note 52, at 187 (suggesting that $1 million is a low estimate of litigation costs); *see also* Josh Lerner, *Patenting in the Shadow of Competitors*, 38 J. L. & ECON. 463, 470–71 (1995) (reviewing evidence that, while $3.7 billion was spent on basic research in 1991, $1 billion was spent on direct litigation costs, and that indirect costs of patent litigation for public firms averaged $20 million per case).

million[65] for the challenger and challenging the $60 million patent yields the challenger $20 million.[66] If the challenger can only challenge one of these patents, it will choose the higher valued patent. Thus the holder of the higher value patent will face greater litigation risks than the holder of the lower value patent.

Empirical evidence supports the conclusion that challengers target the most valuable patents. For example, generic drug companies tend to infringe on patents that protect markets with the highest sales revenues.[67] As a result, blockbuster drugs have experienced significant reductions in their effective patent life in recent years.[68] Whereas the median loss to a drug patent holder from a successful challenge is roughly $400 million,[69] the average loss to such a firm is roughly $1 billion in firm value. This indicates a strong rightward skew in losses. We can think of no reason why patterns of litigation would differ in other industries.[70]

The result is that delegating challenges to private parties not only reduces the share of social surplus that patent holders appropriate from their innovation but also reduces that share disproportionately for the highest value patents. In other words, the nature of private challenges is such that it disincentivizes the most valuable innovations the most, directly undermining the justification for employing the patent system over one that employs government rewards.

---

[65] $30 \times (2/3) \times (1/2) = 10$. The one-half is because we assume that the two firms split any profits equally.

[66] $60 \times (2/3) \times (1/2) = 20$.

[67] C. Scott Hemphill & Bhaven N. Sampat, *Evergreening, Patent Challenges, and Effective Market Life in Pharmaceuticals*, 31 J. HEALTH ECON. 327, 336 (2012); Laura E. Panattoni, *The Effect of Paragraph IV Decisions and Generic Entry Before Patent Expiration on Brand Pharmaceutical Firms*, 30 J. HEALTH ECON. 126 (2011).

[68] Henry G. Grabowski & Margaret Kyle, *Generic Competition and Market Exclusivity Periods in Pharmaceuticals*, 28 MANAGERIAL DECISION ECON. 491, 497 (2007). Although Hemphill and Sampat, *supra* note 67, at 3–4, claim this is because blockbuster drugs use lower quality patents to extend their patent life (so called "evergreening" strategies), their argument fails to explain why Grabowski and Kyle find that the total market exclusivity period for higher-sale new-molecular entities (NME) is lower, whether measured by mean or median, than that for lower sale NMEs.

[69] Indeed, this value is larger than the average cost of R&D up to the point of market approval. Joseph A. DiMasi et al., *The Price of Innovation: New Estimates of Drug Development Costs*, 22 J. HEALTH ECON. 151, 154 fig. 1 (2003).

[70] Indeed, challenging a patent may be easier in the drug industry than in other industries because the Hatch–Waxman Act gives the first generic firm to file an ANDA 180 days of market exclusivity against other generic entry. *See supra* note 47.

3. Disproportionate Impacts on Smaller Firms

Challenges also disproportionately discourage innovations by smaller firms. Commercially successful firms are not the only entities that obtain and hold valuable patents, and thus they are not the only ones to become targets when they innovate productively. Smaller firms—startup companies and the like—also frequently see their valuable patents attacked, and because the firms are less well-equipped to defend themselves the attacks can be all the more pernicious.

These assaults take two typical forms. First, rather than license or purchase valuable IP from smaller entities, large firms often simply attempt to engineer around it.[71] For instance, imagine that Small Startup has designed a valuable new semiconductor chip. It does not possess the necessary manufacturing capability to actually produce the chip, but it hopes to license the technology to a firm such as Intel. If Intel believes that Small Startup's patent is weak, or that Small Startup will not have the resources to successfully prosecute a suit for patent infringement, Intel could elect to engineer around the patent rather than licensing it. It could simply begin manufacturing a very similar technology with only minor, cosmetic alterations, gambling that it will be able to convince a court that it has not infringed. The more that firms like Small Startup fear this outcome, the less incentive they will have to innovate and enter markets occupied by large incumbents in the first place.

This is not to say that all instances of engineering around a patent are cases in which the patent holder is not receiving a fair return on its invention; to the contrary, many patent holders deserve only narrow patents that are relatively easy to design around. However, there are many instances in which a small firm has in fact introduced a new and useful innovation that a larger entity intends to copy. In these cases, a poorly drafted patent, or

[71] John M. Golden, *Principles for Patent Remedies*, 88 TEX. L. REV. 505, 544 (2010).

simply the application of significant litigation resources,[72] can allow the larger firm to avoid paying for the technology it is borrowing.[73]

This possibility is exacerbated by the resource differential between the two entities. If the fair market value for a startup's patent (or portfolio of patents) is in the tens or even hundreds of millions of dollars, a larger firm might think nothing of spending a few million dollars to defeat it or engineer around it.[74] In theory, the startup should be able to use its patent to obtain the capital necessary to defend the patent. In practice, however, that is often impossible—capital constraints can make a battle with a larger firm very difficult for a startup to win.[75] And if a large company can pay its engineers $3 million to find a way around a patent it would otherwise license for $10 million, it will often do so. The patent is still worth something—the large firm is paying millions to evade it—but the startup is not capturing any of that value.

This is not to say that designing around a patent is always or necessarily a bad thing. Ideally, patents would possess clear boundaries.[76] A great deal of scholarship has been devoted to the virtues of narrow patents that do not claim excessively broad inventive scopes.[77] The point is not that every patent holder should be entitled to capture rents from a wide swath of

---

[72] Michael J. Meurer, *Inventors, Entrepreneurs, and Intellectual Property Law*, 45 HOUS. L. REV. 1201, 1227 (2008) ("Anticompetitive IP lawsuits may succeed because the small firm defendant lacks the information to prove noninfringement or invalidity. Other defendants may settle to avoid litigation costs even though they are confident the plaintiff would lose the lawsuit.").

[73] Stuart J.H. Graham & Ted Sichelman, *Why Do Start-Ups Patent?*, 23 BERKELEY TECH. L.J. 1063, 1086 (2008) ("The rational would-be infringer, when confronted with a patent held by an individual inventor or a small company with limited resources, would likely be more willing to engage in infringing behavior, calculating that the risk of enforcement is lower.").

[74] *See* Golden, *supra* note 71, at 544 (describing the incentives to design around patents rather than license them).

[75] Andrew Beckerman-Rodau, *The Choice Between Patent Protection and Trade Secret Protection: A Legal and Business Decision*, 84 J. PAT. & TRADEMARK OFF. SOC'Y 371, 405 (2002) ("Additionally, even if adequate funds exist to obtain patent protection sufficient capital must exist to enforce patent rights against infringers<no space>. . . . This enables accused infringers to aggressively exploit the limited funds available to a patent owner.").

[76] Tun-Jen Chiang, *Fixing Patent Boundaries*, 108 MICH. L. REV. 523, 560 (2010).

[77] *See, e.g.*, Michael J. Meurer & Craig Allen Nard, *Invention, Refinement and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents*, 93 GEO. L.J. 1947, 1955 (2005) (arguing that broad patents may be used for anticompetitive behavior).

following inventions. Rather, we have in mind situations in which a firm is only trying to patent the invention it has already created, but its patent leaves open the possibility that a competitor will find some way to circumvent the intellectual property right. The competitor is still borrowing the key idea—the "point of novelty"[78]—but has managed to evade the patent. When large firms adopt this approach in lieu of purchasing or licensing the patent, it diminishes the incentives for startups to innovate in the first instance.

The second type of assault on small firms with valuable patents is more direct. In many cases, larger firms will threaten to sue small startups with their own (large) patent portfolios as a means of forcing the startup to license its IP on favorable terms.[79] Imagine two firms operating in the same market: a large firm *L* and a small startup *S*. Suppose *S* invents and patents a new, valuable device that will compete with *L*'s products. If *L* has a large patent portfolio, it can threaten to sue *S* for infringement even if *S*'s new device would not actually infringe *L*'s patents. The very threat of suit—not to mention actual scorched-earth litigation—can be enough to hamper *S*'s ability to attract investors and bring its product to market; venture capitalists and banks will be wary of investing in a firm with the threat of litigation hanging over its head.[80] Accordingly, *L* can force *S* to license its patent to *L* on favorable terms in exchange for cross-licenses to *L*'s patents (which *S* does not necessarily need). *L* then becomes *S*'s competitor, despite *S*'s original patent. This practice has become known as "patent bullying,"[81] and it can diminish the value of innovations made by small startups (to those startups) if they do not have the resources or the patent portfolios to defend themselves.

[78] *See* Mark A. Lemley, Essay, *Point of Novelty*, 105 N.W. U. L. REV (2011).

[79] *See* Graham & Sichelman, *supra* note 73, at 1068 (noting that "some companies use patents to bully their competitors in order to drive up their costs, to gain access to their technology, or to push them out of the market").

[80] *See* Colleen V. Chien, *Of Trolls, Davids, Goliaths, and Kings: Narratives and Evidence in the Litigation of High-Tech Patents*, 87 N.C. L. REV. 1571, 1587–88 (2009) ("The strategic use of patent litigation by established companies to impose distress on their financially disadvantaged rivals has been called patent predation. Such litigation can damage a defendant's credit rating, its relationship with customers, and its reputation with investors, regardless of how the suit is ultimately resolved.").

[81] *See id.* at 1588; Ted Sichelman & Stuart J.H. Graham, *Patenting By Entrepreneurs: An Empirical Study*, 17 MICH. TELECOMM. & TECH. L. REV. 111, 125–27 (2010) (explaining that "it may be that incumbents can strategically exploit weak patents to prevent competition from potential entrants").

These two practices—engineering around and patent bullying—bend the reward curve downward for small firms that successfully innovate. And as with the mechanisms we described in section II.A.2, the more valuable a firm's innovation, the more pronounced this effect will be. More valuable inventions are more valuable targets to competitors; larger firms will be willing to invest greater resources in engineering around a valuable innovation or threatening the startup that created it. Worse still, threats by large competitors will scare away capital and commercial partners from small firms, depriving them the resources they require to fight back on more equal terms. The result will be a diminution of rewards to small firms for successful innovation below the socially optimal level, and consequently a reduction in these firms' innovative efforts.

B. CHALLENGES AGAINST INVALID PATENTS

In some cases, challenges go too far, raising costs on valid patents and discouraging innovation. In other cases, however, they do not go far enough: the PTO grants an unnecessary or undeserved patent and a court nonetheless upholds it. In such cases, an invalid patent continues to generate significant economic costs without the compensating benefit of encouraging innovation. The failure of private parties to challenge all invalid patents is not a direct cost of challenges as much as a failure to fully accomplish the institutional objective of challenges.

The reasons why patent challenges underperform mirror the reasons why they sometimes go too far. Either private parties fail to challenge an invalid patent or courts incorrectly uphold such patents when challenged. Moreover, the failure of private parties can often be attributed to their inclination to challenge only the most profitable patients, because they offer larger rewards if successful, or the smallest patent holders, because they are most easily defeated in court. We address these points in turn.

1. Insufficient Challenges and Validations

Just as private parties lack the complete information required to avoid challenging valid patents, they also may lack the information required to challenge all invalid patents. It is difficult, however, to quantify the extent to which private parties fail to challenge invalid patents. In general, one only observes challenges that are actually filed. Situations in which invalid patents are not challenged are "censored" to scholars.

The second and more important source of leakage with challenges is that, even if an invalid patent is challenged, a court might mistakenly validate it. As previously noted, scholars have criticized the Federal Circuit for weak standards of patentability, courts in general for varying degrees of fluency with patent cases, and judges for failing to learn from experience.[82] The 55% overall court-validation-rate provides some information on the rate at which courts mistakenly validate invalid patents. It is likely that this 55% contains at least some truly valid patents; therefore, this is probably an upper bound on the rate of incorrect validations by courts. In addition, challengers might settle cases rather than litigating them fully, depriving the public of the value of the patent challenge. And here, too, there is likely a disproportionate impact on smaller firms, which are more likely to lack the resources for scorched-earth litigation and thus more likely to settle earlier.

When an invalid patent is never challenged—or, worse, when it is validated by a court—it imposes several types of costs on consumers and other firms. First, if the patent is protecting a commercial good, those goods will continue to be sold at monopoly prices, creating deadweight losses for consumers who cannot afford them.[83] Second, when a court incorrectly finds that an invalid patent is valid and infringed by another inventor, the second (true) inventor must pay damages to the holder of the invalid patent. This functions as a tax on genuine innovation, paid by true innovators to holders of invalid property rights. The result will be a diminution of incentives to innovate on account of this tax. And third, success with invalid patents will cause firms to invest money in acquiring, asserting, and litigating those patents. They will hire lawyers, demand licensing and settlement fees, and litigate at substantial cost.[84] If the patents underlying these activities are invalid and socially worthless, then licensing and litigating them will generate no social value either—they represent pure rent-seeking. The more that courts err and validate invalid patents, the more that they will encourage the wasting of resources on these socially worthless activities.

[82] *See supra* notes 57–62 and accompanying text.

[83] *See, e.g.*, Michael A. Carrier, *Unraveling the Patent-Antitrust Paradox*, 150 U. PA. L. REV. 761, 768 (2002) (explaining the problem of monopolies in the context of patents).

[84] *See* AM. INTELLECTUAL PROP. LAW ASS'N, REPORT OF ECONOMIC SURVEY 22 (2003) (estimating that patent litigation with between $1 million and $25 million at stake costs each side $2 million).

2. Disproportionate Impacts

The tendency of patent challenges to target the most profitable patents may exacerbate the cost of challenges when an underlying patent is truly valid, but it is a positive attribute of challenges when the underlying patent is truly invalid. The most profitable patents are the ones that generate the most deadweight loss because monopoly and oligopoly pricing models suggest that deadweight loss is proportional to monopoly or oligopoly profits.[85] If the underlying patent is invalid, this larger deadweight loss is not offset by a larger incentive to innovate, as it might be with valid patents.[86] Thus, the tendency of challenges to seek out the more profitable invalid patents implies that costly challenges are being appropriately rationed to patents with the largest social cost.

The glass-half-empty view, however, is that the propensity to challenge more profitable patents implies fewer challenges against less profitable patents. Yet even less profitable invalid patents impose deadweight loss. It would improve social welfare if there were more challenges and these challenges targeted the less profitable patents, so long as the social costs from those invalid patents are greater than the cost of litigation against those patents. It is unlikely that all such challenges are occurring because the gains to a private party from challenging such a patent may be less than the deadweight loss from that patent. If the patent challenger wins, competitors other than the challenger may enter the market, lowering profits of the challenger below the level of the deadweight loss. It is this concern that explains why, for example, the Hatch–Waxman Act grants the first generic producer to challenge a drug patent 180 days of market exclusivity (as against other generic producers) if the generic producer prevails in its challenge.[87]

Finally, a corollary of the claim that challengers focus on the valid patents of smaller firms is that challengers tend to avoid challenges against invalid patents held by large firms. These large firms can credibly threaten

[85] See TIROLE, *supra* note 21, at 56.

[86] Again, we are operating under the assumption that legally valid patents are economically valuable in the sense that they encourage innovation. Other scholars have challenged the value of modern patent standards and thus this assumption. We do not dispute their claims. However, we have nothing to add to them and these claims do not undermine our claim that patent challenges are problematic even if legal standards of patentability are largely efficiency promoting. *See supra* note 7.

[87] 21 U.S.C. § 355(j)(5)(B) (2006).

large litigation costs to discourage challengers and sustain invalid patents. Large firms can often extract more rent from any given intellectual property right than smaller firms because, for example, they have greater market share and more information about consumers.

III. ENHANCED PATENT REMEDIES

In the preceding Parts, we described the manner in which patent challenges can result in what amount to taxes on valuable innovation or subsidies for the assertion of invalid patents. In particular, these taxes often fall most heavily upon the innovations that are most valuable and on smaller firms. We suggested that the inevitable result of such taxes and subsidies will be to diminish incentives to innovate for the most productive inventors and to encourage rent extraction by non-innovative firms.

Here, we offer a counterintuitive solution to this problem: raise the stakes of patent lawsuits. Patent holders who manage to prevail against challengers should receive enhanced rewards—heightened damages or extensions of their patent terms—while patent holders who lose at trial should be penalized for suing on the basis of invalid patents. This approach might seem misguided on its face. If courts are liable to err in patent lawsuits, the more appropriate response would seem to be to reduce the impact of those lawsuits. Scholars have suggested such reforms,[88] and in recent years courts appear to have taken steps in this direction.[89]

---

[88] *See, e.g.*, Peter Lee, *The Accession Insight and Patent Infringement Remedies*, 110 MICH. L. REV. 175, 202 (2011) (suggesting that the doctrine of accession, which transfers ownership of property to an individual who increases its value and compensates the prior owner, be applied to patent law to limit injunctive relief and the amount of damages available); Simone A. Rose, *On Purple Pills, Stem Cells, and Other Market Failures: A Case for a Limited Compulsory Licensing Scheme for Patent Property*, 48 HOW. L.J. 579, 584 (2005) (suggesting "a limited compulsory licensing scheme"); Ranganath Sudarshan, *Nuisance-Value Patent Suits: An Economic Model and Proposal*, 25 SANTA CLARA COMPUTER & HIGH TECH. L.J. 159, 177 (2009) (exploring the possibility of damage caps in patent law).

[89] *See, e.g.*, eBay, Inc. v. MercExchange L.L.C., 547 U.S. 388 (2006) (adopting a conservative approach to granting permanent injunctions against parties found to be infringing a patent); Therasense, Inc. v. Becton, Dickinson & Co., 649 F.3d 1276 (Fed. Cir. 2011) (en banc) (establishing a higher, clear-and-convincing-evidence standard for invoking the defense of inequitable conduct, which if established invalidates the entire patent).

Yet as we will demonstrate below, increasing the stakes of patent litigation can have tremendously beneficial effects on private firms' incentives. Firms with valid, valuable patents will realize greater profits on those patents, providing them with additional incentives to innovate and correcting for the costs imposed by improper patent challenges. Firms with invalid patents will face steep penalties if they lose at trial. Those penalties will in turn dissuade them from filing suit in the first place and diminish their ability to extract licensing and settlement fees. The result will be a patent system that comes closer to rewarding genuine innovators but not the holders of socially worthless property rights.

A. CANONICAL SOLUTIONS

The problems we have described in the preceding Parts all center around judicial error.[90] In the face of such problems, commentators have typically suggested the most straightforward solution: invest in accuracy. Proposals for making courts more accurate abound.[91] However, there are well-documented practical and theoretical impediments to this solution. Courts, particularly courts staffed by generalist judges, will always struggle with highly technical patent cases.[92] The judicial process and the limitations it imposes upon gathering outside information and accessing expertise will

[90] For a nice review of the relationship between court errors and optimal tort rules, see Giuseppe Dari-Mattiacci, *Errors and the Functioning of Tort Liability*, 12 SUP. CT. ECON. REV. 165, 167 (2005). An important distinction between Dari-Mattiaci's analysis and ours is that the former examines, in the tort context, how court errors in determining due care and the magnitude of damages, and the relationship between the two, affect precautions, while we examine how errors in determining patent validity (liability) can be overcome by intentionally multiplying damages.

[91] *E.g.*, Rochelle Cooper Dreyfuss, *In Search of Institutional Identity: The Federal Circuit Comes of Age*, 23 BERKELEY TECH. L.J. 787, 804 (2008) ("Another idea would be to abolish the Federal Circuit and reconstitute it as a trial court."); Arti K. Rai, *Engaging Facts and Policy: A Multi-Institutional Approach to Patent System Reform*, 103 COLUM. L. REV. 1035, 1102 (2003) ("I discuss whether the best solution would involve abolishing the Federal Circuit, and having a system of specialized trial courts reviewed by generalist appellate courts.").

[92] Jonathan S. Masur, *Regulating Patents*, 2010 SUP. CT. REV. 275, 307–09 (2011) (discussing the relevant level of expertise among federal judges with patent law); *see also* Peter Lee, *Patent Law and the Two Cultures* 120 YALE L.J. 2, 20–25 (2010) (describing courts' view of their own institutional role).

also impede judicial accuracy.[93] This is not to say that there is no value to investing in greater judicial accuracy; such investments may well be worthwhile. But they are no panacea.

If there are limits to judicial accuracy, perhaps the government could instead invest in accuracy at the Patent and Trademark Office. The patent literature is rife with calls for improving accuracy at the PTO and suggestions for achieving that goal.[94] However, even if this were possible, it would not eliminate the costs involved with patent challenges. Holders of valuable, valid patents might still bear costs as those patents were challenged in court. The necessary second step would be to eliminate post-grant validity challenges entirely, whether in federal court, before the PTO, or elsewhere. Once a patent had been issued by the PTO, it would be considered per se valid and not subject to question in any future proceeding. The elimination of post-grant challenges could be coupled with enhanced review at the PTO, with additional resources devoted to screening out invalid patents before they were ever issued. In theory, then, the costs of patent challenges would be borne most heavily by parties with questionable or invalid patents, not successful innovators.

Yet there are serious problems with this option. The first is that the examination performed by the PTO may never be terribly efficient or effective at weeding out bad patents because PTO examiners have misaligned incentives.[95] As we explained above, they have no incentive to conduct thorough searches of prior art and, even if they did, they would still have greater incentives to grant, rather than reject, patent applications.[96]

[93] *See* Masur, *supra* note 92, at 310–11 (discussing the procedural limitations of courts).

[94] *See, e.g.*, John R. Allison & Starling D. Hunter, *On the Feasibility of Improving Patent Quality One Technology at a Time: The Case of Business Methods*, 21 BERKELEY TECH. L J. 729 (2006); Shubha Ghosh & Jay Kesan, *What Do Patents Purchase? In Search of Optimal Ignorance in the Patent Office*, 40 HOUS. L. REV. 1219 (2004). There have been legislative attempts as well. *See, e.g.*, America Invents Act, 125 Stat. 284 *et seq.* (2011) (creating several mechanisms for post-grant review and reexamination).

[95] *See* JAFFE & LERNER, *supra* note 7, at 136 (describing the incentives facing patent examiners); *cf.* John Bronsteen, *Against Summary Judgment*, 75 GEO. WASH. L. REV. 522, 540–41 (2007) (noting that judges' preferences for leisure time will incline them to grant more motions for summary judgment than would otherwise be appropriate). One study found that patent approval rates spike in September—the month in which the PTO's accounting year closes and examiners are awarded bonuses for processed applications. Gajan Retnasaba, Why It Is Easier To Get a Patent in September? (May 23, 2008) (unpublished manuscript), *available at* http://www.ssrn.com/abstract_id=1121132.

[96] *See supra* section I.B.

Moreover, even if it were possible to correct these incentive problems, it would be tremendously costly to conduct a thorough search of the prior art on each and every patent filed each year.[97] There are simply too many patent applications, and too many of them are economically insignificant and will never be litigated or licensed.[98] As inefficient as patent challenges may be, conducting a full-scale examination of every patent would be even worse. Lastly, in many cases the owners of valuable patents are frustrated not by rulings that their patents are invalid, but instead by rulings that they are not infringed by important competitors.[99] Banning challenges to a patent's validity could hardly solve this problem, and there is no correlative solution to the problems caused by non-infringement. Most importantly, if the patent system is *generally* functioning correctly, it makes little sense to entirely prohibit a set of challenges which will be welfare-enhancing more often than not. Such a remedy is overbroad.

