

IMPROVING PATENT QUALITY WITH APPLICANT
INCENTIVES

Stephen Yelderman*

TABLE OF CONTENTS

| | |
|--|-----|
| I. INTRODUCTION | 78 |
| II. BACKGROUND | 81 |
| III. APPLICANT INCENTIVES IN SELECTING CLAIM SCOPE | 85 |
| A. <i>A Simplified Model of Claim Drafting</i> | 87 |
| 1. Claim Value as a Function of Scope | 87 |
| 2. Prosecution Cost Constraints and the Selection of Claims for Filing | 91 |
| B. <i>The Presumption of Patent Validity</i> | 95 |
| 1. Assumptions About the Presumption of Validity | 95 |
| 2. The Effects of the Presumption of Validity | 97 |
| 3. Assessing the Presumption of Validity | 101 |
| C. <i>Substantive Patent Examination</i> | 102 |
| 1. The Effect of Substantive Examination on Expected Costs | 103 |
| 2. Counterbalancing the Presumption of Patent Validity | 105 |
| 3. The Benefits of a Substantive Examination System | 107 |
| D. <i>Other Asymmetries in the Treatment of Validity and Infringement</i> | 108 |
| 1. Re-Litigation of Validity and Infringement | 108 |
| 2. Procedural Preferences for Narrowing Amendments | 111 |
| IV. REFORMS FOR IMPROVING PATENT QUALITY | 112 |
| A. <i>Rationalizing Application Fees</i> | 113 |
| B. <i>Penalizing Amendments and Abandonment</i> | 120 |
| C. <i>Replacing the Presumption of Validity with a Presumption of Infringement</i> | 121 |
| D. <i>Penalizing Invalid Claims After Examination</i> | 125 |
| V. REFINING THE MODEL | 128 |
| A. <i>Some Applicants Have No Intention of Enforcing Their Patents</i> | 129 |

* Associate Professor of Law, Notre Dame Law School. I would like to thank Dennis Crouch, Richard Epstein, Lee Fennell, Roger Ford, Timothy Holbrook, Daniel Kelly, Brian Leiter, Mark Lemley, Alexandra Levy, Jonathan Masur, Mark McKenna, Adam Mossoff, Gary Myers, Greg Reilly, Michael Risch, Lior Strahilevitz, and Katherine Strandburg for very helpful comments and suggestions, and Ravi Fernando and Dillon Cawley for excellent research assistance.

| | |
|---|-----|
| <i>B. Enforcement Outcomes May Diverge from Expected</i> | |
| <i>Litigation Outcomes</i> | 133 |
| <i>C. Even Invalid Patents Have Portfolio Value</i> | 135 |
| VI. CONCLUSION | 136 |

I. INTRODUCTION

Everyone, it seems, wants to improve the quality of issued patents. Over the last decade, legal scholars from every corner have come forward to decry the unacceptably high number of invalid patents issued by the U.S. Patent and Trademark Office (“USPTO”) and to suggest reforms to reduce it.¹ Some commentators have sought to improve the USPTO itself, calling for more funding, better administrative oversight, and changes to the way individual patent examiners are compensated.² Others have looked to third parties, seeking to make it easier for competitors and members of the public to submit prior art and participate in opposition proceedings.³ Whatever the means, these reforms share a common and urgent goal: making the examination process more effective at identifying and rejecting low-quality patent applications.⁴

However, none of these proposals have focused on improving the quality of the applications themselves. Applications (of whatever

1. See generally Joseph Farrell & Robert P. Merges, *Incentives to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help*, 19 BERKELEY TECH. L.J. 943 (2004); Jay P. Kesan, *Carrots and Sticks to Create a Better Patent System*, 17 BERKELEY TECH. L.J. 763 (2002); Michael Meehan, *Increasing Certainty and Harnessing Private Information in the U.S. Patent System: A Proposal for Reform*, 2010 STAN. TECH. L. REV. 1 ¶1 (2010), available at <https://journals.law.stanford.edu/sites/default/files/stanford-technology-law-review/online/meehan-increasing-certainty.pdf>; