Most patent scholars agree that the Patent and Trademark Office grants too many invalid patents and that these patents impose a significant tax upon industry and technological innovation. Although policymakers and scholars have proposed various ways to address this problem, including better ex ante review by patent examiners and various forms of ex post administrative review, district courts invalidating patents in litigation remain a core defense against bad patents.

This article analyzes a previously unidentified impediment to the use of district courts to invalidate patents. Nearly every patent lawsuit rises or falls on one of two defenses: invalidity or noninfringement. Invalidity and noninfringement are distinct legal and factual issues that are usually analyzed separately by scholars. Yet as the article explains, the two issues are closely related, creating a series of tradeoffs and asymmetries that lead many patent defendants to focus on noninfringement instead of invalidity. The net effect of these tradeoffs and asymmetries is that patent defendants often have an incentive to argue noninfringement instead of invalidity, leading courts to invalidate fewer patents than they should. This exacerbates the problem of invalid patents, making it harder for individuals and companies to create new products and services.

The article concludes by proposing three reforms to help restore the balance between invalidity and noninfringement. First, eliminating the elevated burden of proof for invalidity would remove one significant asymmetry that makes it harder to prevail on invalidity. Second, a bifurcation rule giving defendants the option to defer infringement issues until after validity has been decided would help litigants develop coherent trial narratives while allowing them to focus on validity issues early in a case. And third, a new cause of action for an accounting, brought against industry competitors by a litigant that successfully invalidates a patent, would help eliminate the collective-action problem posed by invalidity’s public-good nature.

† Harry A. Bigelow Teaching Fellow and Lecturer in Law, the University of Chicago Law School. For helpful comments and discussions, I am indebted to Alexander Boni-Saenz, Anthony Casey, Anya Bernstein, Daniel Abebe, Douglas Baird, Eduardo Peñalver, Gerard Magliocca, Glynn Lunney, Greg Reilly, Jennifer Berk, Jonathan Masur, Kate Darling, Lior Strahilevitz, Lisa Larrimore Ouellette, Randal Picker, Sarah Wasserman Rajec, Saul Levmore, Victoria Schwartz, and Vincent Buccola, and to participants at the University of Chicago Law School Research Colloquium, the 2013 Works in Progress in Intellectual Property conference at Seton Hall University School of Law, and the 2013 Patent Conference at the IIT Chicago-Kent College of Law.
INTRODUCTION .............................................................................................................. 3

I. PATENT LITIGATION AND INVALID PATENTS ........................................... 6
   A. Patent Defenses .............................................................................................. 6
      1. Invalidity ................................................................................................. 8
      2. Noninfringement .................................................................................. 11
      3. The (Partial) Independence of Invalidity and Noninfringement .......... 14
   B. Invalid Patents ............................................................................................ 16
      1. The Problem ......................................................................................... 17
      2. Three Solutions .................................................................................... 20

II. THE DEFENDANT’S CHOICE: INVALIDITY VERSUS NONINFRINGEMENT ........... 23
   A. Tradeoffs .................................................................................................. 23
      1. Claim-Construction Strategy ................................................................. 24
      2. Trial Narratives .................................................................................... 27
      3. Resource Constraints .......................................................................... 30
   B. Asymmetries .............................................................................................. 31
      1. Invalidity’s Elevated Burden of Proof .................................................... 32
      2. Information and Timing Advantages ..................................................... 33
      3. Asymmetric Outcomes ......................................................................... 38
   C. What About the Merits? ........................................................................... 40

III. THREE REFORMS ............................................................................................. 43
   A. Eliminate the Elevated Burden of Proof for Invalidity ............................ 43
   B. Permit Defendants to Bifurcate Patent Litigation .................................... 44
   C. Allow a Successful Patent Challenger to Recover from Other Beneficiaries of an Invalidity Judgment .......................................................... 47

CONCLUSION ............................................................................................................. 50
INTRODUCTION

Archers hate vibrations. An archer releasing a bowstring and firing an arrow unleashes a great deal of force in a short time, much of which sends the arrow toward the target. Some of the force, however, is transferred to the bow, causing it to vibrate and throwing off the accuracy of the shot. To address this problem, inventors have created numerous damping systems to absorb and dissipate these vibrations. One such system was created by Steven Sims, who received for his invention a patent titled “Archery Bow Accessories with Bow Vibration Decay Pattern Modifiers for Improving Accuracy.” The patent describes a “mushroom-like” attachment that screws into each end of the bow, made out of a soft polymer that can “wiggle and jiggle” when the bow is fired, absorbing and dissipating the vibrations.

Armed with this patent, Sims and his company, Sims Vibration Laboratory, Inc., targeted Bow Jax, Inc., asserting that its products infringed the ‘842 patent and two others. Bow Jax makes damping systems similar to the one described in the ‘842 patent, including one that attaches to each end of a bow and dissipates vibrations by wiggling. The Bow Jax system is not identical to the one described in the ‘842 patent, though; instead of the mushroom shape described and claimed in the patent, Bow Jax’s dampers have an X shape, with four arms that vibrate when an arrow is fired.

Faced with the threat of patent litigation, Bow Jax, like most accused infringers, had two major defenses on which it could have relied: invalidity and noninfringement. An invalidity defense asserts that the patent holder did not
satisfy the basic requirements to obtain a patent, usually because the claimed invention was not novel or would have been obvious when it was invented.\(^6\) A noninfringement defense, on the other hand, asserts that the defendant’s accused product or method does not fall within the scope of the invention claimed in the patent.\(^7\) Both defenses were plausible for Bow Jax: damping systems have been used by archers for decades,\(^8\) and Bow Jax’s damping devices are significantly different from those described in the Sims patents. So with two potentially winning arguments, Bow Jax did what many accused patent infringers do: it chose to rely on noninfringement. It filed a declaratory-judgment claim against Sims and his company, seeking a declaration that its products did not infringe the ‘842 patent.\(^9\)

This is all too common a story in patent law: a patent is granted on an invention that is not really new; the patent holder goes after a potential infringer; and the accused infringer relies on noninfringement arguments instead of invalidity arguments.\(^10\) This is troublesome because invalid patents are arguably the single biggest problem in modern patent law.\(^11\) It is not unusual for a new technology product to be covered, or arguably covered, by thousands of distinct patents, owned by hundreds of different patent holders, many likely to be invalid. It is essentially impossible to analyze all these patents, so patent holders often ignore them until sued, and then settle, whether or not the patents

(footnote continued from previous page)

\(^6\) See 35 U.S.C. §§ 282 (defenses to patent infringement) & 102 (invalidity).

\(^7\) See id. §§ 271 (definition of patent infringement) & 281 (remedy for patent infringement); Seal-Flex, Inc. v. Athletic Track & Court Constr., 172 F.3d 836, 842 (Fed. Cir. 1999) (infringement).

\(^8\) See infra note 15.

\(^9\) First Amended Complaint, supra note 3.


are valid.\textsuperscript{12} This imposes a substantial tax on innovation, raising its cost and reducing the output of new products and services.\textsuperscript{13}

The Bow Jax story had a happy ending for those worried about invalid patents: though the lawsuit between Sims and Bow Jax was settled,\textsuperscript{14} the Patent and Trademark Office also initiated a reexamination of the ’842 patent and eventually rejected the relevant claims as obvious.\textsuperscript{15} But administrative processes to invalidate patents can only do so much; courts also play a significant role in invalidating bad patents. And if defendants avoid arguing invalidity, then courts will invalidate fewer bad patents than is optimal.

Ideally, defendants would analyze invalidity and noninfringement separately, relying on each defense if and when it is strong enough to be cost-efficient. District courts would then be able to invalidate bad patents, creating benefits for innovators and consumers. And the legal literature typically treats them this way, analyzing the two defenses separately. Yet invalidity and noninfringement are interrelated, with tradeoffs and asymmetries that have a substantial effect on parties’ strategic behavior and on the structure of patent doctrine more broadly.

This article provides the first detailed examination of the interaction between invalidity and noninfringement. It argues that these tradeoffs and asymmetries result in too few patents being found invalid, because defendants are often better off pursuing noninfringement defenses, while society is often better off if they pursue invalidity defenses. It then concludes with a proposal for reforms that would help create balance between invalidity and noninfringement.

After a brief background in Part I, Part II discusses the tradeoffs and asymmetries between invalidity and noninfringement. Although in many cases


\textsuperscript{15} Ex parte Sims Vibration Laboratory, Inc., No. 2013-001458, 2013 Pat. App. LEXIS 784 (Feb. 19, 2013). The examiner found the invention obvious in view of two broad categories of prior art: different kinds of damping systems that were attached to archery bows, see, e.g., U.S. Pat. No. 4,936,283 (filed Mar. 17, 1989); U.S. Pat. No. 3,412,725 (filed Mar. 29, 1965); and similar damping systems used in other contexts, including other kinds of sporting goods, see, e.g., U.S. Pat. No. 5,362,046 (filed May 17, 1993); U.S. Pat. No. 5,314,180 (filed Dec. 16, 1992); U.S. Pat. No. 3,941,380 (filed July 12, 1973).
only one of the two defenses will have plausible merit, in a surprising number of cases both defenses are plausible, at least at the beginning of the case. In these cases, tradeoffs lead most defendants to focus their efforts on just one of the defenses, and asymmetries bias that choice in favor of noninfringement. Tradeoffs between invalidity and noninfringement include the strategic choice between broad and narrow claim constructions, the need for a single coherent narrative for trial, and limited litigation resources. Asymmetries between invalidity and noninfringement include the elevated burden of proof that applies to invalidity defenses, information and timing advantages that make it comparatively easier for a defendant to argue noninfringement, and the difference in outcomes between a successful noninfringement judgment and a successful invalidity judgment.

Part III proposes reforms. First, these tradeoffs and asymmetries provide another reason to eliminate the elevated burden of proof for invalidity defenses. Second, giving defendants the option to bifurcate cases and defer consideration of infringement issues until after validity is determined would eliminate the conflict between trial narratives, would give litigants enough time to focus on the invalidity case, and would eliminate the need in many cases to ever litigate infringement. And third, a new cause of action would allow a successful patent challenger to collect, from competitors who benefit from that judgment, a portion of those benefits. This cause of action, akin to an action for an accounting, would better align the incentives of a patent defendant with the public interest in its defense.

I. PATENT LITIGATION AND INVALID PATENTS

This Part provides background, discussing the roles of invalidity and noninfringement defenses in patent litigation and the problem of invalid patents.

A. Patent Defenses

A patent grants to an inventor the exclusive rights, for a limited time, to make, use, sell, offer to sell, and import his or her invention. A patent represents a bargain between an inventor and society: in return for inventing something new and disclosing it to the world, the inventor gets the reward of a temporary monopoly over that invention. An inventor can only obtain that

---

16 35 U.S.C. § 271(a); see also U.S. CONST. art. I, § 8, cl. 8 (granting Congress the power “[t]o 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”).

17 This is the traditional economic account for patent law: that it provides ex ante incentives both to create inventions and to disclose those inventions to the world. The classic statement of (footnote continued on next page)
monopoly, however, if he or she fulfills the patent bargain: the invention must be novel and nonobvious, and it must be fully described, enabled, and claimed in the patent document, so that others can make and use the invention once the patent has expired.

A patent holder who believes that someone else is making, using, selling, offering for sale, or importing a product or process that incorporates that invention can sue for patent infringement. The defendant accused of patent infringement then has two principal defenses, invalidity and noninfringement, that correspond to the two sides of the patent bargain. An invalidity defense focuses on what the inventor contributed to obtain the patent: whether he or she invented something novel and nonobvious and disclosed it to the world in compliance with the patent laws. A noninfringement defense, on the other hand, focuses on the monopoly the inventor gets in exchange for that invention and disclosure: whether the defendant’s product or process falls within that monopoly.

(footnote continued from previous page)

The principal alternative view is provided by prospect theory, which asserts an ex post justification for patents: that by granting a monopoly over a new technology after the technology has been invented, patents encourage investment in developing and commercializing the new invention, while avoiding wasteful duplication of effort. See John F. Duffy, Rethinking the Prospect Theory of Patents, 71 U. Chi. L. Rev. 439 (2004); Edmund W. Kitch, The Nature and Function of the Patent System, 20 J. L. & Econ. 265 (1977). See also Mark A. Lemley, Ex Ante versus Ex Post Justifications for Intellectual Property, 71 U. Chi. L. Rev. 129 (2004).

18 35 U.S.C. § 281; see also Seal-Flex, Inc. v. Athletic Track & Court Constr., 172 F.3d 836, 842 (Fed. Cir. 1999).
1. Invalidity

An invalidity defense asserts that even though the Patent and Trademark Office granted a patent, that patent is invalid because the inventor failed to comply with the basic requirements for patentability. The focus of any invalidity decision, then, is on the state of the world when the patent was granted, rather than the details of the defendant’s accused product or process. There are several invalidity doctrines, but they can be placed into three broad categories: doctrines that ensure an inventor has created something meaningfully new; doctrines that ensure an invention is fully disclosed to the public; and doctrines that govern the types of inventions to which the law extends patent rights.

First, doctrines that ensure an inventor has created something meaningfully new include the novelty and nonobviousness requirements, which are the core doctrines at issue in most invalidity challenges. Their purpose is to help ensure that an inventor has meaningfully contributed to society before he or she is rewarded with a monopoly. Novelty doctrine generally requires that the claimed invention not have been known, used, or described by others before the patent was granted. Their purpose is to help ensure that an inventor has meaningfully contributed to society before he or she is rewarded with a monopoly. Novelty doctrine generally requires that the claimed invention not have been known, used, or described by others before the patent was granted.

---


20 This taxonomy borrows from that given in Andres Sawicki, Better Mistakes in Patent Law, 39 FLA. ST. U. L. REV. 735, 742–44 (2012). Sawicki breaks patentability doctrines into four categories: scope, covering patentable subject matter and utility; the invention itself, covering nonobviousness; disclosure, covering enablement and best mode; and the definiteness doctrine. Id. I include the written-description requirement and group it with the disclosure and definiteness doctrines, all of which concern different aspects of the patentee’s disclosure to the world.

21 See Allison & Lemley, supra note 10, at 208 (finding that 138 out of a sample of 191 patents found invalid in patent litigation were invalidated on the basis of 35 U.S.C. § 102 (novelty) or § 103 (obviousness)).

22 The requirement that an inventor contribute something new to society has been a fixture since the early days of patent law. See, e.g., 1 WILLIAM C. ROBINSON, THE LAW OF PATENTS FOR USEFUL INVENTIONS § 221 (1890) (“The consideration for the grant of [an inventor’s] exclusive privilege is the benefit which he confers upon the public….”); 2 WILLIAM BLACKSTONE, COMMENTARIES *407 n.15 (“The grant of a patent … is in the nature of a purchase for the public, to whom the patentee is bound to communicate a free participation in the benefit of his invention, at the expiration of the time limited”). Indeed, the requirement may be constitutional: the Patent Clause of Article III authorizes Congress to extend exclusive rights only to “Authors and Inventors,” U.S. CONST. art. I, § 8, cl. 8, in contrast with the British practice of granting patents to individuals favored by the Crown, inventors or not. See Graham v. John Deere Co., 383 U.S. 1, 5–6 (1966); but see Tun-Jen Chiang, First-to-File as a Rule of Evidence, 30 YALE J. REG. ONLINE 11 (2012) (arguing that even if the Constitution limits patents to inventors, a rule granting a patent to the first filer is a constitutional rule of evidence).
fore the patent applicant came up with the claimed invention.\textsuperscript{23} Nonobviousness doctrine adds to this by requiring that an invention not have been obvious to a person having ordinary skill in the art as of the time of invention.\textsuperscript{24} These doctrines help ensure that the patent bargain is a good one for society: because society generally gains little or nothing from the independent invention of technologies that already exist, or that are obvious even if they have not been assembled in the precise form contemplated by the inventor, no reward is necessary to encourage their creation.\textsuperscript{25} Moreover, such monopolies would be counterproductive to the prospect nature of patent rights, since independent inventors, each with his or her own patent rights, would lead to just the sort of wasteful duplication of effort in developing and commercializing an invention that the patent system seeks to avoid.

Second, doctrines that ensure an invention is fully disclosed to the public include the written-description, enablement, best-mode, and definiteness requirements. These doctrines all require clear disclosures of specific aspects of the claimed invention, though for somewhat different purposes. The written-description and definiteness requirements are relevant during the term of the patent: they require the disclosure of information that helps readers know the boundaries of the patentee’s exclusive rights, so they can more easily predict whether their activities would infringe the patent.\textsuperscript{26} The enablement and best-

\textsuperscript{23} See 35 U.S.C. § 102. This description necessarily simplifies a complex set of statutory rules; the statutory limits of novelty are both broader and narrower than this description implies. For example, the United States has traditionally used a first-to-invent rule, awarding patent rights to the first inventor to create an invention, even if he or she filed a patent application later than another inventor. This rule was not absolute, however: under the rule, various kinds of prior art counted or did not count to show earlier invention, depending on what kind of prior art they were, what country they were from, when they were from, and so forth. See 35 U.S.C. § 102 (2006). Under the Leahy-Smith America Invents Act, Pub. L. 112-29, 125 Stat. 284 (2011), which will take effect for patent applications filed on or after March 16, 2013, the United States will switch to a first-to-file rule, under which the first inventor to file a patent application will be entitled to a patent unless the invention was previously patented, described, or used in public, with various exceptions. See id. § 3(b)(1), (n), 125 Stat. at 293.


\textsuperscript{25} The current rule is not the only possible balance; patent law could drive a harder bargain. See, e.g., Michael Abramowicz & John F. Duffy, \textit{The Inducement Standard of Patentability}, 120 \textit{Yale L.J.} 1590 (2011) (proposing that the nonobviousness requirement be construed to limit patents to “those inventions which would not be disclosed or devised but for the inducement of a patent” (quoting Graham v. John Deere Co., 383 U.S. 1, 11 (1966))).

\textsuperscript{26} 35 U.S.C. § 112(a) (written-description requirement) & (b) (definiteness requirement); Ariad Pharmaceuticals, Inc. v. Eli Lilly & Co., 598 F.3d 1336, 1346 (Fed. Cir. 2010) (en banc) (observing that one purpose of the written-description requirement is “to inform the public during
mode requirements, on the other hand, are directed to conduct after a patent has expired: they help ensure that others can make use of the invention after the limited patent monopoly has ended, and that others can improve and build upon the invention without “undue experimentation.” They do so by requiring that a patent provide enough information to enable someone who is skilled in the art of the invention to make and use the invention, while also disclosing the “best mode” contemplated by the inventor for practicing the invention. Together, these disclosure doctrines help ensure that the patentee has fulfilled his side of the patent bargain by making available whatever new thing he or she has invented, in a form so others can make use of the knowledge.

Third, doctrines that govern the types of inventions to which the law extends patent rights include the utility and patentable-subject-matter requirements. These doctrines prevent an inventor from obtaining a patent on an invention that is overly conceptual or abstract, because it would preempt too much subsequent innovation. Thus, under the patentable-subject matter doctrine, an inventor may not patent “laws of nature, physical phenomena, 

(footnote continued from previous page)
the life of the patent of the limits of the monopoly asserted, so that it may be known which features may be safely used or manufactured without a license and which may not”) (quoting Schriber-Schroth Co. v. Cleveland Trust Co., 305 U.S. 47, 57 (1938)); Athletic Alternatives, Inc. v. Prince Mfg., Inc., 73 F.3d 1573, (Fed. Cir. 1996) (observing that the primary purpose of the definiteness requirement is “to guard against unreasonable advantages to the patentee and disadvantages to others arising from uncertainty as to their [respective] rights” (alteration in original)) (quoting General Electric Co. v. Wabash Appliance Corp., 304 U.S. 364, 369 (1938)); Michael Risch, A Brief Defense of the Written Description Requirement, 119 YALE L.J. ONLINE 127 (2010), http://yalelawjournal.org/2010/03/09/risch.html. To this list of doctrines, one could add the basic requirement that a patentee write claims setting forth his or her invention. See 35 U.S.C. 112(b)–(f).


Likewise, under the usefulness doctrine, an inventor may not patent an invention for which no practical use is yet known, so that it can only be used for further research. These doctrines also have a secondary purpose of reinforcing the requirement that an inventor create something new, since laws of nature, physical phenomena, and abstract ideas are more likely than other inventions to have been discovered, created, or used in similar or different contexts.

2. Noninfringement

A noninfringement defense argues that even if a patent is valid, it does not cover what the accused infringer makes or does. This is an argument that the accused product or process does not fall within the monopoly granted to the inventor as the other half of the patent bargain. The scope of the patent monopoly is determined by the patent’s claims, which must "particularly point[] out and distinctly claim[] the subject matter which the inventor or a joint inventor regards as the invention." Under the modern law of patent infringement, claims define the scope of the invention, and thus the limits of the inventor’s monopoly rights. This was not always so: until 1836, patents did not

---

31 Chakrabarty, 447 U.S. at 309.
33 See id. § 282(b)(1) (noninfringement defense); Seal-Flex, 172 F.3d at 842.
34 35 U.S.C. § 112(b).
even have to have claims, and the claim did not become the core language defining an invention until the late 1800s.36

A patent claim generally contains two pieces: a preamble setting forth a general description of the invention and a listing of limitations—“elements, steps, and/or relationships”—that describe the invention.37 Determining if an accused product or process infringes a claim requires two steps: first, the language of the claim must be construed, defining the legal scope of the claim; and second, the accused product or process must be compared to the claim to see whether each limitation of the patent claim, as construed by the court, is satisfied by the accused product or process.38 The accused product or process must satisfy every limitation of the claimed invention, either literally or under the doctrine of equivalents, to fall within the scope of the patent claim; if any limitation is not satisfied, there is no infringement.39

It is worth briefly illustrating how this works with an example. One of the patents at issue in the ongoing Apple/Samsung smartphone patent litigation is U.S. Patent No. 7,469,381, which is titled “List scrolling and document translation, scaling, and rotation on a touch-screen display.”40 As Apple explained in a court filing, the patent relates to a “method for displaying images on touch screens: when one uses a finger to drag a displayed page past its bottom edge, for example, and releases the finger, the page bounces back to fill the full screen.”41 Claim 1 of the ’381 patent, in the typically impenetrable language of patents, claims the following invention:

1. A computer-implemented method, comprising:
   at a device with a touch screen display:
   displaying a first portion of an electronic document;
   detecting a movement of an object on or near the touch screen display;
   in response to detecting the movement, translating the electronic document displayed on the touch screen display in a first direction to display a se-

36 ROBERT PATRICK MERGES & JOHN FITZGERALD DUFFY, PATENT LAW AND POLICY: CASES AND MATERIALS 798 (5th ed. 2011); see also Merrill, 94 U.S. at 570.
37 37 C.F.R. § 1.75(e); MPEP § 608.01(i)(e).
38 Seal-Flex, 172 F.3d at 842.
39 This is sometimes called the “All Limitations Rule.” E.g., Johnston v. IVAC Corp., 885 F.2d 1574, 1577 n.3 (Fed. Cir. 1989); ROBERT L. HARMON, CYNTHIA A. HOMAN & CHARLES M. MCMAHON, PATENTS AND THE FEDERAL CIRCUIT 467–68 (10th ed. 2011).
cond portion of the electronic document, wherein the second portion is different from the first portion;

in response to an edge of the electronic document being reached while translating the electronic document in the first direction while the object is still detected on or near the touch screen display:

displaying an area beyond the edge of the document, and

displaying a third portion of the electronic document, wherein the third portion is smaller than the first portion; and

in response to detecting that the object is no longer on or near the touch screen display, translating the electronic document in a second direction until the area beyond the edge of the electronic document is no longer displayed to display a fourth portion of the electronic document, wherein the fourth portion is different from the first portion.\(^\text{42}\)

In its effort to prove that one accused product, the Samsung Infuse 4G smartphone, infringed this claim, Apple used a claim chart to compare each of the claim’s limitations to the way the accused phone worked. One line of the chart, for example, showed:\(^\text{43}\)

<table>
<thead>
<tr>
<th>Claim 1</th>
<th>Samsung Infuse 4G</th>
</tr>
</thead>
<tbody>
<tr>
<td>in response to an edge of the electronic document being reached while translating the electronic document in the first direction while the object is still detected on or near the touch screen display: displaying an area beyond the edge of the document, and displaying a third portion of the electronic document, wherein the third portion is smaller than the first portion; and</td>
<td>In response to reaching the edge of an electronic document while it is moving in the first direction and the finger is still on the touch screen, the Infuse 4G displays a black region beyond the edge, and thus displays a smaller “third portion” of the document. (See Balakrishnan Decl. (\S) 42-46, Ex. 13a.)</td>
</tr>
</tbody>
</table>

\(^{42}\) '381 Patent cl. 1.

In response, Samsung argued that the accused smartphone failed to meet this limitation because the “area beyond the edge of the document”—the black area outlined in green in Apple’s photograph—was not “displayed,” since no light was emitted by the smartphone’s AMOLED screen when it showed solid black.\(^44\) (They also made another argument related to the requirement that the document be translated in a “first direction.”\(^45\))

The dispute over the “displaying” requirement, then, focused on the first step of the infringement analysis, claim construction. Apple asserted that the accused smartphone “display[ed] an area beyond the edge of the document” by showing a black region when the photograph (the “document”) was pulled partially off-screen by the user. Samsung responded that it did not “display” such an area, because the smartphone did not emit light when it displayed black, and the term “display” requires emitting light. And the court sided with Apple, concluding that “displaying” meant “showing or revealing to the viewer,” since nothing in the patent required the device to emit light as it did so.\(^46\) It thus found that Apple was likely to be able to prove infringement of the ‘381 patent.\(^47\)

This pattern is typical for many, if not most, infringement disputes: with little or no dispute about how the accused product or process works, the entire ball game comes down to what exactly the patent claim requires. If the patent is construed broadly enough to encompass the accused product or process, the court is likely to find infringement. If, instead, the patent is construed more narrowly, the court is unlikely to find infringement.

3. The (Partial) Independence of Invalidity and Noninfringement

At the highest level, invalidity and noninfringement are independent defenses focused on different issues. Invalidity is focused on the patent and the universe that existed when the inventor applied for the patent, without regard to the accused product or process. Noninfringement looks to the patent as it exists now and the accused product or process. The two defenses are thus formally independent: nothing prevents a defendant from arguing that the plaintiff’s asserted patent is invalid, and that regardless its accused product or


\(^{45}\) Id.


\(^{47}\) Id. at *106–08.
process does not infringe the patent. A patent issued today on the broad category of flying vehicles would be invalid, because prior art like the Wright brothers’ airplane, helicopters, and jetpacks would render its claims non-novel, but a carmaker sued on that patent would also have an airtight noninfringement defense.

That independence is limited, however, because both defenses depend on the scope of the patent claims, and thus on how the claims are construed. The basic question of a noninfringement defense—whether the accused product or process satisfies every limitation of the patent claim—depends on what those claims mean. Likewise, each of the three basic invalidity arguments depends on how a claim is construed, albeit to varying degrees. First, determining whether a claim was novel and nonobvious requires comparing the construed claims to the prior art. If the claimed invention, as construed by the court, existed in the prior art or was obvious in view of the prior art, then the claim is invalid.\footnote{Indeed, the Supreme Court has expressly compared the processes of determining invalidity and infringement, both of which involve comparing something—the prior art, or the accused product or process—to the patent claims. See Peters v. Active Mfg. Co., 129 U.S. 530, 537 (1889) (quoting Peters v. Active Mfg. Co., 21 F. 319, 321 (C.C.S.D. Ohio 1884)) (“That which infringes, if later, would anticipate, if earlier.”); see also Int’l Seaway Trading Corp. v. Walgreens Corp., 589 F.3d 1233, 1239 (Fed. Cir. 2009) (“[I]t has been well established for over a century that the same test must be used for both infringement and anticipation.”)} But it is impossible to know whether the prior art disclosed any specific claim limitation without knowing what exactly that claim limitation means. Second, determining the scope of the inventor’s required disclosures under the written-description, enablement, best-mode, and definiteness requirements also requires construing the claims. The written-description and enablement requirements, for example, require the inventor to fully describe his or her invention and enable others to make full use of it. If a claim is construed to cover a particular form of the invention, and that form was not described or enabled, then the patent may be invalid. If, however, that claim is construed not to cover that form of the invention, then it doesn’t matter if the patent described or enabled it.\footnote{This necessarily simplifies a complex question about the scope of a patent’s written description and enablement. For a more detailed discussion of the effects of claim construction on the written-description and enablement requirements, see Chiang, supra note 35, at 1131–34. See also Dan L. Burk & Mark A. Lemley, Patent Quantum Mechanics, 9 LEWIS & CLARK L. REV. 29 (2005).} And third, determining whether a claim has utility and falls within the statutory patentable subject matter likewise requires knowing what those claims cover, i.e., how they’ve been construed by a court. In \textit{Bilski v. Kappos},\footnote{130 S. Ct. 3218 (2010).} for instance, the Supreme Court construed two claims in
the petitioner’s patent application to claim the “concept of hedging, described in claim 1 and reduced to a mathematical formula in claim 4.” The Court concluded that these proposed claims were not patentable, because granting them would “[a]llow [petitioners] to patent risk hedging,” a “basic concept” and “abstract idea” too broad to constitute a “process, machine, manufacture, or composition of matter.” If, however, the proposed claims had been construed to require more than the naked “concept of hedging,” then the claims might have been patentable. As Justice Stevens’ concurring opinion put it, “the Court artificially limit[ed] petitioners’ claims to hedging, and then conclude[d] that hedging is an abstract idea rather than a term that describes a category of processes including petitioners’ claims. Why the Court [did] this [was] never made clear.”

It is easy to overstate the degree to which invalidity and noninfringement are related. Invalidity is, at its core, about what came before the invention—the prior art—and what is disclosed in the patent document—the written description and other specification. Noninfringement, in contrast, is most basically about the accused product or process, which has little or nothing to do with the prior art and the patent’s specification. But claim construction plays an important role—sometimes the critical role—in each determination, so the two issues are not wholly independent.

B. Invalid Patents

Patents are granted after a review by the Patent and Trademark Office, which must make sure an applicant “is entitled to a patent under the law” before granting the patent. To make this determination, patent examiners review the patent application, perform their own search for prior art, and determine if the proposed patent claims meet the patentability rules. If a proposed claim fails one of the specific requirements of patentability set forth in 35 U.S.C. §§ 101, 102, or 103, the examiner must reject the application; otherwise the PTO must issue the requested patent.
Despite this review, many invalid patents make it through prosecution, creating significant problems for companies and individuals that want to make and sell new products and services. Though scholars—and much of the public—basically agree that this is a serious problem, there is far less agreement about the best solution.

1. The Problem

Among patent scholars, there is almost unanimous agreement that patent examiners do not do this job particularly well, with the PTO issuing many invalid patents.\(^{58}\) Indeed, nearly half of litigated patents are invalidated by the courts.\(^{59}\) And this statistic may understate the number of invalid patents in force, since very few patents are litigated and since patent holders have an opportunity and an incentive to choose those patents and claims that are more likely to survive judicial review.\(^{60}\)

There are different explanations for the PTO’s less-than-effective screening system. First, patent examiners have little incentive to reject bad patent applications, and every incentive to grant more patents. Rejecting an application creates more work for an examiner, since he or she must justify the rejection, and since it likely results in another round of dialogue with the applicant. Approving an application, however, gets the file off the examiner’s desk for


\(^{59}\) Allison & Lemley, supra note 10, at 205–07. This does not mean, by the way, that district courts are effectively doing their job of invalidating bad patents, since very few cases make it to a final decision on validity.

\(^{60}\) See generally Lemley, supra note 58, at 1501–03. Conversely, it might be the case that only the close calls make it to a final determination because most patents are valid and defendants choose to license the patents in settlement. See generally George L. Priest & Benjamin Klein, The Selection of Disputes for Litigation, 13 J. LEGAL STUD. 1 (1984). But parties settle cases for many reasons, including reducing risk of an adverse outcome, so it’s hard to know how much we can infer from the statistics.
good. And the PTO has at times encouraged this approach, famously asserting that its “primary mission” is “to help customers get patents.” Second, even if they wanted to, patent examiners may just be too overwhelmed to catch bad patents. In recent years the PTO has received more than 500,000 patent applications a year and granted almost 250,000 patents per year. A patent examiner spends just eighteen hours on the average patent application—spread out over a year or two—in which he or she must review the application; conduct a prior-art search; review both the prior art submitted by the applicant and that uncovered in the search; compare the prior art to the claimed invention; go through multiple rounds of office actions (written statements finding claims patentable or not, and if not, explaining why not) and responses to office actions; perhaps conduct an interview with the applicant; and ensure the application complies with various formalities. And third, patent examiners often lack the information needed to determine whether a patent application should be approved. Though patent examiners do specialize in a particular technology area, they are not experts; most have a bachelor’s or master’s degree and less than three years experience examining patent applica-

---

61 Lemley, supra note 58, at 1496 n.3; see also Corinne Langinier & Philippe Marcoul, Monetary and Implicit Incentives of Patent Examiners (U. Alberta Dep’t of Econ. Working Paper No. 2009-22), available at http://www.economics.ualberta.ca/~/media/economics/FacultyAndStaff/WPs/WP2009-22_Langinier.pdf (concluding that rewarding patent examiners for rejecting applications, rather than for allowing patents, would give examiners greater incentives to search for information). There is evidence that examiners act more quickly, and grant more patents, when it is economically advantageous to the PTO for them to do so. See Michael Frakes & Melissa F. Wasserman, Does Agency Funding Affect Decisionmaking?: An Empirical Assessment of the PTO’s Granting Patterns (unpublished working paper), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1986542.


64 Lemley, supra note 58, at 1496 n.3.
tions. Nor can examiners benefit from the perspectives provided by an adversarial process, since patent prosecution is conducted ex parte.

The resulting invalid patents have pernicious effects on individuals and companies that create new products and services, effects that are exacerbated by the sheer scale of the modern patent system. It is no longer the case, if it ever was, that a single invention is likely to be covered by a single patent; today it is far more likely that a new product is covered by thousands of patents. Such patent thickets impose a significant tax on innovation, since a company must, at least in theory, obtain licenses to all those thousands of patents before it introduces a new product. Invalid patents multiply the size of this tax, since

65 See Tamara Dillon, Patent work: The other side of invention, OCCUPATIONAL OUTLOOK Q., Fall 2009, at 18; Dennis Crouch, Patent Examiner Experience Levels, PATENTLYO (Feb. 5, 2010), http://www.patentlyo.com/patent/2010/02/patent-examiner-experience-levels.html. Turnover is a significant problem at the agency, with one examiner leaving for every two hired even as the office tries to expand significantly. See U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-07-1102, U.S. PATENT AND TRADEMARK OFFICE: HIRING EFFORTS ARE NOT SUFFICIENT TO REDUCE THE PATENT APPLICATION BACKLOG (2007).


67 About 9 million U.S. patents have been issued since the launch of the patent system in the 1790s, of which more than 2 million are still in force. See USPTO PATENT FULL-TEXT AND IMAGE DATABASE, http://patft.uspto.gov/netahm/PTO/search-adv.htm (searched Feb. 12, 2013) (finding 9,095,989 patents issued since 1790); Dennis Crouch, How Many US Patents are In-Force?, PATENTLYO (May 4, 2012), http://www.patentlyo.com/patent/2012/05/how-many-us-patents-are-in-force.html.

68 This is especially the case in the information-technology industry, where innovation is an iterative process by which new features and technologies build incrementally on each other. In contrast, the pharmaceutical industry may be a notable exception, because pharmaceutical patents can precisely claim a particular molecule, because pharmaceutical-patent claims that are broad and vague enough to apply to many innovations are rare, and because the pharmaceutical industry relies more than many industries on concentrated high-fixed-cost investment that is a better match for patent protection. See generally BURK & LEMLEY, supra note 11, at 37–65 (2009) (describing the industry-specific nature of the patent system); see also Alexander Tabarrok, Patent Theory versus Patent Law, 1 CONTRIB. ECON. ANAL. & POL’Y 9 (2002).

the holder of an invalid patent will want a royalty just as much as the holder of a valid patent will. And potential infringers are risk-averse, choosing to license even invalid patents and to settle even dubious patent lawsuits. Potential infringers do so both because of the sheer scale of the patent-thicket problem—sorting thousands of potentially relevant patents into groups of valid and invalid patents would be a monumental undertaking—and because patent litigation is uncertain enough that paying the tax is simpler and cheaper than fighting. Fighting is a risk because any patent holder can block, or attempt to block, the accused infringer’s activity; although injunctions are no longer almost automatic when a patent holder prevails in court, they are still a significant risk, especially when multiplied across thousands of patents.

2. Three Solutions

Scholars and patent practitioners have proposed three types of solutions to the problem of invalid patents.

First, better review by the PTO could lead to fewer bad patents being granted. Expanding the number of patent examiners, the amount of time examiners spend on a patent application, and the amount of funding available to the Patent and Trademark Office could all improve the review process and reduce the number of invalid patents granted. This, in turn, would reduce the number of invalid patents that are licensed and settled.

One surprisingly common response to patent thickets is simply to ignore them. See Lemley, supra note 12.

As in most types of litigation, nearly all patent lawsuits are settled—not counting the many patent disputes that are settled before litigation. See Kesan & Ball, supra note 10, at 271–74.


In the three years after eBay was decided, district courts granted permanent injunctions in 48 out of 67 cases in which successful patent plaintiffs sought them. Ernest Grumbles III, Rachel C. Hughey & Susan Perera, The Three Year Anniversary of eBay v. MercExchange: A Statistical Analysis of Permanent Injunctions, INT’L PROP. TODAY, Nov. 2009, at 25.

aminers can spend on each patent application,\textsuperscript{75} or the ability of examiners to track down prior-art references, could make it easier for examiners to figure out when a patent application claims something that is truly new, and thus to distinguish between valid and invalid patents. Requiring patent applicants to provide examiners with more-detailed information, such as the results of a comprehensive prior-art search, would also give examiners more information to work with. Likewise, finding ways to reduce examiner turnover could lead to examiners who are more experienced and knowledgeable and thus better able to distinguish between valid and invalid patents.\textsuperscript{76} But all these measures would be expensive, and it is unclear they would be worth the expense, since most patents are never enforced.\textsuperscript{77}

Second, various forms of administrative review by the PTO, after a patent has been granted, could invalidate bad patents.\textsuperscript{78} This approach provides a form of adversarial argument, with members of the public—usually potential infringers—playing a role in initiating, and sometimes litigating, the administrative review. It also shifts some of the burden of identifying bad patents to the public, who can provide information to the PTO and challenge a patent’s validity without undergoing the greater burden of litigating the dispute. It also allows the PTO to concentrate its resources on the patents that are important enough to merit further review. These advantages have led many scholars and

\textsuperscript{75} The Leahy-Smith America Invents Act, which was signed into law in 2011, makes one kind of information more easily available to examiners: it permits third parties to submit prior-art information relevant to any application, before the examiner issues a notice of allowance. \textit{See} 35 U.S.C. § 122(e).

\textsuperscript{76} The Patent and Trademark Office has itself taken this approach in recent years, hiring scores of new examiners and opening or announcing new satellite offices, staffed with patent examiners, in Detroit, Denver, Dallas-Fort Worth, and Silicon Valley. \textit{See} David Kappos, \textit{Progress Report on Satellite Office Openings, DIRECTOR’S FORUM: A BLOG FROM USPTO’S LEADERSHIP} (Sept. 12, 2012), \textit{http://www.uspto.gov/blog/director/entry/progress_report_on_satellite-office}.


\textsuperscript{78} As modified by the America Invents Act, there are three forms of post-grant administrative review available in the U.S. patent system: \textit{ex parte} reexamination, \textit{see} 35 U.S.C. §§ 302–07; \textit{inter partes} review, \textit{see id.} §§ 311–19; and post-grant review, \textit{see id.} §§ 321–29. \textit{See also} Andrei Iancu & Ben Haber, \textit{Post-Issuance Proceedings in the America Invents Act}, 93 J. PAT. & TRADEMARK OFF. SOC’Y 476 (2012).
policymakers to embrace various forms of post-grant review. But this approach could apply only to the few patents that are important enough to receive extra scrutiny, and not the far-larger category of patents, many invalid, that are never litigated or seriously contested. It would thus do little to combat the patent-thicket problem.

Third, district courts can and do invalidate patents in litigation, either standard patent-infringement litigation or declaratory-judgment claims brought by potential infringement defendants. Like administrative review, litigation concentrates on a small fraction of patents and benefits from an adversarial process, but unlike administrative review it is extraordinarily expensive and time-consuming. And it is imperfectly suited for invalidating bad patents, since much of a typical case is consumed by unrelated infringement issues and since district courts must apply a presumption of validity. District courts also lack expertise and have other structural shortcomings that limit their ability to decide complex technological cases like patent lawsuits. Moreover, potential infringers have imperfect incentives to challenge patents as invalid. These and other factors mean that litigation challenges to patent validity have high error costs. Scholars have nevertheless proposed ways to make litigation a more useful tool in invalidating bad patents.

Which of these three mechanisms is best is, to some degree, beside the point. We need not choose a single mechanism for weeding out bad patents; rather, the goal should be to make each mechanism as effective as it can be without sacrificing other goals. If we can make one mechanism more cost-

---


81 See Farrell & Merges, supra note 58, at 950–60.


84 Cf. Sawicki, supra note 20 (discussing potential mistakes at different points in a patent’s lifetime and analyzing which mistakes are better and worse for society).
effective at preventing an invalid patent from being issued, or eliminating it after the fact, without affecting the other mechanisms or producing other adverse consequences, we should do so. Accordingly, the next Part discusses a substantial impediment to the use of litigation to invalidate patents.

II. THE DEFENDANT’S CHOICE: INVALIDITY VERSUS NONINFRINGEMENT

This Part introduces the tradeoffs and asymmetries between patent invalidity and noninfringement that lead many defendants to choose to focus on just one of the two defenses, and to prefer focusing on noninfringement.

A. Tradeoffs

An accused infringer fighting a patent lawsuit makes many important strategic decisions in the course of litigating the case, but probably the most critical one is what arguments to make in defense. In nearly every patent case, the two leading candidates are invalidity and noninfringement. And as discussed above, invalidity and noninfringement are distinct issues that can be argued simultaneously. Yet several tradeoffs exist between invalidity and noninfringement, so that patent defendants face pressure to focus on one or the other.

For both invalidity and noninfringement, the defendant is usually the party that decides whether the issue will be seriously contested. This is straightforward for invalidity, which is an affirmative defense that must be proved by clear and convincing evidence, and so can only come up as an issue if raised by the defendant. The point is less obvious with noninfringement, since a plaintiff has the burden of proving infringement and must meet that burden to win the case. In practice, though, meeting this burden is relatively simple if the defendant does not put effort into a counterargument; the plaintiff can typically rely on an expert witness, testifying at a relatively high level of generality, that the accused product or process satisfies every limitation of the asserted claims. The issue is only really contested when the defendant argues that the accused product or process is missing certain claim limitations.

---

85 See supra Part I.A.3.
86 35 U.S.C. § 282(a) (“A patent shall be presumed valid. ... The burden of establishing invalidity of a patent or any claim thereof shall rest on the party asserting such invalidity.”); Microsoft Corp. v. i4i Ltd. Partnership, 131 S. Ct. 2238 (2011) (holding that invalidity must be proved by clear and convincing evidence).
87 See 35 U.S.C. § 281 (“A patentee shall have remedy by civil action for infringement of his patent.”).
That the defendant chooses what issues to contest is a rather banal point, but it has profound consequences in patent litigation, because defendants usually choose to focus on one or the other of the two defenses. Several tradeoffs penalize patent defendants who try to keep both defenses alive, even if they both could have merit in any particular case. These tradeoffs include tradeoffs in claim-construction strategy, the need for a coherent trial narrative, and resource constraints.

1. Claim-Construction Strategy

The biggest factor pushing patent litigants toward focusing on one defense is the need for a claim-construction strategy that supports that defense. One of the important strategic decisions a patent litigant makes is what claim constructions to propose and argue for. It is often difficult or impossible to construe claims in a way that is compatible with both invalidity and noninfringement arguments, so the choice of claim constructions can act as a fork in the road, forcing defendants to choose one or the other. And because claim construction is such a malleable and uncertain process, and unusually dependent on facts for a question of law, the parties’ arguments have a disproportionate impact on the court’s ultimate claim constructions.\(^\text{88}\)

In many patent cases—perhaps most patent cases—the defendant must choose between arguing for broad claim constructions that support its invalidity arguments and arguing for narrow claim constructions that support its noninfringement arguments.\(^\text{89}\) Narrow claim constructions, in general, are better for noninfringement arguments, since fewer products and processes will fall into narrow claims than into broad ones. Narrow claim constructions make it harder to prove invalidity, though, because fewer prior-art references will describe elements of a narrower claim. Similarly, broad claim construc-


\(^{89}\) In the American patent system, the meaning of a claim’s language is a question of law resolved by the court. Markman v. Westview Instruments, Inc., 517 U.S. 370 (1996). Generally, partway through a case, the parties will meet and confer to discuss which claim terms are in need of construction and will propose constructions. For terms for which the parties cannot agree on constructions, the parties will submit proposed claim constructions and supporting briefs to the court. The court will hold a hearing, called a Markman hearing; consider the parties’ arguments; and construe the claims. See, e.g., N.D. Cal. Loc. Pat. R. 4-1 to 4-7 (revised Dec. 1, 2009); E.D. Tex. Loc. Pat. R. 4-1 to 4-6 (revised July 16, 2012); INTELLECTUAL PROP. OWNERS ASS’N, IPO MODEL LOCAL PATENT RULES §§ 4-1 to 4-7 (2009); Phillips v. AWH Corp., 415 F.3d 1303, 1311–19 (Fed. Cir. 2005) (en banc); id. at 1332–33 (Mayer, J., dissenting). Claim construction is essentially the midterm of a patent case, with expert reports that rely on the construed claims, dispositive motions, and trial usually coming later.
tions make it harder to argue noninfringement, since it’s easier to prove that a product or process falls within a broader category than a narrower category, while broad claim constructions make it easier to argue invalidity, since more prior art can anticipate or render obvious a patent claim.\(^90\)

A classic example of the importance of broader and narrower claim constructions is *Nystrom v. TREX Co.*\(^91\) *Nystrom* concerned a patent covering boards used to build outdoor decks.\(^92\) The patent claimed to solve a problem with standard flat deck boards, which would accumulate water, leading to decay. The patent described a slightly rounded board that, when oriented so the convex side faced up, would shed water to the sides of each board. Figure 2 of the patent showed a cross section of the invention.\(^93\)

---

\(^90\) These tradeoffs between narrow and broad claim constructions have been repeatedly recognized by scholars and practitioners. See, e.g., Tun-Jen Chiang, *The Advantages of Inter Partes Reexamination*, 90 J. PAT. & TRADEMARK OFF. SOC’Y 579, 581 (2008) (“[T]he ideal tactical posture for an accused infringer is to receive a broad claim construction for invalidity analysis but maintain a narrower claim construction for infringement analysis; with patent holders preferring the opposite. Keenly aware of such potential gamesmanship by both sides, the Federal Circuit has repeatedly held that claims must be construed consistently between the two analyses.”); Matthew B. Lowrie, *Critical Issues in Managing Patent Litigation*, 44 IDEA 267, 280 (2004) (“The strategy issues related to what claim construction to ask for also fundamentally impact a patent litigation. The trade-off typically involves whether to ask for a broad construction (infringement is easier, but validity more difficult) or a narrow construction (same, but in reverse).”).

\(^91\) 424 F.3d 1136 (Fed. Cir. 2005), withdrawing and superseding 374 F.3d 1105 (Fed. Cir. 2004).


\(^93\) ’831 Pat., Fig. 2.
Although the specification of the patent described a board milled from solid wood logs and repeatedly discussed wooden boards,94 the asserted claims just used the term “board” without specifying the material from which the board was made.95 Nystrom sued TREX, a company that made composite deck boards made out of a composite mixture of wood fibers and plastic.96

The asserted patent claimed a “board for use in constructing a flooring surface for exterior use,” satisfying various other limitations.97 The parties disagreed about the proper construction of the term “board,” among other claim-construction disputes. The plaintiff proposed a broad construction for “board,” suggesting it meant “an elongated piece of construction material for use in building,” regardless of the material from which the board was made.98 The defendant proposed narrower constructions, under which “board” meant “wood board cut from a log” or “solid all wood board.”99

The defendant’s narrower constructions gave it an airtight noninfringement case: since its boards were made from a wood-plastic composite, not wood cut from a log, they would not satisfy every limitation of a claim that required solid-wood boards.100 The plaintiff would have a better shot under its broader construction, since a board can be “an elongated piece of construction

94 E.g., id. at col. 1, ll. 27–37 (“wood flooring for exterior use”); id. at col. 1, ll. 48–55 (“[V]ery little change has been made in the basic design of wood building materials....”); id. at col. 3, ll. 25–35 (“growth rings” and “bark side”).
96 Nystrom, 424 F.3d at 1140.
97 ’831 Pat., cl. 1. The full language of claim 1, one of several similar asserted claims, reads:

1. A board for use in constructing a flooring surface for exterior use, said board having a top surface, a bottom surface and opposite side edges, said top surface being manufactured to have a slightly rounded or curved configuration from a longitudinal center line thereof downwardly toward each side edge, thereby defining a convex top surface which sheds water and at the same time is comfortable to walk on, and said bottom surface having a concave configuration for nesting engagement with the top surface of another board so that a plurality of the boards may be stacked one on top of the other with the stability of conventional boards having flat top and bottom surfaces.

Id. (emphasis added).
99 Id. at *10 & n.2.
100 Indeed, Nystrom conceded noninfringement, 424 F.3d at 1140–41, after the district court construed “board” to mean “piece of elongated construction material made from wood cut from a log,” 2002 U.S. Dist. LEXIS 27501 at *16.
material for use in building” regardless of whether it is made from wood, metal, stone, or a wood-plastic composite. A greater universe of boards would infringe such a broadly construed claim.

Although Nystrom was decided on infringement, the point applies just as well to invalidity. If a court construes “board” to require a solid-wood board milled from a log, then the universe of relevant prior art is limited to solid-wood boards, or, perhaps, boards made from other materials that would make a solid-wood board obvious. If a “board” can be made out of any building material, then the universe of prior art expands significantly. Yet anything in the prior art that satisfies a claim’s limitations can invalidate a patent; the prior art does not need to anticipate or render obvious every possible board covered by the claim.\(^{101}\) If the prior art contained steel planks or stone slabs that met each of the claim limitations at issue in Nystrom, then the claims would be found invalid, regardless of whether the prior art anticipated solid-wood or composite boards. So anything that expands the universe of prior art makes it easier to prove invalidity.

This strategic dynamic means that parties often go all in on either invalidity or noninfringement, arguing for broad claim constructions that shore up their invalidity arguments or narrow claim constructions that shore up their noninfringement arguments. The alternative can be middle-of-the-road constructions that don’t help either defense enough to be successful. Unless the defendant’s case is unusually strong, then, claim-construction strategy can force the defendant to jettison one defense in favor of the other.