A more moderate alternative would be to imbue patents that have been granted by the PTO with a heavy presumption of validity, diminishing the number of incorrect invalidity determinations in the federal courts. Patents are currently presumed valid when granted, and clear-and-convincing evidence is required before they can be found invalid.[100] This presumption might be strengthened further, to the point where (for instance) a patent could only be invalidated if no reasonable person could find it valid. Some scholars have suggested that the law should move in the opposite direction, toward eliminating the presumption of validity on the basis of the PTO's manifest failings in patent examination,[101] but change now appears very unlikely as the Supreme Court recently affirmed the existing standard.[102] Heightening the standard for invalidity would have many of the same advantages and flaws as simply eliminating challenges entirely, though those effects would be more muted. Absent any reason to believe that such an intermediate solution would decrease the costs of eliminating validity challenges more than it would the benefits, it strikes us as no more advisable than a complete ban.

---

[97] Lemley, *supra* note 10, at 1495 & n.1.
[98] *Id.*
[99] *See supra* Part I.
[100] 35 U.S.C. § 282(a) ("A patent shall be presumed valid."); Microsoft Corp. v. i4i Ltd., 131 S. Ct. 2238 (2011).
[101] *See* Lichtman &. Lemley, *supra* note 10.
[102] *Microsoft Corp.*, 131 S. Ct. at 2243, 2252.

B. AN ALTERNATIVE APPROACH: ENHANCED REWARDS AND PENALTIES

The issue that we have identified is partly the result of erroneous decisions to challenge or not to challenge a patent, and erroneous judicial decisions to uphold an invalid patent or strike down a valid patent. If we cannot easily reduce the probability that a valuable patent will be erroneously defeated at trial, the second-best alternative might be to increase the rewards to holders of valuable patents who prevail at trial. So long as courts are more likely than not to uphold a valid patent, the effect would be the same: to increase the valid patent holder's net expected trial outcome. Similarly, if we cannot easily reduce the probability that an invalid patent will be erroneously validated at trial, the second-best alternative might be to impose additional penalties against the holders of patents who fail at trial. So long as courts are more likely than not to strike down an invalid patent, this would reduce the invalid patent holder's net expected trial outcome.

We thus suggest raising the stakes in patent cases. Our basic idea is simple. If a patent holder sues and wins, the court should award enhanced damages above and beyond the normal measure of damages. If a patent holder sues and loses, the court should assess a substantial monetary penalty against the patent holder. This may seem counterintuitive—if patent lawsuits are not perfectly accurate, it would seem to make little sense to increase the costs involved in errors. Yet so long as courts are better than a coin flip at identifying a patent as valid or invalid, a system of properly designed, supplemental rewards and penalties could simultaneously (1) eliminate the downward pressure on innovative incentives caused by errors within the patent system; and (2) dissuade holders of bad patents from filing suit in the first instance. Most importantly, such a system would tend to benefit holders of valid, valuable patents, and diminish incentives to acquire and assert invalid patents.

How would such a system function? Our idea is to apply standard theories of compensation drawn from tort law.[103] Consider first holders of valid, valuable patents. The purpose behind enhanced rewards is to compensate those owners for the risk that their patents will be improperly invalidated and thus counteract the downward bending of the reward curve for the most valuable inventions. What is the cost of that risk? Suppose that

---

[103] *See* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 167–214 (7th ed. 2007) (describing standard economic theories of tort law).

a typical valid, valuable patent has a probability p of being erroneously invalidated (or erroneously found not infringed) in litigation. Suppose further that litigation imposes a fixed cost of c on the patent holder. The cost of litigation to a patent holder is $c + pvy$, where v is the value of the patent (to its owner) per year and y is the number of useful years of patent life lost to the patent owner.[104] The pvy term represents the risk of early invalidation.[105]

The calculation for losing patent owners is quite similar. Suppose that the owner of an invalid patent sues a genuine innovator for infringement. Suppose the probability that the court errs and upholds an invalid patent is p, the same as the probability that the court strikes down an valid patent.[106] The accused infringer will bear litigation costs of c. It will also face expected damages of pvy. The potential cost to the innovator is therefore $c + pvy$. As in tort law, the way to deter the holder of an invalid patent from imposing such costs on blameless innovators is to force the losing patent owner to internalize the costs of her own lawsuit, namely $c + pvy$.

A problem that arises is that, just as a court cannot perfectly identify whether a patent is valid or invalid, it cannot perfectly identify which parties deserve compensation for exposure to litigation and those which

---

[104] In the interests of simplicity, the model in the text assumes that a patent has constant value over time and does not include discounting. A more general formulation of the costs of litigation to the patent holder is $c + pV$ where V is the net present value of the patent over the period during which the patent was infringed but calculated as of the date the suit was resolved in favor of the patent holder. In later sections we relax some of these simplifying assumptions.

[105] These years of patent life would include both the years remaining on the patent term—if the patent is valuable over those years—and any earlier years in which the patent was infringed. That is, suppose Generic Firm *B* begins infringing Pharmaceutical Firm *A*'s patent eleven years into that patent's twenty-year life. Firm *A* sues Firm *B* for infringement, and seven years later (in the eighteenth year of the patent term) prevails before a jury. Firm *A* would be entitled to supplemental damages based on nine years of patent life. Those nine years represent the period of valuable patent life that was effectively at risk during the lawsuit.

[106] It would be easy to generalize and assume the probability of upholding an invalid patent is $q \neq p$. On the assumption that current rules of patentability are correct, we have no reason to suspect one type of error is more likely than the other. Some scholars assert that patent law has too low standards of patentability, *see supra* note 7. This would suggest a greater likelihood of upholding an invalid patent. Because the thesis of this paper—using enhanced penalties to address flaws in patent challenges—does not depend on the specific underlying rates of error, we proceed under the assumption that error probabilities are symmetric, i.e., $p = q$, in order to simplify our exposition.

should be penalized for imposing litigation risks. All courts know is whether they upheld or struck down a patent. Therefore, any compensation or penalties imposed after patent litigation must be conditioned on verdicts. In doing so, the court must account for the fact that a patent that is upheld may not be valid and a patent that is struck down may not be invalid. This implies that the optimal compensation for a patent that has been upheld is

$$\Pr\{Valid|Upheld\} \times (Enhanced\ remedy\ for\ valid\ patent) \\ +\ \Pr\{Invalid|Upheld\} \times (Penalty\ for\ invalid\ patent)$$

where $Pr\{Valid|Upheld\}$ is the probability that a patent is valid given that it was upheld by the court, and $Pr\{Invalid|Upheld\}$ is the probability that a patent is actually invalid even though it was upheld by the court. That is, the additional compensation provided to a winning patent holder must be reduced to account for the probability that the court erred and the patent was not actually valid. Similarly, the optimal penalty for a patent that has been struck down at trial is

$$\Pr\{Valid|Struck\ down\}\,(Enhanced\ remedy\ for\ valid\ patent) \\ + \Pr\{Invalid|Struck\ down\}\,(Penalty\ for\ invalid\ patent)$$

where $Pr\{Valid|Struck\ down\}$ is the probability that a patent is actually valid even though it was struck down by the court, and $Pr\{Invalid|Struck\ down\}$ is the probability that a patent is invalid given that it was struck down by the court. As with enhanced rewards, because judicial verdicts are imperfect, the penalty must be reduced to account for the possibility that the court erred in striking it down.

To calculate the optimal transfers and penalties, we need to estimate how informative court judgments are. To do this, we can use Bayes's Theorem:

$$\Pr\{Valid|Upheld\} = \frac{\Pr\{Upheld|Valid\} \times \Pr\{Valid\}}{(\Pr\{Upheld|Valid\} \times \Pr\{Valid\} + \Pr\{Upheld|Invalid\} \times \Pr\{Invalid\})}$$

where $Pr\{Upheld|Valid\}$ is the probability that a patent will be upheld given that it is valid, which is equal to $1 - p$; $Pr\{Upheld|Invalid\}$ is the probability that a patent will be upheld given that it is invalid; and $Pr\{Valid\}$ and $Pr\{Invalid\}$ are the probabilities that a patent selected at random will be valid or invalid, respectively. If we assume that $Pr\{Valid\} = Pr\{Invalid\}$,[107] then Bayes's Rule suggests that $\Pr\{Valid|Upheld\}$ simplifies to $(1-p)/(1-p+p) = 1-p$. Because a patent must either be valid or invalid, this also implies that $\Pr\{Invalid|Upheld\} = p$. Using the same approach, we can estimate the probabilities that a patent is valid or invalid if it is struck down. By Bayes's Rule,

$$\Pr\{Invalid|Struck\ down\} = \frac{\Pr\{Struck\ down|Invalid\} \times \Pr\{Invalid\}}{\begin{array}{c}(\Pr\{Struck\ down|Invalid\} \times \Pr\{Invalid\} \\ + \Pr\{Struck\ down|Valid\} \times \Pr\{Valid\})\end{array}}$$

where $Pr\{Struck\ down|Invalid\}$ is the probability that a patent will be struck down given that it is invalid, which is equal to $1-p$; $Pr\{Struck\ down|Valid\}$ is the probability that a patent will be struck down given that it is valid, which is equal to p; and $Pr\{Valid\}$ and $Pr\{Invalid\}$ are the probabilities that a patent selected at random will be valid or invalid, respectively. As before, if we assume that $Pr\{Valid\} = Pr\{Invalid\}$, $\Pr\{Invalid|Struck\ down\}$ simplifies to $(1-p)/(1-p+p) = 1-p$. Again by negative implication, $\Pr\{Valid|Struck\ down\} = p$.

If we plug these values into the equations for optimal compensation, we will find that the optimal reward for a patent upheld at trial is

$$(1-p)(c+pvy) + p(-c - pvy) = (1-2p)(c+pvy).$$

Likewise, the optimal penalty for a patent struck down at trial is

$$p(c+pvy) + (1-p)(-c - pvy) = -(1-2p)(c+pvy).$$

The $(1-2p)$ discount reflects the lack of confidence that court verdicts identify truly valid and invalid patents.

In order for this system of enhanced rewards and penalties to have the desired incentive effects, patent holders and challengers must of course

[107] This is equivalent to assuming that 50% of patents asserted in litigation are invalid or not infringed. This may overstate the true percentage of asserted patents that are valid and infringed. *See supra* note 52. Regardless, we set $Pr\{Valid\} = Pr\{Invalid\}$ only to simplify the mathematics here. The formula will generate correct answers so long as the true values are plugged in.

have some sense of whether the patent at issue is valid and infringed or not. If the parties are entirely mistaken as to the validity of the patent, enhanced rewards and penalties will only skew their behavior even further. However, as we will demonstrate at the end of this section, very little accuracy is actually required of patent holders and challengers. So long as litigants are better than a coin flip at determining what sort of patent is involved, the system of enhanced rewards and penalties will improve litigation incentives.

In order to illustrate the effects of these enhanced rewards and penalties, consider a simple numerical example. Suppose that the typical patent litigation costs $10 million, and the error rate in the typical case is 20%. (It will of course be impossible to determine the error rate in a particular case—doing so would be tantamount to determining the outcome with perfect certainty. Courts will necessarily rely instead upon the typical error rate across cases.)[108] Suppose further that pharmaceutical Firm *A* holds a patent that is worth $10 million per year and has seven years of patent life remaining. That patent is being infringed by generic drug Firm *B*. Firm *A* stands to collect $70 million (the value of damages and an injunction) from Firm *B* if it prevails at trial.[109] Under current rules, if Firm *A* were to prevail, its gain would be:

$$\$70 \text{ million in damages} - \$10 \text{ million in litigation costs} = \$60 \text{ million.}$$

If Firm *A* were defeated, it would pay:

$$\$10 \text{ million in litigation costs.}$$

Under our proposed system of enhanced rewards and benefits, if Firm *A* prevailed, it would collect:

[108] We discuss in detail below the issues of how this error rate might be calculated and what actor or institution might be best equipped and positioned to calculate it.

[109] It will not necessarily always be the case that the value of the patent to its owner is equivalent to the damages that will be assessed against the defendant in the event that the patent owner prevails at trial. It will depend upon a number of factors, including the effect that the entrance of the infringer into the market will have on the patent holder's super-competitive profits. Yet the exact numbers are irrelevant. The system of enhanced rewards and penalties we describe will function similarly irrespective of the precise numerical values involved. We employ similar numbers here only to simplify the mathematics.

$$\$70 \text{ million in damages} - \$10 \text{ million in litigation costs} + (\$10 \text{ million} + \$70 \text{ million} \times 0.2) \times (1 - 0.4) \text{ in enhanced rewards} = \$74.4 \text{ million.}$$

If Firm A lost at trial, it would be forced to pay:

$$\$10 \text{ million in litigation costs} + (\$10 \text{ million} + \$70 \text{ million} \times 0.2) \times (1 - 0.4 \text{ in penalties} = \$24.4 \text{ million.}$$

Now consider the effects that these enhanced rewards and penalties will have upon litigant behavior. Imagine that Firm *A* has a valid, valuable patent, one based upon legitimate research and covering a socially valuable invention. Suppose that Firm *A* has a 75% chance of prevailing against Firm *B* at trial. Under current law, Firm *A*'s expected payoff from litigating is:

$$(\$60 \text{ million net} \times 0.75) - (\$10 \text{ million} \times 0.25) = \$42.5 \text{ million.}$$

With enhanced rewards and penalties, Firm *A*'s expected payoff is:

$$(\$74.4 \text{ million net} \times 0.75) - (\$24.4 \text{ million} \times 0.25) = \$49.7 \text{ million.}$$

As is evident from the calculations above, the enhanced rewards that Firm *A* receives when it succeeds at trial more than balance out the penalties it would be forced to pay if it fails. This is because Firm *A* has a strong patent, one that is more likely than not to be found valid and infringed. The result is that Firm *A* will be almost fully compensated for the risk it runs that its patent will be found invalid each time it is forced to litigate. Firm *A*'s reward curve will be bent back upward, and its incentives to pursue the most socially valuable inventions will be largely restored.

Suppose that Firm *A* instead owns a worthless, invalid patent, one that it should not be asserting against genuine innovators. Nonetheless, because of the possibility of judicial error, there is a 20% chance that Firm *A*'s patent will be found valid and infringed at trial. Under current law, Firm *A*'s expected payoff from litigating is:

$$(\$60 \text{ million net} \times 0.2) - (\$10 \text{ million} \times 0.8) = \$4 \text{ million.}$$

With enhanced rewards and penalties, Firm A's expected payoff from litigating is:

$$(\$74.4 \text{ million} \times 0.2) - (\$24.4 \text{ million} \times 0.8) = -\$4.64 \text{ million}.$$

The addition of enhanced rewards and penalties thus transforms Firm A's decision to litigate this weak patent from a reasonable gamble with a positive payoff into a losing proposition.[110] This will have feedback effects on Firm A's other uses of the patent. Firm A will not be able to extract substantial concessions during licensing negotiations or settlement if it stands to lose money if it proceeds to trial. This is exactly as it should be: we are better off if this patent is never so much as mentioned in a threatening letter, much less asserted at trial.

The analytic discussion that began this section set forth the optimal measure of enhanced rewards and penalties. And the numerical example that followed was of course just one example. But it is crucial to note that enhanced rewards and penalties will improve trial outcomes for owners of valuable patents and harm trial outcomes for owners of invalid patents *any time that courts and litigants are more accurate than flipping coins*. That is, if courts reach the right outcome at trial more than 50% of the time, and litigants know whether they have a valid or invalid patent with at least 50% accuracy, and the enhanced reward or penalty is greater than zero, then the system we describe will benefit holders of valid patents and harm holders of invalid ones.[111] It is not necessary that private parties have any informational advantages over courts (though it would be helpful). Nor is it necessary that the two parties have asymmetric (or symmetric) beliefs or

[110] If the probability of error is larger, for example, $p = 0.25$, then it is possible that a patent holder should receive a reward even if its patent is struck down. The reason is that the ideal reward for a valid patent holder is much larger than the ideal penalty on an invalid patent holder. Even a slight increase in the error rate increases the probability that a verdict striking down a patent is ensnaring a valid patent holder and thus increases the proper transfer, perhaps making it net positive.

[111] If courts are less than 50% accurate—that is, worse than a coin flip—then it makes no sense to have courts deciding patent cases in the first place. If the courts cannot be improved, we would be better off abolishing them and flipping coins. We cannot prove that courts are better than a coin flip, but we suspect (or at least hope) that this is the case.

information. The mechanism of enhanced rewards and penalties will function properly so long as each actor is more accurate than a coin flip. It is thus robust to an extremely wide range of parameters.

To illustrate this point, consider the following extreme numerical example. Suppose courts are 51% accurate at determining whether a patent is valid or invalid; parties are similarly 51% accurate at determining the validity of the patent at suit; and enhanced rewards and penalties are set at $1,000. The holder of a valid patent will likely receive an enhanced trial award of:

$$\$1000 \times 0.51 - \$1000 \times 0.49 = \$20.$$

The holder of an invalid patent will likely be assessed a penalty of:

$$\$1000 \times 0.51 - \$1000 \times 0.49 = \$20.$$

Suppose that a patent owner owns a valid and valuable patent but only knows this with 51% probability. The patent owner will anticipate an enhanced payoff of:

$$\$20 \times 0.51 - \$20 \times 0.49 = \$0.40.$$

In symmetric fashion, an owner of an invalid patent would anticipate an enhanced penalty of –$0.40. The high degree of inaccuracy depresses the impact of the enhanced rewards and penalties. But the principle nonetheless holds: firms with valuable patents will be compensated in part for their litigation risk, while firms with valueless patents will be punished for threatening suit.

By consequence, firms with valid, valuable patents will be even more likely to file suit; firms with invalid, socially worthless patents will be less likely to file suit or assert those patents against genuine innovators. Most importantly, this mechanism will function without any gains in accuracy by the courts or the PTO. Private parties will adjust simply as a matter of their own incentives and their perceived likelihood of success. The result will be fewer lawsuits based upon invalid patents and greater rewards for owners of valid, valuable intellectual property rights.

C. WHO PAYS WHOM?

The discussion thus far has been directed toward properly setting the patent holder's incentives. The goal is to reward holders of valuable patents in order to incentivize further research and development, while simultaneously dissuading owners of worthless patents from filing suit. We have not yet addressed the question of who should pay for patent owners' enhanced rewards and who should be paid when patent owners are assessed enhanced penalties. We take up those questions in this section, and again we reach a counterintuitive conclusion: the structure of payments should not be symmetric. Successful patent challengers should be paid by the patent holders who litigated against them; but victorious patent holders should be paid by the public, rather than the patent challenger who has been found to infringe. This is contrary to the norm in American civil litigation that, aside from the costs of litigation, damages paid by the defendant are equal to the damages received by the plaintiff.[112]

1. Victorious Patent Owner

When a patent owner in possession of a valuable property right prevails at trial, it would seem obvious at first blush that the patent challenger should be forced to pay for the enhanced rewards. After all, it is the challenger who has created the costs in the first instance. However, this might inhibit valuable challenges to bad patents as well—a losing infringer could face very substantial liability under this rule. Challengers to bad patents are providing public goods: if they invalidate a socially harmful property right, a broad spectrum of innovators will reap the benefits.[113] When a court invalidates a patent, it benefits the consumers of the underlying product as well as all competing firms that might wish to enter

[112] *See* Robert D. Cooter, *Unity in Tort, Contract and Property: The Model of Precaution*, 73 CALIF. L. REV. 1, 4–5 (1985). Criminal penalties are an exception. Defendants pay fines to the state rather than to the victim of their crimes. *See, e.g.*, Andrew M. Kenefick, *Note*, *The Constitutionality of Punitive Damages Under the Excessive Fines Clause of the Eighth Amendment*, 85 Mich. L. Re. 1699, 1704 (1987).

[113] Public goods are goods that are non-rival, in that no one can be excluded from using or enjoying them. The invalidation of a bad patent creates a public good in that any competitor to the patent holder, not just the party that invalidated the patent, can now enter the market. For a general discussion of public goods, see HARVEY S. ROSEN, PUBLIC FINANCE 61 (5th ed. 1999).

the relevant market—not just the firm that prevailed in the lawsuit.[114] As a result, a patent challenger only internalizes a small fraction of the benefits of a successful suit. There will be many instances in which it would be socially productive if a patent were challenged but not privately worthwhile for any individual firm, and the challenge will not take place.[115] For instance, suppose that a semiconductor manufacturer holds a patent on a valuable computer chip. The patent, however, is plausibly invalid—the chip is too similar to one that preceded it. It would be socially productive if another semiconductor firm were to challenge the patent and attempt to invalidate it because the price of the chip would fall if the challenge were successful. But it might not be worthwhile for another manufacturer to do so. That firm would bear the full cost of the challenge, including the damages it would have to pay if it were found to be infringing. But it would capture only a fraction of the value of invalidating the patent. The original manufacturer would still control part of the market, and other semiconductor firms that did not participate in the lawsuit could swoop in and capture market share as well. The challenger would be largely providing a benefit to other parties. This is why there are likely too few patent challenges over all. It makes little sense to impose additional taxes on patent challengers and potentially further dissuade them from producing such public goods.

Another possibility is paying for the additional rewards out of general tax revenues. The government could provide a direct monetary award as part of the remedies phase of the litigation. This would avoid distortions in the behavior of potential patent challengers. However, we think a superior solution would be for future consumers of the innovative firm's products to pay for these rewards. The reason is fairness—or at least distributional neutrality. The reason for creating supplemental rewards is to eliminate the disincentive for future innovation imposed by non-meritorious litigation. The beneficiaries of this future innovation are the future consumers of the firm's products. Thus, it is more fair—and there is less needless

[114] *See* Joseph Scott Miller, *Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents*, 19 BERKELEY TECH. L.J. 667, 688 (2004) ("A court judgment that a patent claim is invalid is a public good.").

[115] *See* Christopher R. Leslie, *The Anticompetitive Effects of Unenforced Invalid Patents*, 91 MINN. L. REV. 101, 114 (2006) ("[E]ven invalid patents can create unacceptable litigation risks for potential entrants, raise entry costs, delay entry, deter customers and business partners from contracting with new entrants, and impose inefficiencies while distorting innovation.").

redistribution of wealth—if future consumers pay for these supplemental rewards.

The more difficult question is how to identify and collect from these future consumers. It is, of course, impossible to know precisely who will purchase a firm's products in the future. But the firm's current consumers (or those people who will be purchasers in the near future) might serve as a reasonable proxy. Individuals who are purchasing Apple products today are probably most likely to purchase them in the future; individuals (or businesses) who buy one Dell computer are more likely to purchase another Dell computer; and so forth.[116] The government could conceivably impose a special tax on current or future purchases of a firm's products, with that tax being paid directly to the firm. But this would be counterproductive. It would amount to a state-imposed price hike on a firm's goods, which would presumably decrease the quantities of those goods sold. It is safe to assume that each firm is pricing its own goods so as to maximize profits—or, at least, that the firm is better at doing so than the government would be.[117] Most firms would simply lower their prices, returning the overall price of the product to its prior level. A separate tax, even one paid directly to the firm, would not be an improvement.

A better solution is to extend the terms of the patents at suit.[118] Consider a change in the law that allowed courts to award additional years at the end of a patent term any time a patent holder won a lawsuit for infringement of that patent. The firm would garner further monopoly profits from this extended term, providing additional rewards for its innovation. These rewards would be paid for by consumers who purchase that firm's products in the near future—again, a reasonable proxy for those consumers who will purchase future products made by the same firm.[119]

---

[116] *Cf.* Gideon Parchomovsky & Peter Siegelman, *Towards an Integrated Theory of Intellectual Property*, 88 VA. L. REV. 1455, 1462–63 (2002) (arguing that patents and trademarks can be used to increase the power of brand loyalty and its profitability).