2. Trial Narratives

Even if a defendant pursues both invalidity and noninfringement defenses at the start of a case, the need for a simple, coherent narrative of the case will often lead that defendant to focus on one defense or the other by the time the case reaches trial. Every (good) trial lawyer tries to simplify his or her case and tell jurors a straightforward, appealing story that just happens to support his or her client’s side.\(^{102}\) The need for this simple narrative is only magnified in patent trials, which otherwise stack complexity upon complexity. Many pa-

---


tents and accused products and processes are technically advanced and difficult even for experts to understand—understandably so, since patents by their nature protect new inventions. Invalidity and noninfringement are themselves complex legal concepts that can be difficult for jurors to understand. And invalidity and noninfringement have little in common, making it even harder to keep track of both simultaneously.

Trial lawyers are smart to focus on a simple narrative, even at the cost of a potential defense. Jurors hear competing stories from the plaintiff and defense and must decide which story is more credible. Yet studies of jury psychology show that the credibility a juror assigns to a story is based not on the strength of the evidence, but on the narrative plausibility of a party’s arguments. A juror constructs his or her own internal narrative representation of a case, based on the evidence and his or her background knowledge. The juror then uses this narrative representation, not the raw evidence, to reach a verdict.103 And a patent holder can always present a simple, plausible narrative: “We invented this great new widget, and the defendant used (or stole) our idea.” A defendant can counter with its own simple narrative of invalidity (“The plaintiff didn’t invent anything”) or noninfringement (“Our product is fundamentally different”)—but only if it focuses on one defense. Otherwise, the response is a story that pulls in multiple directions: “Our product is fundamentally different from this thing the plaintiff invented—which, by the way, wasn’t actually new, but had been invented before by this other inventor, or would have been obvious to any idiot in the field.” Such a narrative is less coherent and less intuitively plausible than the plaintiff’s simple narrative.104

Bench trials could help, since lawyers and judges may be better able than jurors to track multiple complex narratives simultaneously, yet most patent cases that go to trial are tried before juries. There are two reasons for this. First, a patent holder has a Seventh Amendment right to a jury trial, though certain equitable issues—like injunctions and the defense of unenforceability


104 Arguably, trial narratives weigh not just in favor of choosing one defense, but in favor of noninfringement. See Roger Shang & Yar Chaikovsky, Inter Partes Reexamination of Patents: An Empirical Evaluation, 15 Tex. Intell. Prop. L.J. 1, 25 (2006) (“Seasoned patent litigators almost always prefer to focus on the noninfringement argument at trial, because, among other reasons, an argument of ‘we don’t infringe’ sounds more moral to a jury than an argument of ‘maybe we infringe but other people invented this first’...”).
due to inequitable conduct—are decided by the judge. Second, the conventional wisdom among patent litigators is that juries are more sympathetic than judges to patent holders. Although some judicial districts are considered more favorable than others to patent holders, plaintiffs can usually choose where to file patent lawsuits. And patent holders have made aggressive use of that power, favoring districts like the Eastern District of Texas that are believed to be friendly to plaintiffs.

---


106 See, e.g., 5 INTELLECTUAL PROPERTY COUNSELING & LITIGATION § 82.02 (Lester Horwitz, Ethan Horwitz & Lisa Hershman, eds. 2012) (“[A] plaintiff with a weak patent should seek a jury, since jurors will give more deference to the beribboned patent than the Judge.”); ABA SECTION OF INTELLECTUAL PROPERTY, PATENT LITIGATION STRATEGIES HANDBOOK § 6.I.X.B (Barry L. Grossman & Gary M. Hoffman eds., 3d ed. 2010); John M. Griem, Jr. & Emily Jayne Kunz, Jury Trials in Patent Cases: Practical and Legal Considerations, PATENT STRATEGY & MANAGEMENT, Jan. 2007.

107 Although the general rule is that as long as the court has personal jurisdiction over the defendant, a plaintiff’s choice of forum is entitled to significant deference, a defendant can seek to transfer the case to another district court “[f]or the convenience of parties and witnesses, in the interest of justice.” 28 U.S.C. § 1404(a). An extensive case law has developed in the Federal Circuit governing transfer motions in patent cases, as numerous defendants, especially in cases pending in the Eastern District of Texas, have petitioned for writs of mandamus directing district judges to transfer cases to districts preferred by the defendants. See, e.g., In re Verizon Bus. Network Servs., 635 F.3d 559 (Fed. Cir. 2011); In re Microsoft Corp., 630 F.3d 1361 (Fed. Cir. 2011); In re Vistaprint Ltd., 628 F.3d 1342 (Fed. Cir. 2010); In re Acer Am. Corp., 626 F.3d 1252 (Fed. Cir. 2010); In re Zimmer Holdings, Inc., 609 F.3d 1378 (Fed. Cir. 2010); In re Nintendo Co., 589 F.3d 1194 (Fed. Cir. 2009); In re Hoffmann-La Roche Inc., 587 F.3d 1333 (Fed. Cir. 2009); In re Genentech, Inc., 566 F.3d 1338 (Fed. Cir. 2009); In re Volkswagen of Am., Inc., 566 F.3d 1349 (Fed. Cir. 2009); In re TS Tech USA Corp., 551 F.3d 1315 (Fed. Cir. 2008). (Full disclosure: I was counsel to Microsoft Corp. in the Acer America Corp. mandamus petition.)

3. Resource Constraints

Resource constraints also create a tradeoff between invalidity and noninfringement, though the importance of the tradeoff depends greatly on the stakes in the case. Patent cases are notoriously expensive to litigate. A 2011 survey of patent lawyers estimated that the median cost of litigating a patent case to final decision was $2.5 million, when between $1 million and $25 million was at stake.\(^{109}\) If more than $25 million was at stake, this estimated cost doubled to $5 million.\(^{110}\) Besides attorneys’ fees, parties can spend heavily on expert fees, expenses from investigating the accused product or process, and searching for prior art.

Very few of these expenses apply to both invalidity and noninfringement arguments, so the marginal cost of pursuing both defenses can be substantial. Consider, for instance, the steps involved in proving a patent invalid. First, generally, the patent and its prosecution history are analyzed—by patent litigators, technical experts, or both—to determine what the relevant patent claims cover and what the critical issues are likely to be. Second, a prior-art search is conducted to build a universe of relevant prior art. Third, the prior art and the relevant patent claims are compared to determine what limitations of the patent are covered by the prior art, what invalidity arguments can be made, and what the weaknesses are in those arguments. Fourth, proposed claim constructions that support the defendant’s invalidity arguments are developed. Fifth, legal arguments, expert opinions, and supporting evidence are developed. Of all these steps, only the initial examination of the patent and prosecution history and the development of proposed claim constructions are likely to be relevant both for invalidity and noninfringement—and even then, those steps are likely to take longer, and involve more debate, if they must support two different defenses instead of just one. The other steps involving prior art are irrelevant for noninfringement arguments, and expert and legal work will rarely overlap.\(^{111}\) And, of course, proving noninfringement has its own unique steps, including developing evidence about the accused product or process, learning how it works, and comparing its elements to the limitations of the asserted patent claims.


\(^{110}\) Id.

\(^{111}\) The best practice is to have separate paid experts for invalidity and noninfringement, to prevent a single expert from taking inconsistent positions when discussing invalidity and noninfringement. Indeed, it is not uncommon for a defendant’s legal team to have dedicated invalidity and noninfringement teams.
If enough is at stake, then the added cost of developing two defenses may be worth it. But with patent litigation becoming more and more expensive, and many frequent defendants facing dozens or hundreds of patent lawsuits, it is not surprising that defendants look to cut costs where they can. This strategy can be risky, because a defendant losing a patent lawsuit may have to pay large damages or even take its product off the market, if the patent holder can obtain a permanent injunction. But this risk is manageable, because most patent cases settle, and because injunctions are more difficult to obtain since the Supreme Court’s decision in eBay Inc. v. MercExchange, L.L.C.

B. Asymmetries

If tradeoffs between noninfringement and invalidity lead most defendants to focus on just one of the two defenses, then the next obvious question is whether one defense or the other has any systematic advantage. This section discusses three asymmetries that lead defendants to prefer focusing on noninfringement: the elevated burden of proof that applies to invalidity defenses; information and timing advantages enjoyed by plaintiffs arguing validity; and

---

112 The American Intellectual Property Law Association has long conducted a biennial economic survey of its members. The 1991 survey found that the median patent litigation, through trial, would cost each side $396,000. AM. INTELLECTUAL PROP. L. ASS’N, REPORT OF THE ECONOMIC SURVEY 1991 at 29 (1991). In 2001, the survey broke out the average by the amount at stake in the case; with less than $1 million at stake, the median reported cost was $499,000, while if more than $25 million was at stake, the median reported cost was $2,992,000. AM. INTELLECTUAL PROP. L. ASS’N, REPORT OF THE ECONOMIC SURVEY 2001 at 84–85 (2001). In 2011, those numbers had grown to $650,000 (less than $1 million at stake) and $5 million (more than $25 million at stake). AM. INTELLECTUAL PROP. L. ASS’N, supra note 109, at I-153 to I-154.

113 This is especially common in the fields of computer hardware and software. According to searches of the Lex Machina patent-litigation database, for example, as of December 30, 2012, Apple Inc. was a party in 121 open patent cases; Microsoft Corp. was a party in 55 open cases; Google Corp. was a party in 86 open cases; various Sony companies were parties in 94 open cases; and various Samsung companies were parties in 104 open cases.


115 Kesan & Ball, supra note 10, at 271–74

116 547 U.S. 388 (2006). Before eBay, a successful patent plaintiff was almost entitled to a permanent injunction; the Federal Circuit applied a “general rule that courts will issue permanent injunctions against patent infringement absent exceptional circumstances.” Id. at 391 (quoting MercExchange, L.L.C. v. eBay, Inc., 401 F.3d 1323, 1339 (Fed. Cir. 2005)). One study found that after eBay was decided, the proportion of winning patent plaintiffs who obtained an injunction dropped from 100% to 60%. Benjamin N. Simler & Scott McClelland, A Model for Predicting Permanent Injunctions After eBay v. MercExchange, BLOOMBERG LAW REPORTS INTELLECTUAL PROPERTY, Apr. 11, 2011, at 11.
the asymmetric outcomes of successful invalidity and noninfringement defenses.

1. Invalidity’s Elevated Burden of Proof

The most basic asymmetry between litigating invalidity and noninfringement is in their burdens of proof: invalidity must be proved by clear and convincing evidence, while infringement must be proved only by a preponderance of the evidence.\(^{117}\) This difference arises from 35 U.S.C. § 282(a), which requires that patents be “presumed valid” and that a party asserting that a patent claim is invalid has the burden of so proving.\(^ {118}\) Accordingly, to prevail on a noninfringement defense, an accused infringer must only show that the patent holder cannot win by a preponderance of the evidence, but to prevail on an invalidity defense it must show that the patent is invalid by clear and convincing evidence, a substantially greater burden.

The standard justification for the elevated burden of proof is that a patent grant represents an administrative decision of the United States. Since such decisions are made only after extensive examination by expert patent examiners, the story goes, they are likely to be correct and are therefore entitled to some deference.\(^ {119}\) This justification has been criticized by scholars, judges, and the Federal Trade Commission, who point out that patent examiners spend little time on each patent, lack sufficient knowledge of the prior art, often have not even seen all the relevant prior art, and do not benefit from an adversarial presentation on the merits of a patent application.\(^ {120}\) The merits of

---


\(^{118}\) 35 U.S.C. § 282(a).

\(^{119}\) See, e.g., Microsoft, 131 S. Ct. at 2249 (noting the “rationale underlying the presumption—that the PTO, in its expertise, has approved the claim” (quoting KSR Int’l Co. v. Teleflex, Inc., 550 U.S. 398, 426 (2007))); Harmon, Homan & McMahon, supra note 118, at § 1.5(a). Other justifications have also been asserted. See, e.g., The Barbed Wire Patent, 143 U.S. 275, 292 (1892) (observing that it was the patent holder “who first published this device; put it upon record; made use of it for a practical purpose; and gave it to the public, by which it was eagerly seized upon, and spread until there is scarcely a cattle-raising district in the world in which it is not extensively employed,” and concluding on that basis that “the doubts we entertain concerning the actual inventor of this device should be resolved in favor of the patentee”), quoted in Microsoft, 131 S. Ct. at 2249.

invalidity’s elevated burden of proof are beyond the scope of this article; instead, the point is that it has the effect of making it harder to win on invalidity defenses, which makes defendants more likely to rely on noninfringement.