[117] *See, e.g.*, Richard Hynes & Eric A. Posner, *The Law and Economics of Consumer Finance*, 4 AM. L. & ECON. REV. 168, 179 (2002) (explaining that the common effects of price controls are "queuing, unsatisfied demand, and an illegal market . . . .").

[118] A patent is valid for twenty years from the date the patent application was first filed. 35 U.S.C. § 154(a)(2) (2006).

[119] Of course, if the victorious patent holder is not a commercial firm but instead a non-practicing entity that makes profits through patent royalties, it would be the losing firm's customers, rather than the winning firm's customers, who would foot the bill. If the patent was truly novel and innovative and was effectively expropriated by the defendant,

It would not be difficult for a court (or Congress) to properly price the size of this supplemental reward. Recall that the supplemental reward should equal $(c + pvy)/(1-2p)$. The pvy term represents the potential loss of patent term length if that patent is improperly invalidated at trial. If the value of a patent over its lifespan is approximately constant, a court should just extend a patent's term by $(c/v + py)(1-2p)$—the ratio of transaction costs to the annual value of the patent, plus the number of valuable years of the patent term put at risk at trial discounted by the probability that the patent could have been mistakenly invalidated. If litigation costs are small relative to annual profits from a patent—that is, if $c \ll v$—one could ignore the c/v term.[120] This would eliminate any need to calculate precisely the average yearly value of the patent, which a court would have to undertake if it were awarding supplemental monetary damages. The fact that it is the patent itself that is being extended renders this accounting unnecessary. Imagine, for instance, that a patent holder wins an infringement lawsuit based on conduct that began fourteen years after the patent at suit was granted. Six years remain on the patent term. Suppose that the court estimates that the error rate in such cases involving valid patents is 20%.[121] In addition to the usual remedies, the court would extend the patent's term by an additional 0.84 years, or approximately 10 months.[122] There could also be a small additional adjustment (c/v) for the fixed cost c of the patent litigation even if $c \ll v$.[123]

The potential downside of extending the patent term—as opposed to simply paying the patentee from general tax revenues—is that it could lead to increased deadweight economic losses. As we explained above, the virtue

---

this arrangement would be appropriate. But if the patent is not novel and valuable, it presents a problem. Section IV offers a brief sketch of a solution.

[120] For evidence that suggests litigation costs are often substantially lower than annual profits, see *infra* text accompanying note 64.

[121] It may be substantially overoptimistic to think that a court could properly estimate the probability of its own (or the jury's) error. Accordingly, it would probably be best if Congress set this probability by legislation.

[122] (20 year patent term – 14 years elapsed) × 0.2 × (1 – 0.4) = 0.84 years = 10.08 months.

[123] Calculation of this additional adjustment will necessarily be more crude, as the variance in patent values—and thus the variance in the value of additional term length—far exceeds the variance in the cost of patent litigation. Some victorious patent holders will inevitably be paid too much; some will be paid too little. However, in many cases, and for many valuable patents—the ones that will be litigated most frequently—c will be much smaller than v. Accordingly, it will be unnecessary to calculate this additional quantity.

of a patent is that it provides the patentee with a limited monopoly over a good, allowing the patentee to charge monopolistic prices (rather than competitive prices). These higher prices incentivize further research and innovation, but they are also conventionally thought to price some consumers out of the market. When a consumer who would have purchased the good at a competitive price cannot afford it at its monopoly price, there is a resulting deadweight loss in the form of diminished consumer welfare.[124]

This is an important consideration, though it may well be outweighed by the other advantages of extending the patent term. The main case for a twenty-year patent is that the innovative effects from exclusivity of that duration exceeds the deadweight loss from that exclusivity. That argument is typically made while ignoring the litigation costs from patent challenges and assuming no court errors when challenges are litigated. Our patent extension and penalties are designed to return the period of exclusivity to the balance that would be achieved under a twenty-year patent without challenges. Moreover, it is possible that the conventional belief that patents lead to monopoly prices and deadweight losses is overstated. In separate work, we have argued that creative pricing mechanisms can eliminate the deadweight loss associated with patents by ensuring that no (or very few) consumers are priced out of the market for patented goods.[125] We further demonstrate that these pricing mechanisms are in use across a broad spectrum of industries.[126] If we are correct, the case against patent term extensions dissipates substantially.

A second concern with our approach is that it relies upon a questionable assumption: that the value of a patent is approximately constant over time. If a patent declines in value over time, additional years after the end of the typical patent term will be insufficient to compensate the patent holder for the risk of losing earlier years before the end of the patent's life. In the limiting case, a patent may even be worthless by the end of its life. This assumption of constant patent value is fairly conservative for pharmaceuticals and for many types of medical devices, which sell for a higher price and at higher quantities at the end of their life than they do at

[124] *See* Alan J. Meese, *Debunking the Purchaser Welfare Account of Section 2 of the Sherman Act: How Harvard Brought Us a Total Welfare Standard and Why We Should Keep It*, 85 N.Y.U. L. REV. 659, 662 (2010).

[125] *See* Anup Malani & Jonathan Masur, Two-Part Pricing and the Cost of Patents (2012) (unpublished manuscript on file with authors).

[126] *Id.*

the beginning[127] due to advertising.[128] But it does not hold true for most semiconductor and computer patents, which are generally valueless after four or five years as they are outpaced by advances in technology.[129] Accordingly, in designing supplemental remedies it might be necessary to draw distinctions among industries. Owners of pharmaceutical patents would receive an additional patent-term length, while owners of computer-related patents would receive direct monetary payments from the government. We do not pause to dwell on the specifics of this proposal here but instead explore the idea of industry-specific treatment in greater detail in Part IV. In addition, in that Part we suggest modifications to the proposal for supplemental remedies that may obviate the particular issue of whether to grant additional term length or supplemental money damages.

2. Victorious Patent Challenger

As we explained above, there will generally be too few patent challenges because patent challengers cannot fully internalize the benefits of their success. In order to incentivize greater numbers of patent challenges, it makes sense to offer additional rewards or bounties to patent challengers who succeed in court. Accordingly, when a patent challenger prevails and forces a patent owner to pay heightened penalties, those penalties should be paid to the patent challenger. The Hatch–Waxman Act accomplishes this in the context of pharmaceutical patents by offering successful patent challengers 180 days of market exclusivity.[130] In effect, we are suggesting a Hatch–Waxman-type rule for every type of patent.

---

[127] Jayanta Bhattacharya & William B. Vogt, *A Simple Model of Pharmaceutical Price Dynamics*, 46 J.L & ECON. 599 (2003); Richard G. Frank & David S. Salkever, *Generic Entry and the Pricing of Pharmaceuticals*, 6 J. ECON. & MGMT. STRATEGY 75, 86 (1997); Henry. G. Grabowski & John M. Vernon, *Brand Loyalty, Entry, and Price Competition in Pharmaceuticals After the 1984 Drug Act*, 35 J.L & ECON. 331 (1992).

[128] Darius Lakdawalla & Tomas Philipson, *Does Intellectual Property Restrict Output? An Analysis of Pharmaceutical Markets*, 55 J.L & ECON. 151 (2012).

[129] *See* Rosemarie Ham Ziedonis, *Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms*, 50 MGMT. SCI. 804, 808 (2004) (describing the life cycle of high-tech patents).

[130] *See* Hemphill, *supra* note 47, at 1561–67 (2006) (describing in detail the operation of the Hatch–Waxman Act). The Hatch–Waxman Act has several design flaws that make it subject to substantial abuse, *see id.* at 1571–72, but those flaws do not affect the system of enhanced penalties described here.

In addition to the positive incentive effects, forcing defeated patent owners to pay enhanced penalties to victorious challengers would have valuable distributional effects. When a patent owner asserts an invalid patent, it is the competitor—the patent challenger—who stands to be harmed most directly. The invalid patent functions as a mechanism for taxing the genuine innovation in which the competitor has engaged. This in turn harms consumers of the patent challenger's products, who are forced to pay higher prices because of this tax. The enhanced penalties paid to patent challengers would compensate them, in the aggregate, for the risk that they will be unfairly taxed at trial. As the costs of innovation decrease, so too will the prices of patent challengers' goods. The end beneficiaries will be the consumers who have been shouldering the costs of unmeritorious litigation all along.

D. RELATED APPROACHES

1. The English Rule

The solution we offer above bears a family resemblance to a more pedigreed legal mechanism: the "loser pays" or "English Rule." In jurisdictions that have adopted the English Rule, the losing party in litigation must pay the prevailing party's costs and attorneys' fees.[131] If courts and juries in patent cases are accurate most of the time, then applying the English Rule should be generally beneficial: holders of strong patents would see their rewards increase, and holders of weak patents would see theirs diminish. Yet we do not believe that it would be appropriate to straightforwardly apply the English Rule in patent cases.

There is an extensive literature on the advantages and disadvantages of the English Rule, in comparison to the standard American Rule in which both sides bear their own costs,[132] and we will not recapitulate that literature

[131] Thomas D. Rowe, Jr., *The Legal Theory of Attorney Fee Shifting: A Critical Overview*, 1982 DUKE L.J. 651 (1982) ("The English routinely include an assessment for a reasonable attorney's fee in the costs to be borne by a losing party . . . .").

[132] *Id*. ("With its general rule that each side in civil litigation has ultimate responsibility for its own lawyer's fees and that the system will not require the loser to pay anything toward the winner's representation, this country stands in a small minority among the industrialized democracies.").

here.[133] It suffices to note three particular reasons why we do not believe that the English Rule is advisable or adequate. First, it is well understood that the English Rule can cause distortions in litigation behavior by encouraging litigants to increase their litigation expenditures, figuring that their opponents will eventually have to pay.[134] The greater the differential in resources between the two sides, the greater the distortions: a wealthy litigant can threaten to effectively bankrupt a poorer opponent in the event of victory.[135] This is especially important in the context of patent disputes between small start-up companies and large commercial firms, which we described in Part II above. There, the problem is that litigation costs are relatively minor for the large commercial entity but enormous for the much smaller firm, impinging on the smaller firm's ability to effectively defend its patents. Forcing the smaller firm to account for the risk of bearing the larger firm's costs would only exacerbate this problem.

One partial solution might be to institute an "infringer-pays" rule, rather than the neutral English Rule. Under such a rule, a defendant held liable for infringement would pay the plaintiff's costs and fees, but a defeated plaintiff would not be responsible for the defendant's fees.[136] This would shield plaintiff start-up companies from huge losses in the event that they were defeated by larger competitors. However, it would not solve the

---

[133] For a sampling of that literature, see generally John C. Coffee, Jr., *The Regulation of Entrepreneurial Litigation: Balancing Fairness and Efficiency in the Large Class Action*, 54 U. CHI. L. REV. 877, 896 (1987) (explaining that "under the English rule, the interests of attorney and client often can differ: the attorney may want to prosecute a weak or marginal case in order to earn a fee even when the client—who would be liable for the fees of both sides if the action were unsuccessful—would not."); Avery Katz, *Measuring the Demand for Litigation: Is the English Rule Really Cheaper?*, 3 J.L. ECON. & ORG. 143 (1987); John F. Vargo, *The American Rule on Attorney Fee Allocation: The Injured Person's Access to Justice*, 42 AM. U. L. REV. 1567 (1993).

[134] *See* Robert D. Cooter & Daniel L. Rubinfeld, *Economic Analysis of Legal Disputes and Their Resolution*, 27 J. ECON. LIT. 1067 (1989) (summarizing literature on the English Rule and cost-increasing behavior).

[135] *See, e.g.*, Edward F. Sherman, *From "Loser Pays" to Modified Offer of Judgment Rules: Reconciling Incentives to Settle with Access to Justice*, 76 TEX. L. REV. 1863, 1871 (1998) (detailing the wealth differential, but also noting that in the extreme case the effect may be reversed by the presence of a judgment-proof party).