2. Information and Timing Advantages

Information and timing advantages also make it harder for a defendant to win on a noninfringement defense. Litigating invalidity and noninfringement require different kinds of information. Invalidity is a question about the asserted patent, so it depends on information about that patent—its claims, specification, and prosecution history—and information about the state of the world when the patent was granted, to the extent the invalidity argument is based on prior art. Noninfringement, on the other hand, is a question about the accused product or process, so it depends on the features and workings of that product or process.121

Generally, patent plaintiffs and defendants will have different levels of access to these different kinds of information. And as a result of this asymmetry, accused infringers will almost always have better access to the information needed to litigate noninfringement, while patent holders will often have better access to the information needed to litigate invalidity.

Accused infringers will almost always have better access to information needed to litigate noninfringement because they usually have better access to information about their own accused products or processes. The defendant usually developed those products or processes, manufactures or uses them, and knows how they work and what their components and steps are. The plaintiff may have some of this information as well, since publicly available products can be purchased and analyzed, the results of processes can sometimes be obtained, and their makers and users often advertise specifications, schematics, and other details about the products’ ingredients and features. But often facts that are critical to the noninfringement analysis are expensive or impossible to obtain without inside information.

The extent of this information advantage can vary substantially by industry, type of patent, and type of product or process. Simple products and pro-

---

121 Both issues also depend on claim construction, which like invalidity depends on information about the patent and about the state of the world when the patent was issued.
cesses can lead to more symmetrical information, since it is easier to understand just by examining how a simple product or process works than to understand a complex one. It is far easier, for instance, to figure out how a tarp cover for a truck trailer works than to reverse engineer the source code to the graphics algorithms in Microsoft Windows.\textsuperscript{122} Regulated industries like food, pharmaceuticals, and wireless telecommunications, in which competitors must disclose much information to the government and the public—including to patent holders—or more competitive industries in which competitors advertise more information to the public can also lead to more symmetrical knowledge between litigants, since more information may be available to patent holders.\textsuperscript{123} And even when the industry and parties are the same, some types of patent claims lead to greater information advantages. In the pharmaceutical industry, for example, it is easy to tell if a drug infringes a patent on a specific active compound, since a chemist can just examine the accused drug and see if it uses the same compound, and since federal law typically requires that a generic drug contain the same active ingredient and be bioequivalent to the original name-brand drug.\textsuperscript{124} Without taking discovery, though, it may be impossible to tell if a drug was manufactured according to a patented manufacturing process, since the result may be identical even if the compound was manufactured by a different process.\textsuperscript{125}

For invalidity, similar information advantages can benefit patent holders, because patent holders sometimes (but not always) have better access to information about the patent and the prior art. If a patent holder filed the application that led to the asserted patent, then it will have preexisting knowledge about the content of that application and the prosecution of the patent. This might include knowledge about challenges to patentability raised by the patent examiner, or about prior art the patent was drafted to overcome. If the patent

\textsuperscript{122} Cf. Sundance, Inc. v. DeMonte Fabricating Ltd., 550 F.3d 1356, 1365–68 (2008) (holding invalid a patent on a tarp cover system for truck trailers, even after excluding all expert testimony supporting the invalidity determination, because "no such testimony [is] required" when "[t]he technology is simple and neither party claims that expert testimony is required"). (Full disclosure: I was one of DeMonte’s attorneys in the Federal Circuit appeal.)

\textsuperscript{123} E.g., AstraZeneca Pharms. LP v. Apotex Corp., 669 F.3d 1370, 1378 n.5 (Fed. Cir. 2012) (observing that the plaintiff’s complaint had relied upon information contained in the defendants’ FDA filings). See also Benjamin N. Roin, Unpatentable Drugs and the Standards of Patentability, 87 TEXAS L. REV. 503 (2009).


\textsuperscript{125} See, e.g., Julie E. Zink, Shifting the Burden: Proving Infringement and Damages in Patent Cases Involving Inconsistent Manufacturing Techniques, 2 HASTINGS SCI & TECH. L.J. 81 (2010).
holder practices its patented technology, or is an active competitor in the technology field into which the patent falls, then it may also have preexisting knowledge about the history of the field, i.e., about the prior art. An accused infringer may not have the same historical knowledge—especially if it is an upstart, or if it competes in a different market.

To take one example of a patent holder’s advantage in knowledge of prior art, Honeywell is a leading maker of thermostats for HVAC systems and has a long history in the field, having made thermostats for more than a century. Honeywell has many patents covering its product: a search of the Patent and Trademark Office patent database finds 142 patents issued since 1976 assigned to Honeywell with the word “thermostat” in the title. In 2012, Honeywell sued Nest Labs, a startup that makes a high-tech “smart” thermostat, for patent infringement. Nest was founded in 2010 by Tony Fadell, who had previously run Apple’s iPod division; it’s safe to say that Nest does not have the same sort of institutional knowledge as Honeywell has of the thermostat prior art.

These information advantages are not absolute or necessarily dispositive of the choice between arguing invalidity and arguing noninfringement. Sometimes there will be no advantage: a patent holder may have complete information (or at least adequate information) about how an accused product or

---

126 I say “if” because patents can be bought and sold like any other property, and in many cases the patent has been sold or transferred between its prosecution a subsequent patent lawsuit.


131 IBM is another good example. IBM has been the leading recipient of patents for every year in the last two decades, receiving patents in a variety of fields relating to computer software and hardware. Press Release, Int’l Bus. Machs. Corp., IBM Tops U.S. Patent List for 20th Consecutive Year (Jan. 10, 2013) at http://www-03.ibm.com/press/us/en/pressrelease/40070.wss. IBM was issued 6,478 patents in 2012, for a total of almost 67,000 patents between 1993 and 2012. Id. If IBM wanted to assert one of those patents, it would have a wealth of internal knowledge about the history of the computer industry that would help it find prior art and choose patents and claims that would more likely be upheld by the court.
process works, and an accused infringer may have as much or more information as the patent holder about the prior art. Other times the advantage can be overcome in the litigation process, through liberal fact discovery and disclosure rules. But even when the court mitigates the effect of any information advantages through discovery and disclosure requirements, developing invalidity and noninfringement arguments is costly enough that having an information advantage can be quite valuable at the margin.

Moreover, although the litigation process is designed to cure some of these information advantages, timing effects reinforce these information advantages by making it more difficult for discovery and research to fill in the gaps between patent holders and accused infringers. Discovery forces defendants to disclose information about how their products and processes work, while the availability of expert witnesses helps defendants—or their counsel, in the case of information subject to a protective order—understand this information. And the inherent delays in litigation give plaintiffs time to investigate prior art and develop validity arguments, while rules that require each side to disclose invalidity contentions and prior art on which they expect to rely help eliminate any gaps in plaintiffs’ and defendants’ knowledge about invalidity.132

These timing advantages arise because a patent holder typically can decide when to file a patent lawsuit.133 A patent holder can take as much time as needed before filing a lawsuit to research the prior art and position itself to defend against invalidity arguments. Since both a patent’s prosecution history

132 Information about prior art and a defendant’s invalidity contentions can be obtained through normal discovery tools such as document requests ("Produce copies of all prior art upon which you may rely to prove one or more of the asserted claims is invalid") and interrogatories ("State your complete basis for asserting that each asserted claim is invalid, including the specific prior art upon which you rely"), but several district courts have simplified the process by requiring that parties disclose certain contentions, typically by specific court-ordered deadlines. See, e.g., N.D. Cal. Loc. Pat. R. 3-1 to 3-4 (requiring parties to disclose infringement and invalidity contentions by certain deadlines); E.D. Tex. Loc. Pat. R. 3-1 to 3-4 (same); INTELLECTUAL PROP. OWNERS ASS’N, supra note 89, at §§ 3-1 to 3-5.

133 A potential defendant can bring a claim for a declaratory judgment that a patent is invalid, but only when there is a “substantial controversy” over the patent’s validity, “of sufficient immediacy and reality to warrant the issuance of a declaratory judgment.” MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 127 (2007) (quoting Maryland Casualty Co. v. Pacific Coal & Oil Co., 312 U.S. 270, 273 (1941)); see generally Nick Walrath, Expanding Standing in Patent Declaratory Judgment Actions to Better Air Public Policy Considerations, 88 N.Y.U. L. REV. (forthcoming 2013). If a patent holder has not taken “any affirmative action” asserting its patent rights, though, there is probably no such controversy. 8 CHISUM, supra note 105, at § 21.02[1][a][vi][C] (2012). A patent holder can stay under the radar, then, investigating and waiting until it is ready to bring a claim.
and the prior art are, generally, available to the public, a patent holder can effectively prepare much of its invalidity case before receiving disclosures from an accused infringer. The accused infringer, on the other hand, often has much less time to develop an invalidity case while the lawsuit is pending—as little as 45 days, under local rules in two districts that often hear patent cases. Conversely, although a patent holder can spend as much time as it wants investigating the accused product or process before filing suit, in many cases the information necessary to develop infringement arguments is not available before the lawsuit is filed and discovery is available. Yet the accused infringer typically has complete information about the accused product or process well before the lawsuit is filed, along with the expertise needed to quickly harness that knowledge. Once it learns of a lawsuit, then, a defendant can develop non-infringement arguments much more quickly than it can develop invalidity arguments.

As with the underlying information advantages, these timing advantages do not necessarily apply in every case, or even dictate which defenses a defendant will use when they do apply. In many cases, both patent holders and accused infringers have plenty of time to develop their arguments on invalidity and noninfringement. This can be because they anticipated litigation and prepared in advance, because they engaged in licensing negotiations before a lawsuit was filed and so had access to the same sorts of information they would later obtain in discovery, or because litigation proceeds slowly enough that they had time to develop their arguments even with the constraints litigation imposes. And in many cases, one defense is so much more viable than the other that the merits swamp any timing effects. But just as information advantages can have effects at the margin, timing effects that magnify or reinforce these advantages can likewise affect the defenses upon which accused infringers rely.

134 Before the America Invents Act, it was in rare cases possible to rely on prior art that was not available to the public. The America Invents Act expressly eliminates several of these corner cases, and may eliminate all of them. See Dennis Crouch, Did the AIA Eliminate Secret Prior Art?, PATENTLYO (Oct. 10, 2012), http://www.patentlyo.com/patent/2012/10/did-the-aia-eliminate-secret-prior-art.html.


136 See infra Part II.B.3.
3. Asymmetric Outcomes

Successful invalidity and noninfringement defenses also lead to different outcomes, which create post-judgment asymmetries. Both defenses, if successful, leave accused infringers free to continue to produce, use, and sell the accused products or processes, but each defense also has unique collateral consequences. A successful invalidity defense acts in rem on the patent claims that were found invalid: generally, those claims are dead and cannot be enforced in any subsequent patent litigation, against the same or different defendants. An invalidity judgment thus necessarily benefits anyone practicing the invention claimed in the patent (because he or she would otherwise be liable in a patent-infringement lawsuit) or practicing in the technology field to which the patent relates (because even if he or she does not infringe the patent, there is nevertheless a risk that the patent holder will sue). A successful noninfringement defense, on the other hand, means only that the defendant’s accused product or process does not satisfy every element of the asserted claims; the patent remains in force, limiting both the defendant’s later activities and those of its competitors. Such a judgment can benefit others, if it results in claim interpretations that make the patent less likely to be infringed. But in general, the benefits of a noninfringement judgment are local to the defendant in that case.

As others have recognized, this means that a successful invalidity defense is a public good. Any competitor can make use of the technology claimed in an invalid patent without infringing the patent, and the use by one competitor cannot prevent other competitors from simultaneously using the technology. And just as participants in a market economy tend to produce fewer public goods because they cannot internalize all of their benefits, defendants choosing between defenses will make fewer invalidity arguments because the benefits will flow both to the defendants themselves and to their competitors. All else being equal, then, a defendant would prefer to win with a noninfringement defense than with an invalidity defense.