[136] An analogous rule in federal constitutional litigation is that parties that prevail on constitutional claims against a state or local government actor receive attorney's fees, while those that lose do not have to pay the government's fees for defending against those claims. *See* 42 U.S.C. § 1988 (2006) (awarding reasonable attorney fees to parties prevailing in §1983 actions).

correlative problem of larger firms using their extensive portfolios to threaten smaller competitors who possess valuable patents.[137] If those threats became litigation, the smaller firms would still be at risk of financial ruin in the event that they lost. Of course, a finding that the small firm had infringed the larger firm's patents might have exactly the same effect, making the addition of attorneys' fees irrelevant. Accordingly, an infringer-pays rule might be preferable.[138]

Yet this in turn raises the second problem with the English Rule, which is that it could unreasonably diminish incentives to bring patent challenges, including worthwhile challenges to invalid patents.[139] As we explained above, worthwhile patent challenges produce public goods.[140] Consequently, there are generally fewer patent challenges than would be optimal. The English Rule would exacerbate this problem by increasing the penalties for unsuccessful patent challenges, further dissuading potential challengers from litigating. It is for this reason that we advocate paying successful patent owners through patent extensions, rather than forcing patent challengers to shoulder the cost. One could imagine instead instituting a "plaintiff pays" rule, in which only unsuccessful patent plaintiffs must shoulder the other side's costs. But this would merely return to the problems described in the paragraphs above.

It is additionally worth noting that at the time of this writing, a bill has been introduced in the House of Representatives that would establish a similarly asymmetric fee-shifting rule. The "Shield Act" would allow courts to order patent plaintiffs to pay alleged infringers' attorneys' fees—"upon making a determination that the party alleging the infringement of the patent did not have a reasonable likelihood of succeeding."[141] The bill is incomplete and flawed as currently written. It would do nothing to

[137] *See supra* sections II.A.3 & II.B.2.

[138] Another potential wrinkle is an exception to the English Rule for small firms. This would eliminate the possibility that a larger commercial entity could drive a small startup out of business simply by running up litigation costs. Nonetheless, we do not believe that is the paramount concern, as we explained above. A finding that a small startup has infringed a larger firm's patents will likely have the same effect. This wrinkle also would not solve the problem described below.

[139] *See, e.g.*, Thomas D. Rowe, Jr., *Study On Paths to a "Better Way": Litigation, Alternatives, and Accommodation: Background Paper*, 1989 DUKE L.J. 824, 888 (1989) (describing the incentive effects of the English Rule).

[140] *See supra* section III.C.

[141] The Saving High-Tech Innovators from Egregious Legal Disputes (SHIELD) Act of 2012, H.R. 6245, 112th Cong. § 2(a).

compensate holders of valuable patents for the risk that their patents might be erroneously overturned, and in fact it would only reduce the value of those patents further. But it nonetheless represents a promising step in the right direction.

None of these formulations of the fee-shifting rule eliminates the third shortcoming of that approach, which is that litigation costs constitute an extremely small fraction of the potential lost value of a patent in the event of an erroneous judicial decision. Holders of valuable patents face two sorts of costs when they become targets: litigation costs and costs associated with the possibility that a patent will be mistakenly held invalid (or not infringed). The former cost is not insignificant, but it can be dwarfed by the latter. The average patent case that is litigated to final judgment costs each side on the order of $5 million.[142] Yet a single patent—particularly a patent on a successful pharmaceutical—could be worth hundreds of millions or even billions of dollars per year. Consider a patent valued at $500 million that is 10% likely to be invalidated at trial. Each time that patent's owner goes to trial, the litigation costs represent less than 10% of the total expected loss that litigation presents.[143] The English Rule, standing alone, is thus far from a full solution.

2. Insurance Claims and Fraud

To the best of our knowledge there is no legal system that formally replicates the structure of enhanced rewards and penalties that we have described here. However, a rough facsimile has developed in the field of insurance law.[144] If an insurance claimant files a fraudulent claim, she runs the risk of not only having that claim denied but in addition facing criminal penalties for insurance fraud.[145] On the other hand, if an insurance claimant files a valid claim and the insurer unreasonably or fraudulently denies the claim, the claimant can in some cases collect punitive damages from the insurance company.[146] Insurance law thus functions as a two-sided system

---

[142] *See* sources cited in *supra* note 64.

[143] The cost to the patent holder in terms of the risk that the patent will be invalidated is $500 million x 10% = $50 million. If the litigation costs another $5 million, the total cost is $55 million, of which the litigation cost accounts for 9%.

[144] We thank Todd Henderson for suggesting this point.

[145] *See, e.g.*, United States v. Irons, 53 F.3d 947 (8th Cir. 1995).

[146] *See, e.g.*, State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 413–15 (2003).

of enhanced rewards and penalties, at least in extreme cases. We can think of the insurance claimant as standing in the shoes of the patent holder, and the insurance company in the role of the alleged infringer.

There is a potential concern that the possibility of enhanced rewards—punitive damages for fraudulent denial of an insurance claim—could lead insured parties to sandbag their insurers. The insured would withhold key information in an attempt to induce the insurer to deny the claim, and then deploy that information in order to convince a court that the insurer's actions had been fraudulent. A system of enhanced rewards and penalties could conceivably become counterproductive if such behavior migrated to patent law.

To the extent that this concern even exists in insurance law,[147] we do not believe that it would be present in patent law. The premise behind such chicanery is that insured parties possess private information about their claims or themselves. It is this private information that they are able to first withhold and then deploy. But patentees possess no such private information: the patent, its prosecution history, and all relevant prior art are all public. There is nothing for the patentee to withhold, and thus no opportunity for such strategic behavior.[148]

*   *   *

If verdicts in patent cases tend to be inaccurate, it would seem misguided to suggest raising the stakes of those cases. But that is precisely what we propose here. Providing enhanced rewards for patent owners who succeed at trial and enhanced penalties for owners who fail would force owners of valid and invalid patents to self-sort. Owners of valid, valuable

---

[147] Insurance law has mechanisms to deal with this type of problem, such as the fact that the cause of action requires that the insurer have acted in bad faith.

> To show a claim for bad faith, a plaintiff must show the absence of a reasonable basis for denying benefits of the policy and the defendant's knowledge or reckless disregard of the lack of a reasonable basis for denying the claim. It is apparent, then, that the tort of bad faith is an intentional one.

Anderson v. Continental Ins. Co., 271 N.W.2d 368, 376 (1978).

[148] Of course, the patentee could refrain from publicizing the very existence of the patent. But this is precisely what patentees are already doing. *See* Clarisa Long, *Information Costs in Patent and Copyright*, 90 VA. L. REV. 465, 471 (2004).

patents would realize greater rewards from asserting those patents, and thus greater incentives to innovate in the first instance. Owners of invalid patents would have substantially less to gain at trial, and thus less ability to extract rents from genuine innovators. Such a system of heightened rewards and penalties would have substantial salutary effects, even if courts never became more accurate.

IV. Refinements and Caveats

So far we have laid out the basic theory behind our proposal for raising the stakes in patent cases. As we have explained, policymakers need not estimate courts' error rates precisely or calculate reward and penalty multipliers perfectly for our proposal to improve the patent system. All that is necessary is (1) that courts and parties are more than 50% accurate—that is, better than a coin flip; and (2) that the supplemental penalties and rewards are greater than zero. So long as these two conditions are met, any implementation would represent an improvement from the status quo.

But policymakers could in fact do much better if they so chose. In the sections that follow, we highlight a number of refinements that would enhance the effectiveness of our proposal in substantial ways. We also address a number of potential problems with such a system and provide workable solutions. First, we describe how policymakers could tailor a system of enhanced rewards and benefits to particular inventive industries, applying enhanced rewards in industries where patents are generally valuable and the risk of erroneous invalidation is highest, and enhanced penalties in industries characterized by excessive patenting and patent holdup. We then discuss how our proposal might be implemented, and by which institutional actors. We analyze the effects of enhanced rewards and penalties on settlement behavior. And we close by addressing the problems that arise if patent plaintiffs are judgment-proof.

A. Industry-Specific Treatment

The system of enhanced rewards and penalties that we propose will effectively separate holders of valid and invalid patents by adjusting their incentives at trial. But it might be possible to increase the specificity and precision of this system by applying it piecemeal. Certain types of lawsuits would be eligible for enhanced rewards but not penalties; other types of lawsuits could be opened to enhanced penalties but not rewards.

For instance, consider the role of non-practicing entities (NPEs). These are firms that do not actually produce or market any product or service, and often do no research, either. Instead, they simply own patents and use those patents to secure licensing fees or litigation judgments against productive commercial firms. It might be possible to take advantage of the fact that NPEs file a disproportionate share of the lawsuits in which invalid patents are asserted.[149] That is to say, a lawsuit brought by an NPE is more likely to involve an overbroad or invalid patent, or one that contributed no useful innovation, than a lawsuit brought by a commercial firm. If the goal is to avoid benefitting holders of these sorts of patents, NPEs should be separated from other types of patent plaintiffs.

The most direct method for accomplishing this would be to not award enhanced rewards to any patent plaintiff that has not produced a product in the technological area covered by its patent in suit. One could even imagine very particular tests, such as: plaintiffs will not be eligible for enhanced remedies unless they have made $*X* million in sales of a product covered by the patent at suit.

There are a number of problems with this approach. The first is that it would disadvantage small start-up companies and other firms that might eventually *become* commercial entities but have not yet produced products. (It also might be seen as an invitation to infringe patents that have simply not yet been commercialized.) However, this will be an issue with any sorting mechanism based around whether or not a firm is an NPE. A firm's NPE status is not a perfect proxy for the true variable of interest—whether the firm is asserting an invalid or overbroad patent. Using it as a proxy will inevitably lead to errors of over-inclusion (start-up firms) and under-inclusion (commercial firms asserting invalid patents).

The larger flaw with this sorting mechanism is that it would incentivize firms to evade it by simply transferring their patents to other companies. Consider a true patent troll, *P*, a firm that exists only to hold patents and assert them against commercial entities. Imagine that it holds a patent that could plausibly cover a product produced by both Firm *A* and Firm *B*, two large commercial firms. Under this rule, *P* could not obtain enhanced remedies against either Firm *A* or Firm *B*. Instead, it could choose to sell the patent to Firm *A*. Firm *A* would be willing to pay *P* the expected value of

---

[149] Robert P. Merges, *The Trouble with Trolls: Innovation, Rent-Seeking, and Patent Law Reform*, 24 BERKELEY TECH. L.J. 1583, 1602 (2009) (describing the role of non-practicing entities and the types of lawsuits they initiate).

P's suit against Firm *A*, plus the expected value of Firm *A*'s potential suit against Firm *B*. The expected value of that suit *would* involve enhanced remedies, because Firm *A* manufactures a product covered by the patent. *P* would have enhanced the value of its property right simply by transferring it to a different type of party.[150] Not only would this frustrate the purpose of a rule excluding NPEs from accessing enhanced patent remedies, it would also create incentives for firms to expend resources on transfers of intellectual property rights that create no social wealth whatsoever but produce transaction costs.

This concern is mitigated somewhat by the rarity of large commercial competitors litigating infringement suits against one another to judgment.[151] The reason is that they have too much to lose: in many cases, each firm holds a substantial portfolio of patents that the other firm is plausibly infringing.[152] If one firm were to file suit, it would risk a countersuit that could be just as damaging.[153] Both firms would expend millions of dollars in litigation costs without gaining a clear advantage. For this reason, larger commercial firms typically prefer to enter into cross-licensing agreements with one another, rather than litigating to judgment.[154] In addition, the patents held by trolls may be duplicative of patents already held by these large commercial firms—particularly when it comes to devices that are covered by hundreds or even thousands of patents.[155] In the hands of a

[150] Firm *A* and *P* would presumably split the value of the enhanced remedies between them, according to their relative bargaining power. This would mean that part of the "tax" being paid by Firm *B* would go to Firm *A*, where it might well be redirected towards valuable research. A smaller share would go to *P* as a true tax on valuable innovation. This makes this particular solution somewhat more appealing.

[151] This is not to say that large firms do not file suit against one another. *See* Chien, *supra* note 80, at 1572 ("I found that public and large private companies initiated 42% of all lawsuits studied, 28% of the time against other large companies . . . ."). The difference is that monetary awards are rarely the outcome or the objective of large-firm litigation. *See infra* notes 152–54 and accompanying text.

[152] *See* Colleen V. Chien, *From Arms Race to Marketplace: The Complex Patent Ecosystem and Its Implications for the Patent System*, 62 HASTINGS L.J. 297, 299 (2010).

[153] *Id.* ("To guard against the risk of patent litigation, companies acquire patents so they can retaliate against or neutralize threats of suits brought by their competitors.").

[154] *See id.* at 307–10.

[155] *See, e.g.*, David Drummond, *When Patents Attack Android*, GOOGLE OFFICIAL BLOG, http://googleblog.blogspot.com/2011/08/when-patents-attack-android.html (Aug. 3, 2011) ("A smartphone might involve as many as 250,000 (largely questionable) patent claims, and our competitors want to impose a 'tax' for these dubious patents that makes Android devices more expensive for consumers.").

major commercial firm, one additional patent may confer little additional value. Nonetheless, it is always possible that a patent troll would be able to find a higher value commercial buyer for its patents, resulting in wasteful transactions and evasion of the limits on heightened remedies.

In light of this, an industry-focused approach might be superior. This approach takes advantage of the fact that certain industries and areas of technology are characterized by substantial activity by NPEs and patent trolls, and others are not. One could consider awarding enhanced rewards only to victorious patent plaintiffs who hold patents in industries and technical fields that do not involve significant activity by trolls: pharmaceutical drugs, biotechnology, medical devices, chemicals, optics, machinery, and the like.[156] Victorious plaintiffs in industries with significant activity by patent trolls—software, computers, electronics, semiconductors, and similar fields—would be denied access to these enhanced remedies. In symmetric fashion, one could consider limiting the availability of enhanced penalties to industries with substantial troll activity.

This proposal draws upon a literature suggesting that courts are already creating different patent rules for different industries[157] and recommending that Congress or the PTO do the same even more explicitly.[158] There is also a direct analogy to the Supreme Court's approach to injunctive remedies in *eBay v. MercExchange*. There, several concurring justices noted that not all industries, and not all patent plaintiffs, are equivalent.[159] Where there is an especially high risk of patent holdup, or where there is reason to believe that the plaintiff's assertion of its patent rights will hinder rather than promote innovation, the Court hinted that it disfavored injunctive relief.[160] To date, these types of industry-by-industry adjustments have been made largely by

---

[156] *See, e.g.*, Robert E. Thomas, *Vanquishing Copyright Pirates and Patent Trolls: The Divergent Evolution of Copyright and Patent Laws*, 43 AM. BUS. L.J. 689, 693 (2006) ("Large biotechnology, medical, and pharmaceutical companies (biotech/pharma) do not face the same threat that their info-tech counterparts face.").

[157] *See generally id.*

[158] *See* Masur, *supra* note 92, at 277.

[159] eBay, Inc. v. MercExchange L.L.C., 547 U.S. 388, 396 (2006) (Kennedy, J., concurring) ("An industry has developed in which firms use patents not as a basis for producing and selling goods but, instead, primarily for obtaining licensing fees . . . .").

[160] *Id.* at 396–97 ("When the patented invention is but a small component of the product the companies seek to produce and the threat of an injunction is employed simply for undue leverage in negotiations, legal damages may well be sufficient to compensate for the infringement and an injunction may not serve the public interest.").

the courts,[161] though Congress[162] and the PTO have intervened on very limited occasions.[163] It may be appropriate for the courts to take the lead again here, or it might be advantageous for Congress or an administrative agency to play a significant role. The institutional details are interesting and important but beyond the scope of this Article. What is important is that industry-by-industry distinctions such as the one we are proposing are hardly foreign to patent law.

Could this arrangement similarly be gamed by opportunistic patent trolls? One option would be for trolls to simply cease activity in a given industry, goading courts into offering enhanced remedies, before resuming litigious activities. Yet this is highly improbable for any number of reasons. If trolls could convince courts to allow enhanced remedies by ceasing activity, courts would presumably turn the spigot back off once trolls resumed litigating. Patent trolls would also cost themselves a tremendous amount of money by ceasing activity simply in order to tap into greater enhanced remedies at some future date. And patent trolls would also have to engage in a significant amount of concerted action (actually, nonaction) in order to implement this plan. This creates a severe collective action problem—any given patent troll would benefit enormously from defecting from an agreement and continuing to litigate.

A more likely possibility is that trolls might migrate from the technical fields they currently inhabit to other industries (such as pharmaceuticals) where the remedies are more generous. If it became standard practice for patent trolls to "follow the money" in this fashion,[164] any strategy that relied upon distinctions between industries would be quickly eroded.

However, it is extremely difficult—if not impossible—for patent trolls to take up residence within another industry or technical field. The reason has nothing to do with the expertise within those firms, or the types of patents owned by trolls. If those were the barriers, trolls could simply hire experts in other technical areas and purchase other patents. Rather, some industries are simply more conducive to predatory patent behavior than others. The reason appears to be that it is easier in some fields than in others to specify an invention for purposes of a patent. In the pharmaceutical and

---

[161] *See* DAN L. BURK & MARK A. LEMLEY, THE PATENT CRISIS AND HOW THE COURTS CAN SOLVE IT 8–11 (2009).

[162] *See* 21 U.S.C. § 355(j)(5)(B) (2006).

[163] *See* Examination Guidelines, 66 Fed. Reg. 1092, 1092 (Jan. 5, 2001).

[164] *See Memorable Quotes For* All the President's Men, INTERNET MOVIE DATABASE, http://www.imdb.com/title/tt0074119/quotes (last visited Nov. 1, 2012).

chemical industries, for instance, a patentee can specify a drug or chemical with a great deal of precision by describing the molecule involved. Any given invention is usually covered by only a small number of significant patents[165]—hence the often-stated principle of "one molecule, one patent."[166] Consequently, old patents can rarely be re-interpreted in broad fashion to cover new inventions. The opportunities for trolls are greatly limited. It is for this reason that these industries—and others, such as machinery and optics—are not generally thought to have many trolls currently operating. If trolls could gain a foothold litigating in these fields, they would already have done so; there is no reason for them to have artificially confined their activity to certain industries. The relative absence of troll-like behavior is therefore best understood as a function of the way in which patents interact with and describe the relevant technology.

Accordingly, we believe that it will be possible to obtain the advantages of enhanced remedies while minimizing the harm done by patent trolls by limiting these enhanced remedies by industry. There will be some definitional issues at the margins—parties may argue over whether a particular patent covers computers or machinery, for instance—but these are the types of issues that courts are well-equipped to decide.[167] The distinctions we seek to draw are necessarily crude, but here these crude distinctions may function better than either finer distinctions, which can be gamed, or the status quo.

B. IMPLEMENTATION AND MEASUREMENT

The proposal we have described is not one that could be easily implemented by courts under current law. To begin with, there is no

---

[165] A drug might have multiple patents on the form in which it is delivered or the dosage rate, but only one patent on the underlying molecular form. *See, e.g.*, Mark A. Lemley, *Ten Things To Do About Patent Holdup of Standards (and One* Not *To*), 48 B.C. L. REV. 149, 150 (2007) ("In particular, the one central fact about the information technology ("IT") sector—including the Internet, semiconductors, telecommunications, computer hardware, and computer software—is the multiplicity of patents that developers must deal with. This is not a problem pharmaceutical companies generally encounter.").

[166] *Id.*

[167] For that matter, the PTO classifies every patent by technology area as a matter of course. It does so in order to assign patent applications to the proper examiners when they are filed. If courts prove incapable of drawing consistent and meaningful distinctions between technological fields, the PTO might prove to be a worthy substitute. *See* Masur, *supra* note 92, at 312.

provision in law that would allow courts to assess the types of enhanced rewards and penalties that we advocate. The Patent Act permits courts to "increase the damages [found by a jury] up to three times the amount found or assessed,"[168] and courts have employed this provision to assess treble damages in cases of "willful" infringement.[169] However, this provision only allow courts to increase the amount paid by the defendant to the plaintiff.

The Patent Act also permits courts to award attorneys' fees in "exceptional cases."[170] The problem with this provision of the statute is that we propose awarding enhanced awards and penalties in essentially every case, rather than just exceptional ones. Also, as we explained during our discussion of the English Rule, attorneys' fees will generally be too small to have the desired effect.[171]

In addition, our proposal relies on determining the (approximate) accuracy of courts across patent cases. We have no illusion of courts' ability to ascertain this figure themselves. The judges of a court do not have the time to scrutinize one another's opinions for error, nor would they be eager to point out their colleagues' errors even if they discovered them.[172]

Accordingly, legislative or administrative action will be necessary. Congress could implement such an arrangement by legislation, or (perhaps

---

[168] 35 U.S.C. § 284 (2006).

[169] *See, e.g.*, *In re* Seagate Tech. LLC, 497 F.3d 1360, 1367 (Fed. Cir. 2007) ("Absent a statutory guide, we have held that an award of enhanced damages requires a showing of willful infringement.").

[170] 35 U.S.C. § 285 (2006).

[171] *See supra* Part III.D.1. There is also the possibility of Rule 11 sanctions. *See* FED. R. CIV. P. 11(c). Yet Rule 11 sanctions are typically quite small (in the thousands or tens of thousands of dollars), rather than the millions we suggest awarding in enhanced penalties. Herbert Kritzer, Lawrence Marshall & Frances Kahn Zemans, *Rule 11: Moving Beyond the Cosmic Anecdote*, 75 Judicature 269, 270 (1992) ("The median sanction is $2,500--that is to say 50 per cent of monetary sanctions awarded were for less than $2,500. A full two-thirds of the sanction awards were for $5,000 or less."); *see also* Joseph J. Brecher, *The Public Interest and Intimidation Suits: A New Approach*, 28 SANTA CLARA L. REV. 105, 137 (1988) ("Furthermore, the typical sanction award is a relatively minimal amount and is usually imposed against an attorney, rather than his client."). Moreover, Rule 11 sanctions are traditionally applied only in extraordinary circumstances, whereas we propose enhanced penalties in every case where a patent is found invalid. Robert Kasunic, *Preserving the Traditional Contours of Copyright*, 30 COLUM. J.L. & ARTS 397, 427 nn.95–96 (2007) (noting that that the "imposition of [Rule 11] sanctions is rare" and courts have "typically reserved [sanctions] only for egregious cases").

[172] *See* Richard A. Posner, HOW JUDGES THINK 123–35 (2008) (describing the value that judges place on collegiality).

preferably) could delegate the task to an administrative agency.[173] In either event, a panel of outside experts should be tasked with reviewing a random sample of completed patent cases and determining courts' error rate.

C. SETTLEMENT

Thus far our discussion has focused on outcomes at trial, and we have only alluded to licensing and settlement negotiations. These negotiations make up a significant fraction of the economic activity surrounding patents, but our general neglect of them has been deliberate. The reason is simple: settlement and licensing occur in the shadow of expected trial outcomes.[174] The more that a plaintiff and defendant believe the plaintiff will win at trial, the greater the amount they will settle for (in the event that they settle).[175] The less the plaintiff has to gain at trial, the lower the settlement price.[176] Accordingly, a system of enhanced rewards and penalties will increase the amount that owners of valid, valuable patents will earn in licensing negotiations and decrease the amount that owners of invalid patents will be able to extract. These changes in settlement outcomes will mirror the changes in expected trial outcomes.[177]

Nor should a symmetric system of enhanced rewards and penalties affect the likelihood of settlement. Settlement is valuable because it allows both sides to avoid the substantial expense involved in litigating.[178] When parties fail to settle, it is typically because they disagree on the likely outcome of the case.[179] If either party has private information that leads her

[173] *See* Masur, *supra* note 92, at 279 (suggesting that the PTO be afforded general rulemaking authority).