139 In economics terms, an invalidity judgment is both non-excludable and non-rivalrous. This does not necessarily mean, by the way, that competitors are entirely free to use the technology claimed in the invalid patent. It is often the case that multiple patents apply to a particular technology, and even if one patent has been invalidated, others may prevent, or impose a cost upon, competitors. See, e.g., Shapiro, supra note 13.
The comparative advantage from winning a noninfringement judgment instead of an invalidity judgment can be quite substantial. The pharmaceutical industry provides a particularly clear example. Drugs can be protected by several kinds of patents, such as patents on active ingredients, patents on formulations for combining active ingredients with inactive ingredients to form finished drugs, and patents on processes by which drugs are made. Under the Hatch-Waxman Act, a company that develops a new drug must declare to the FDA a list of every patent protecting the drug or its use.\footnote{See generally 21 U.S.C. § 355(b).} A competitor that wants to sell a generic version of the drug cannot do so until those patents expire, unless it can show that the patents are invalid or not infringed by the proposed generic drug.\footnote{See generally id. § 355(j)(2)(A)(vii). Technically, the potential generic-drug maker must only certify that the patent is expired (or will expire), or that it is invalid or not infringed by the proposed generic drug. That declaration then acts as a technical act of patent infringement; the patent holder then has 45 days to file a patent lawsuit, automatically blocking a approval of the generic drug for thirty months, or until the patent case is decided. See generally id. § 355(j)(5)(iii).} And a potential generic competitor may be far better off proving noninfringement than invalidity. If it proves that the proposed generic version of the drug does not infringe the patent (for instance, because it is made by a different process, or it uses a different formulation), it may become the only generic competitor in the market.\footnote{See generally id. § 355(j)(5)(iii). It is not as strange as it may seem to say that a proposed generic drug would not infringe a (valid) patent on the name-brand drug, even for a patent that covers the drug or its formulation, instead of a manufacturing method that could differ between competitors. Under the Hatch-Waxman Act, a generic competitor must prove that its proposed generic drug is bioequivalent to the name-brand drug. Id. § 355(j)(2)(A)(iv); see also id. § 355(j)(8)(B) (defining bioequivalence). But generic drugs can and do have differences in their inactive ingredients, and those differences can cause the active ingredients to be released and absorbed at different rates. Those differences are normally small enough not to matter, but for drugs with a narrow therapeutic index—a narrow window between the amount of the drug necessary to be effective and the amount that would be toxic—they can be significant. See, e.g., Michelle Hottinger & Bryan A. Liang, Deficiencies of the FDA in Evaluating Generic Formulations: Addressing Narrow Therapeutic Index Drugs, 38 AM. J.L. & MED. 667 (2012); Lesley Alderman, Not All Drugs Are the Same After All, N.Y. TIMES, Dec. 19, 2009, at B6.} If it proves that the patent is invalid, on the other hand, other generic makers become free to enter. As economic theory predicts, duopoly and competition give dramatically different prices: an FDA study found that when a single generic drug competes with the brand-name drug, the generic drug costs 94% of what the brand-name drug costs, but when just two generic drugs compete with the brand-name drug, the
ratio falls to 52%. And the ratio keeps falling as the number of generic competitors increases: with more than five generic competitors, generic drugs cost, on average, less than 30% as much as brand-name drugs. A generic competitor that wins on noninfringement, then, may get substantially higher prices for its generic drug than it would if it won on invalidity.

C. What About the Merits?

It may seem that all this is much ado about nothing, that patent defendants will make the arguments that are most likely to win in any particular case. Even if defendants are forced to choose between invalidity and noninfringement, that may not lead to distortions if defendants just pursue whichever defense is stronger, given the law and facts applicable in a particular case. If merits asymmetries swamp the asymmetries discussed in this article, then litigation may already produce an optimal level of invalidity judgments. And this argument has some force, since the success rates of invalidity and noninfringement defenses seem to be closely linked to changes in the underlying substantive law. Indeed, as Glynn Lunney has shown, these success rates essentially flipped when the Federal Circuit was created: before 1982, about 75% of successful defenses were based on invalidity or unenforceability, compared to about 25% based on noninfringement, while after 1982, about 65% were based on noninfringement, compared to 35% based on invalidity and unenforceability. It would not be terribly surprising, then, if defendants switched their focus from invalidity to noninfringement in view of apparent changes in the two defenses’ strengths that came with the creation of the Federal Circuit.

Though it is surely the case that differences in their merits will weigh heavily on an accused infringer’s choice between defenses, there are nevertheless at least three reasons to think that other factors play an important role.

First, patent claims are ambiguous enough, or difficult enough to interpret, that in many cases both defenses are plausible and efforts put into one defense can substitute for efforts put into the other. In these cases, accused infringers can effectively decide which defense will be more plausible. Scholars have long

144 Id.
lamented the difficulty and uncertainty involved in construing patent claims, with the Federal Circuit reversing nearly half the claim constructions they review.\footnote{See, e.g., Schwartz, supra note 35, at 234–37, 259–60.} This gives parties leeway to argue strategically for, and obtain, their own preferred claim constructions.\footnote{Attorneys have significant agenda-setting power in litigation, and can affect not only the arguments that courts consider, but the likelihood that courts will reach specific outcomes. See, e.g., Kevin T. McQuire, Repeat Players in the Supreme Court: The Role of Experienced Lawyers in Litigation Success, 57 J. Pol. 187 (1995).} And different claim constructions will favor invalidity or noninfringement arguments.\footnote{See supra Part II.A.1.} This effect can be substantial: the difference between a broad claim construction and a narrow one can be the difference between a strong invalidity defense and a nonexistent one, and the inverse is equally true. As a result, then, of the indeterminacy of claim language, defendants have a substantial ability not only to decide which defense to pursue, but often to decide which defense will have merit.

This does not mean that every defendant is free to choose between invalidity and noninfringement; in many cases it is surely correct to say that one defense is so much stronger than the other that litigation and doctrinal asymmetries have little effect. But nor is it correct to say that there is always one clearly superior defense. The range of potential claim constructions is broad enough in many cases that a patent defendant has significant room to maneuver. Accordingly, the tradeoffs and asymmetries between invalidity and noninfringement can have an effect on the margin in surprisingly many cases.

Second, there is evidence that patent defendants care about the effects of their judgments on the post-judgment marketplace. Generic pharmaceuticals, which I described above,\footnote{See supra text accompanying notes 140–144.} provide an obvious example, since the price difference between duopoly—one possible result of a noninfringement verdict—and competition—the likely result of an invalidity verdict—is so substantial. But we also see signs in other industries. One famous example comes in the case of the Dutch electronics company Philips N.V. and patent-holder Gilbert Hyatt, who held dozens of patents for electronic components. Mr. Hyatt’s patents—including one patent on the basic concept of a microprocessor—were greeted with skepticism by the electronics industry when they were first granted.\footnote{Andrew Pollack, Inventor Finds Ally in Philips, N.Y. Times, Nov. 7, 1991. As the Times reported, “[t]he [microprocessor] patent was awarded 20 years after Mr. Hyatt applied for it, and he had never brought his idea to practice and was not considered to have played any important role in the subsequent development of the electronics industry.” Id.}
Regardless, Mr. Hyatt filed patent lawsuits against numerous electronics manufacturers, and in 1991 Mr. Hyatt settled with Philips. The Philips settlement had one striking feature: Philips agreed to partner with Mr. Hyatt to license the rest of the industry, with Philips getting a cut of Mr. Hyatt’s future licensing revenue.\textsuperscript{151}

Philips was one of Mr. Hyatt’s early targets, and it had substantial reason to think that many of Mr. Hyatt’s patents—including the microprocessor patent—were invalid. Other companies, like Intel and Texas Instruments, had done pioneering work to develop the first microprocessors well before the priority date claimed in Mr. Hyatt’s patents. Indeed, in the Patent and Trademark Office later invalidated Mr. Hyatt’s patent in an interference, finding that Mr. Hyatt had not reduced his invention to practice—a core patentability requirement—until years after another inventor did so.\textsuperscript{152} Yet rather than rely on these invalidity arguments, Philips licensed Mr. Hyatt’s patents and agreed to become his partner. Perhaps Philips just believed Mr. Hyatt’s patents were valid, or wanted to avoid the downside risk of a judgment of infringement.\textsuperscript{153} But another explanation is that Philips wanted to avoid conveying a gift on its competitors. Indeed, by helping Mr. Hyatt license his patents in return for part of the revenue, Philips imposed a tax on the rest of the electronics industry—a tax it estimated would be in the range of $100 million.\textsuperscript{154}

The Philips-Hyatt joint licensing venture is not unique; litigants settling patent disputes have devised several similar structures, including cross-licensing agreements, patent pools, and other joint ventures.\textsuperscript{155} What these structures all have in common is that they are all means by which patent holders and accused infringers can share the benefits of the patent monopoly. But such deals only make sense if both sides bring something to the table. Patentees, of course, contribute their patent rights. Although accused infringers do sometimes make substantive contributions to such ventures—including their own patents or other technical or business expertise—they can also contribute by forgoing their opportunity to invalidate the patent holder’s patents.

\begin{footnotes}
\item[151] \textit{Id.}
\item[153] \textit{See generally} Farrell & Merges, \textit{supra} note 58, at 948–60.
\item[154] Pollack, \textit{supra} note 150 (quoting a Philips employee saying total licensing revenue of $100 million “is quite possible”).
\item[155] \textit{See, e.g.,} Carl Shapiro, \textit{Antitrust limits to patent settlements}, 34 RAND J. ECON. 391 (2003).
\end{footnotes}
Third, even if a defendant does choose between defenses based solely on their merits, those merits do not exist in a vacuum. The relative strengths of invalidity and noninfringement defenses are products not only of the facts in a particular case, but of procedural and substantive rules that apply in the case, and the results of each defense. Rules like the presumption of validity and the requirement in many courts that the defendant provide invalidity contentions relatively early in the litigation make it harder or easier to succeed on a defense, changing its merits. The outcome of a successful defense also affects its strength, all else being equal, since parties will invest more to obtain a more favorable outcome. A defendant might prefer an invalidity judgment, for example, because it would prevent the patent holder from asserting the patent against future products, while a noninfringement judgment applies only to the specific products litigated in the case. Accordingly, it is impossible to separate the merits of a defense from the pre-judgment rules and post-judgment outcomes of that defense.

III. THREE REFORMS

This Part proposes three reforms that would help restore the balance between invalidity and noninfringement and make litigation a better tool for invalidating bad patents.

Since this imbalance is a product of the tradeoffs and asymmetries between invalidity and noninfringement, efforts to level the playing field between invalidity and noninfringement should work to eliminate, or counteract, one or more of these tradeoffs and asymmetries. Yet some of these tradeoffs and asymmetries cannot easily be changed, or should not be changed because they exist for good reasons. The three proposals in this Part would make it easier or more advantageous for defendants to argue invalidity without producing significant adverse side effects.

A. Eliminate the Elevated Burden of Proof for Invalidity

One useful reform would be to eliminate the elevated burden of proof that applies to invalidity, which is perhaps the most striking doctrinal asymmetry between invalidity and noninfringement. That burden, which stems from the statutory presumption that a patent is valid unless proved otherwise, makes it relatively more difficult to win an invalidity defense than a noninfringement defense, even if the two defenses would otherwise have similar merits.\textsuperscript{156} Scholars and others have long argued that this elevated burden of proof should

\textsuperscript{156} See infra Part II.B.1.
be eliminated, at least when the invalidity case rests on prior art that was not considered by the patent examiner, since the policies underlying it seem inconsistent with the modern reality of patent examination.  

This article provides a different justification for eliminating the elevated burden of proof for invalidity, one that does not depend on patent examiners’ expertise or the quality of their review. Instead, eliminating the elevated burden of proof for invalidity would help level the playing field for defendants choosing between noninfringement and invalidity defenses. It would not completely eliminate the difference between the two defenses, since the burden of proving invalidity would presumably still be on the defendant, while the burden of proving infringement is on the patent holder. But both issues would be decided by a preponderance of the evidence, making it easier for a defendant to focus on an invalidity defense.

B. Permit Defendants to Bifurcate Patent Litigation

Another useful reform would be to enact rules giving defendants the option to bifurcate patent cases, deferring infringement litigation until after the court has reached a final decision on validity. Such a bifurcation rule would directly eliminate only the trial-narrative tradeoff. But the rule would also have indirect effects, helping to counter the tradeoff stemming from resource constraints and the litigation-phase asymmetries stemming from information advantages and timing effects.

---

157 See id. To be sure, this view is not unanimous; many patent holders and practitioners have argued that eliminating the elevated burden of proof for invalidity would unduly weaken patent rights and the incentives to innovate. See generally Microsoft v. i4i Limited Partnership, SCOTUSBLOG, http://www.scotusblog.com/case-files/cases/microsoft-v-i4i-limited-partnership/ (linking to numerous amicus briefs in support of the elevated burden of proof for validity). These arguments have generally been rejected by legal scholars. Notably, of the 24 amicus briefs filed in Microsoft v. i4i in support of a elevated burden of proof for invalidity, none were filed by legal scholars. See id.

158 One could argue that if patent examiners were more reliable determiners of patent validity, there would be less need for litigation to invalidate patents, and so less need to level the playing field by eliminating the elevated burden of proof. Even if that were the case, though, an invalid patent that made it through the examination system would be all the more harmful, since the public would assume that most or all patents are valid and its holder would be able to extract correspondingly greater rents. And it is unlikely patent examiners will ever be able to approach 100% accuracy in determining validity, since the task of judging whether complex technologies are new or not is necessarily difficult.

159 Such a rule would standardize an area of practice that has previously been quite heterogeneous, with judges exercising their case-management discretion in many different ways. Some authorities argue that because invalidity is an affirmative defense, it should only be decided after the claim for relief—patent infringement—has been decided in a way that makes it necessary to
The effect of a bifurcation rule on trial strategy is straightforward. One reason patent litigants focus on a single defense is that it is hard to tell a simple and coherent trial narrative that embraces both invalidity and noninfringement. But if validity and infringement were decided in separate proceedings with separate trials, there would be no need to jettison one defense to maintain a coherent trial narrative.