[174] *See* Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L. J. 950 (1979) (setting forth the seminal model of settlement bargaining); *see also* Robert Cooter et al., *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior*, 11 J. LEGAL STUD. 225 (1982) (expanding on the Mnookin & Kornhauser model); Steven Shavell, *Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs*, 11 J. LEGAL STUD. 55 (1982) (same).

[175] Shavell, *supra* note 174, at 67.

[176] *Id.*

[177] *See* Mnookin & Kornhauser, *supra* note 174, at 971 (elaborating on this point).

[178] Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 418 (1973).

[179] George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 9–12 (1984).

to believe that she is more likely to win, the two sides will not be able to come to an agreement.[180] That is, the decision to litigate rather than settling is driven by differences in information. Here, the system of enhanced rewards and penalties introduces no private or asymmetric information. Both parties will have the same information regarding the multipliers used to calculate enhanced rewards and penalties, and both parties will be able to perform the same calculations to the same degree of accuracy. If the parties would be inclined to settle absent a system of enhanced rewards and penalties, the introduction of that system will not dissuade them.

The one factor that could impact settlement is the proposal we outlined above for paying victorious patent plaintiffs from general tax revenues or via a patent term extension, rather than forcing losing defendants to pay. The reason is that this third-party payment to plaintiffs decouples plaintiffs' and defendants' incentives, making trial worth more to plaintiffs if they have valuable patents. Consider a simple numerical example. Suppose that a plaintiff is 80% likely to prevail at trial, damages in the case will be $100 million, and the cost of litigating will be $10 million for each side. Absent a system of enhanced rewards and penalties, plaintiff's expected payoff from going to trial would be:

$$\$100 \text{ million} \times 0.8 - \$10 \text{ million} = \$70 \text{ million.}$$

Defendant's expected payoff from going to trial would be:

$$-\$100 \text{ million} \times 0.8 - \$10 \text{ million} = -\$90 \text{ million.}$$

Plaintiff and defendant thus should be willing to settle for any amount between $70 million and $90 million.

[180] *See id.* at 9 (describing the role of differential beliefs in settlement). Note that in the text we only consider the case where there is symmetric uncertainty that leads to differential beliefs. We do not consider the case where one side has systematically better information. Even if one thought, for example, that the patent holder had better information on validity, the outcome of the settlement game depends on who is able to make a final take-it-or-leave-it offer. *See, e.g.*, Lucian Bebchuk, *Litigation and Settlement under Imperfect Information*, 15 RAND J. OF ECON. 404, 405, 414 (1984) (describing model in which the less informed party makes final offer); Jennifer F. Reinganum & Louis L. Wilde, *Settlement, Litigation, and the Allocation of Litigation Costs*, 17 RAND J. ECON. 557, 558–59 (1986) (describing model in which the better informed party has ability to make final offer). We certainly do not know a priori who has that sort of bargaining power.

Now consider the effect of adding enhanced rewards and penalties equal to $10 million, with the reward coming via a patent term extension rather than from defendants. Plaintiff's expected payoff from going to trial would be:

$$\$110 \text{ million} \times 0.8 - \$10 \text{ million} \times 0.2 - \$10 \text{ million} = \$76 \text{ million}.$$

Note that the plaintiff obtains an additional $10 million if she wins (with probability 0.8), but loses $10 million if she loses (with probability 0.2). Defendant's expected payoff from going to trial would be:

$$-\$100 \text{ million} \times 0.8 + \$10 \text{ million} \times 0.2 - \$10 \text{ million} = \$88 \text{ million}.$$

Whereas the plaintiff's enhanced gain does not come from the defendant's pocket, her loss does go to the defendant. Plaintiff and defendant would then be willing to settle only for any amount between $76 million and $88 million. The available bargaining range has shrunk from $20 million ($90 million – $70 million) to $12 million ($88 million – $76 million). This will decrease the likelihood of settlement.[181]

The result will be fewer settlements on the whole and more patent trials. This would most likely increase social welfare with respect to invalid or dubious patents, because those patents will be invalidated at trial and cannot later be asserted against other alleged infringers. But it will decrease social welfare with respect to valid patents by forcing the parties to incur litigation costs that might otherwise have been avoided. The net effects will be ambiguous and will depend on the percentages of asserted patents that are valid and invalid. This might constitute an argument for eschewing the asymmetric mechanism we recommend in which plaintiffs pay enhanced penalties to patent challengers but are paid enhanced remedies by the public.

The important point is that enhanced rewards and penalties will improve settlement outcomes for holders of valid patents and worsen them for holders of invalid patents. For instance, in the numerical example above, the

[181] *See* Russell Korobkin, *Aspirations and Settlement*, 88 CORNELL L. REV. 1, 6 (2002); *see also* Samuel Issacharoff & George Loewenstein, *Second Thoughts About Summary Judgment*, 100 YALE L.J. 73, 101 (1990) ("[I]t is natural—as well as customary in the legal and economic literature—to assume that the likelihood of settlement is positively related to the width of the settlement zone.").

midpoint of the bargaining range absent enhanced rewards and penalties is $80 million. That midpoint rises to $82 million once enhanced rewards and penalties are introduced. The salutary effects of enhanced rewards and penalties will persist irrespective of whether litigants settle their cases or proceed to trial.

D. RISK AVERSION

One concern with our proposal for enhanced rewards and penalties is that they increase the amount of litigation risk to which the patent holder and challenger are exposed. To the extent that these parties are risk averse,[182] our proposal would appear to reduce these parties' welfare. This is not a critical objection to our basic proposal because parties can insure against litigation risk by settling cases. In the basic settlement model we employed in the last section, a reduction in utility or corporate welfare due to risk aversion to larger litigation stakes operates just like a litigation cost—it reduces the patent holder's gain from litigating and increases the patent challenger's loss to litigating. The combined effect is to reduce the parties' utility from going to trial, making settlement even more attractive. If the parties settle, they do not face the same litigation risk.

Interestingly, risk aversion offsets part of the problem raised by the variant of our proposal that has enhanced rewards funded partly by taxpayers or future consumers. In the previous section we explained that this decoupling of enhanced rewards from defendant payments increases the amount that the plaintiff expects to obtain at trial more than it increases the amount that the defendant expects to pay at trial. Because the scope for settlement is proportional to the amount that the defendant expects to pay above and beyond what the plaintiff expects to receive, decoupled enhanced rewards reduce the scope for settlement. For reasons we gave in the last paragraph, however, decoupled enhanced rewards also increase the plaintiff's cost of litigation due to risk aversion. This reduces the payoff that patent holders can expect from trial, offsetting some of the increased return from decoupled enhanced rewards. The offset is not complete because the enhanced reward is similar to a strictly positive value lottery. Positive value

[182] Corporate parties may behave in a risk-averse fashion either because owners or managers are actually risk averse or because they face a risk that threatens a company's existence. In general, smaller firms are more likely to behave in a risk averse fashion than larger firms, either because smaller firms have fewer and perhaps less-diversified owners or because any given level of risk is more likely to bankrupt a small firm than a big firm.

lotteries—for example, heads you win $5 and tails you win $10—have positive value to both risk-averse and non-risk-averse individuals. That is, both individuals would pay a positive amount to face the lottery, though the risk-averse individual would be willing to pay less than the non-risk-averse individual to face it. In summary, risk aversion reduces some (but not all) of the anticipated returns to enhanced rewards and penalties but also ameliorates some of the reduction in settlement rates caused by decoupled enhanced rewards.

E. INSOLVENT PLAINTIFFS AND SHAM LAWSUITS

We close with two relatively discrete but important issues. First, it is essential that patent plaintiffs have the capacity to pay enhanced penalties if they lose at trial. Patent plaintiffs could conceivably evade their responsibility for enhanced penalties by transferring their patents to under-funded shell corporations and then using those corporations to bring suit. If the suit failed, the corporation would not have the resources to pay the penalty judgment.

This is a real problem but one that is easily addressed. Patent plaintiffs should be forced to either post a litigation bond or purchase insurance against being assessed an enhanced penalty.[183] The bond or insurance would be pegged to the damages demanded by the plaintiff: the greater the damages, the greater the bond or insurance.

Second, and lastly, it is always possible that patent holders will take advantage of the prospect of supplemental rewards by arranging sham lawsuits, which they then win at trial. Sham suits present a real concern, but they are hardly unique to this context—patent law offers numerous opportunities for patent holders and challengers to gain advantages via sham lawsuits.[184] As in other areas of patent law, they can be policed through other means, principally an examination of connections between the plaintiff and defendant in a given suit and the parties with economic interests on both sides of the case.[185] The PTO has already begun to take

---

[183] *See generally* Robert J. Rhee, *Bonding Limited Liability*, 51 WM. & MARY L. REV. 1417 (2011) (describing the use of litigation bonds and similar instruments).

[184] *See* Timothy Denny Greene, *"All Substantial Rights": Toward Sensible Patent Licensee Standing*, 22 FED. CIR. BAR J. 1 (2012).

[185] Federal law already requires that any civil action be "prosecuted in the name of the real party in interest." FED. R. CIV. P. 17.

steps in this direction.[186] The threat of fraud thus does not provide an adequate basis for rejecting supplemental patent remedies.

CONCLUSION

In this paper we have documented how patent challenges, patent law's *ex post* attempt to correct mistakes made by the PTO, can undermine the initial goals of patents. Because of mistakes in the selection of patents by challengers and errors by courts in determining which patents are valid, patent challenges can penalize valid patents, thereby discouraging innovation, and protect some invalid patents, sustaining deadweight loss and taxing true innovation. One solution to the problem is to make patent challengers and courts more accurate. Where that is difficult, we propose an alternative, counterintuitive solution: increase the stakes in patent challenges. We recommend that patent owners whose patents are upheld at trial be given a reward—in the form of a patent extension—on top of the damages they usually get in court. Similarly, patent owners whose patents are held invalid by a court should be forced to pay a penalty to patent challengers. This will increase the wedge between the payoffs of having a patent upheld in courts and having it struck down by a court. So long as courts are better than a coin flip at determining whether a patent is truly valid, this approach will reward holders of truly valid patents and punish holders of truly invalid patents. Courts need not be perfectly accurate for our solution to work. Indeed, the higher stakes are a substitute for more accuracy.

Although our proposal seems bold, it is actually fairly narrow. Whereas we only use higher stakes to correct for skewed incentives created by imperfect patent challenges, higher stakes can also be used to correct for other flaws in the patent system, including the possibility that supra-competitive profits from market exclusivity may not fully capture the full social gains from innovation[187] or that market exclusivity may discourage follow-on innovation.[188] In some sense, this is not at all surprising. Because the threat of damages awarded by court is ultimately how patent laws are

---

[186] *See* Tony Dutra, *Stakeholders at PTO Roundtable Recommend Reporting Real Party in Interest Changes Only*, 85 Patent, Trademark & Copyright Journal 419 (January 24, 2013).

[187] *See* Shavell & van Ypersele, *supra* note 1, at 529.

[188] *See* Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCIENCE 698, 698 (1998).

enforced, those laws can be changed substantially by altering damages that courts award.

Although manipulating the stakes in patent challenges can be used to tackle broader issues of patent policy, the choice between addressing the problem of imperfect challenges by increasing the accuracy of courts or by changing the patent approval process does not depend on the resolution of those underlying questions about patent policy. Whether one thinks that current patent law undercompensates for innovation or that its breadth deters future innovation, improvements in accuracy and increases in stakes will both address the problem. The point we wish to highlight is that increased accuracy and increased stakes are substitutes.

Although we make our two basic observations—that litigation is mistake-prone and that higher stakes can correct some of that error—in the context of patent challenges, it also applies outside patent law. In general, it is important to model not just the incentive effects of a given legal rule but also the incentives to litigate that rule. Those litigation incentives can introduce errors into application of the basic legal rule, reducing the efficacy of the rule. Moreover, litigation errors can be corrected either by directly improving the accuracy of litigation or, surprisingly in some cases, by increasing the stakes in litigation. Increasing stakes is a plausible substitute for greater accuracy when litigants and courts, while not perfectly accurate, are at least better than random at identifying truly legal and illegal behavior.