A bifurcation rule would counteract resource constraints by deferring the cost of infringement litigation to a second litigation phase that may never occur. Such a rule would not completely eliminate the resource constraints that lead many litigants to focus on just one defense; patent litigants would still have limited litigation budgets, and litigating infringement and validity would still have costs that are largely separate. Indeed, such a rule might increase the cost of some patent cases, since whatever expenses that do apply to both invalidity and noninfringement litigation would now be separated across two phases. But it would also put off many costs associated only with infringement litigation—including particularly expensive costs like infringement experts and discovery about the defendant’s product—until the second phase, if one ever becomes necessary. This would allow defendants with limited litigation budgets to focus on invalidity without spending simultaneously on a noninfringement case that could be much more expensive.

(footnote continued from previous page)

consider the defense. Others respond that courts should consider validity first because leaving validity undecided, as would happen if a court does not find infringement, inevitably leaves some invalid patents in force. See generally 6 CHISUM, supra note 105, at § 19.02[1]. Both ordered sequences are common, as is simultaneous litigation of both invalidity and noninfringement; the venerable CHISUM on Patents hedges its bets, opining that “[w]hen both validity and infringement are at issue, the better practice is to consider the question of validity first or concurrently.” Id. at § 19.02.

A bifurcation rule for patent cases in American courts would not be unique; Germany has long decided validity and infringement in separate proceedings in separate courts. Infringement claims are considered private-law matters and are decided in ordinary civil courts. Invalidity claims, however, are consider public-law matters and are decided by a special patent court, the Bundespatentgericht. See Sarah R. Wasserman Rajec, Evaluating Flexibility in International Patent Law 22 n.83 (Feb. 23, 2013) (unpublished manuscript); Matthew A. Smith et al., Arbitration of Patent Infringement and Validity Issues Worldwide, 19 HARV. J.L. & TECH. 299, 334 (2006).

160 See supra Part IIA.2.

161 It may seem odd to discuss trials of validity issues, since patent validity is a question of law. Several forms of patent invalidity, however, turn on underlying factual questions, such as the content and scope of the prior art. Jury trials of invalidity issues are thus relatively common. See, e.g., 2 CHISUM, supra note 105, at § 5.04[3].
A bifurcation rule likewise cannot eliminate information advantages and timing effects, but it can mitigate them. Just as a bifurcation rule would let litigants defer spending on noninfringement arguments, it would allow them to focus their limited attention on developing the evidence needed for an invalidity defense. It would also help if courts, or rule drafters, were careful to give the parties enough time to develop their invalidity arguments.

While it is only a partial solution, the advantage to litigants from a bifurcation rule should not be minimized. Patent lawsuits are complex affairs with many moving parts. It is not unusual for a defense team to have lawyers working on areas as diverse as understanding the accused technology; analyzing the asserted patents and their prosecution histories; working with engineers from the defendant company; working with invalidity and noninfringement experts; analyzing prior art; analyzing documents from, and deposing, the inventor; analyzing documents from, and deposing, witnesses from the plaintiff company; and working on a whole set of remedy issues, including working with experts and fact discovery to make arguments about damages and a potential injunction. Splitting a case into two parts would cut the number of simultaneous moving parts in half, turning it into a far more manageable coordination problem. This, in turn, would make it much easier for defendants to do the research into the prior art and the asserted patents’ prosecution histories necessary to present an invalidity defense.

The main objection to a bifurcation rule is likely to be that it will just make patent litigation slower, less efficient, and more costly for litigants. This objection has some force, since patent litigation is already a notoriously slow and expensive process. It is possible, though, that a bifurcation rule would reduce the average length and cost of patent litigation, because patent holders might bring fewer cases or because in many cases there would be no need to ever reach the infringement phase. Indeed, several studies have found that

---

162 The median patent case takes more than 30 months to reach trial, and almost 20% of cases take more than four years to reach trial. PRICEWATERHOUSECOOPERS LLP, 2012 PATENT LITIGATION STUDY 20–21, available at http://www.pwc.com/en_US/us/forensic-services/publications/assets/2012-patent-litigation-study.pdf.

163 Studies of bifurcation rules in ordinary civil litigation, in which phase 1 decides liability and phase 2 decides damages, predict that plaintiffs will be more likely to bring lawsuits when all issues will be decided by a single unitary trial, and that significant cost savings can be achieved from avoiding phase 2 when phase 1 finds no liability, at least when economies of scale from combining the two phases are insubstantial. See Jef De Mot, Sequential trials and the English rule, 34 EUR. J.L. & ECON. 31 (2012); Kong-Pin Chen, Hung-Ken Chien & C.Y. Cyrus Chu, Sequential Versus Unitary Trials with Asymmetric Information, 26 J. LEGAL STUD. 239 (1997); DOUGLAS G. BAIRD, ROBERT H. GERTNER & RANDAL C. PICKER, GAME THEORY AND THE LAW 255— (footnote continued on next page)
almost half of patents litigated to a final determination are ultimately invalidated by courts.\textsuperscript{164} And even if some cases become more expensive to litigate, the savings in cases when patents are invalidated may outweigh that added expense, since litigating infringement is likely more expensive than litigating invalidity.\textsuperscript{165}

C. Allow a Successful Patent Challenger to Recover from Other Beneficiaries of an Invalidity Judgment

Finally, another useful reform would be to create a way for accused infringers who successfully challenge and invalidate patents in litigation to recover some of the value they create by eliminating an invalid patent. I proposed to do so by creating a new cause of action for an accounting, which would permit a party that invalidates a patent to recover from industry competitors that otherwise would have been susceptible to a claim for infringing that patent. This would help eliminate the asymmetry in outcomes between successful invalidity and noninfringement defenses.

Several scholars have proposed similar bounty systems to overcome the public-good nature of invalidity judgments and reward successful challenges to patents’ validity.\textsuperscript{166} These bounty systems would essentially pay for successful invalidity challenges, counteracting, but not eliminating, the asymmetrical incentives to argue invalidity and noninfringement. While these bounty sys-

\textsuperscript{164} See Allison & Lemley, supra note 10, at 205–06 & n.52 (finding an invalidity rate of 46% and summarizing prior studies).

\textsuperscript{165} This is true for three reasons. First, the fact discovery needed to litigate noninfringement is more extensive than that needed for invalidity, since document discovery and testimony concerning the defendant’s accused products or processes is typically much more voluminous than information about a patent’s prosecution history and prior art. Second, discovery about damages and remedy is only relevant the infringement phase. And third, invalidity can be decided on summary judgment in far more cases than for infringement, since the scope and content of the prior art is often undisputed. Expensive trials, then, are needed far more often to decide infringement than to decide validity.

\textsuperscript{166} See Miller, supra note 83, at 704–30 (proposing that a patent’s holder be required to pay a bounty to a successful challenger of that patent); Michael J. Meurer, Controlling Opportunistic and Anti-Competitive Intellectual Property Litigation, 44 B.C. L. Rev. 509, 535–38 (2003) (proposing fee shifting to discourage “opportunistic” patent litigation); Kesan, supra note 74, at 786–96 (proposing a “one-way, pro-defendant, fee-shifting system” when patent claims are invalidated based on specific kinds of prior art); Thomas, supra note 83, at 340–47 (proposing that the PTO pay patent bounties to members of the public who provide information leading patent applications to be rejected).
tems would help strengthen an accused infringer’s incentive to rely on an invalidity defense instead of noninfringement, they have two significant limitations. First, most of the proposals rely on the holder of the invalid patent to pay the bounty. Many patent plaintiffs, however, are judgment-proof, and this is especially true for the non-practicing entities (or patent trolls) that bring many patent-infringement claims. And second, they would still leave a defendant’s competitors in a position to capture a windfall, benefiting from a patent being invalidated without having to undertake the effort or risk of attempting to invalidate the patent. Although a large-enough bounty could make up for this shortcoming, simple fee shifting and similar bounties would not do the job.

The proposed cause of action for an accounting would provide a bounty that would be paid instead by the most direct beneficiaries of an invalidity judgment: the businesses and individuals, typically competitors of the successful patent challenger, who would otherwise be vulnerable to a claim for infringing that patent. These businesses and individuals benefit directly from an invalidity judgment because they can no longer be sued for infringing the patent or, if they have previously licensed the patent, would no longer have to pay royalties for that patent. By transferring some of this benefit to the party that successfully challenged the patent, the accounting action would reduce the asymmetry in outcomes between successful invalidity and noninfringement defenses, while reducing the benefit to a defendant’s direct competitors from an invalidity judgment.

One objection to this proposal is likely to be that it would just create a system of secondary patent-infringement litigation, with a successful defendant turning around and having to prove that its competitors infringe a now-invalid patent. This issue can be avoided, however, by basing liability under the ac-

---

167 John Thomas has proposed instead that the PTO pay bounties. Id. Though this would avoid the problem of judgment-proof patent holders, it would require potential bounty collectors to come forward far earlier, well before it becomes clear that a patent poses a realistic litigation threat. Miller, supra note 83, at 698. It would also likely pay bounties that are too small to change incentives meaningfully.


counting action not on whether one actually would have infringed the now-invalid patent, but on whether the patent holder could have stated a plausible claim for patent infringement. Such plausibility standards are common in other areas of the law. And such a standard is a better match for the real benefits of an invalidity judgment, since an invalid patent imposes a tax not only on those who would infringe, but on any party that could be plausibly sued for patent infringement, and so would have to pay royalties or incur litigation costs.

Another objection to this proposal is that it conflicts with the public policy favoring settlement of private disputes. It is undoubtedly true that an accounting action would give defendants more to gain from refusing to settle, though it does not follow that settlement will become more likely. Instead, one possible outcome is that it could change the relative bargaining positions of plaintiffs and defendants, leading to fewer lawsuits or to settlements that are more favorable to defendants accused of infringement. But even if the accounting action did make settlements less common, that might not be a bad thing. The public policy favoring settlements is motivated by judicial economy, the need to reduce the risk of litigation, and the feeling that privately crafted resolutions are more likely to respond to private litigants’ real-world needs. But these reasons are less compelling when there is a significant countervailing public interest—such as the need to eliminate invalid patents—in the litigation. Indeed, the Supreme Court has repeatedly recognized that the public interest in free competition, unencumbered by invalid patent monopolies, can overcome traditional equitable considerations.

---

170 The Rule 11 standard for sanctionable litigation misconduct provides a good example. See Fed. R. Civ. P. 11; Ashcroft v. Iqbal, 556 U.S. 662 (2009); Bell Atlantic Corp. v. Twombly, 550 U.S. 544 (2007). Another good example comes from insurance-coverage disputes, in which the question is typically not whether the policyholder is liable under one of the grounds covered by the policy, but whether the policyholder has been sued for any claim that is potentially covered by the policy. See generally Steven P. Inman & Robert Kinder, Securing Insurance Coverage For Patent Infringement Lawsuits Under CGL Insurance Policies Could Save Millions, INTEL. PROP. TODAY, Mar. 2012, at 13–15.


172 See Lear, 395 U.S. at 670 (holding that estoppel does not prevent a patent licensee from challenging the patent’s validity, in part because “the equities of the licensor do not weigh very heavily when they are balanced against the important public interest in permitting full and free competition in the use of ideas which are in reality a part of the public domain”); Pope Mfg. Co. v. Gormully, 144 U.S. 224 (1892) (similar).
Although this proposal would help counteract the collective-action problem stemming from the public-good nature of invalidity judgments, it would not completely align a potential challenger’s incentive with society’s interest in eliminating bad patents. As Joseph Farrell and Robert Merges have recognized, the public-good problem is not the only reason for incomplete incentives to challenge patents; potentially as troublesome is the pass-through problem, under which royalty costs are passed along to consumers.\textsuperscript{173} But reducing one problem is better than reducing no problems, and combined, the proposed reforms could significantly reduce the tradeoffs and asymmetries that lead to courts invalidating too few patents.

CONCLUSION

Most patent scholars and lawyers agree that there are too many bad patents in force, imposing a tax on creators of valuable goods and services, and that the legal system needs effective mechanisms to eliminate those bad patents. One classic mechanism for invalidating patents is infringement litigation. Yet, as this article has shown, the patent system makes it harder than necessary for defendants to rely on invalidity defenses, and thus for district courts to invalidate bad patents. Through a series of tradeoffs and asymmetries between invalidity and noninfringement, defendants are encouraged to focus their efforts on just one defense, and usually to make that defense noninfringement. Because of this, defendants challenge, and district courts invalidate, fewer patents than they should; patent holders are able to extract more rents using invalid patents; and more invalid patents restrict the behavior of the public.

Though there is no one perfect solution, a series of reforms could significantly reduce the effect of these tradeoffs and asymmetries. Eliminating the elevated burden of proof for invalidity would make it relatively easier for defendants to rely on invalidity defenses. A bifurcation rule, requiring district courts to decide validity before infringement, would make it easier for litigants to develop coherent trial narratives for both validity and infringement issues, while helping them focus their resources on validity issues early in a case. And a new cause of action for an accounting would reduce the collective-action problem posed by the public-good nature of an invalidity judgment.

\textsuperscript{173} Farrell & Merges, supra note 58, at 953–55.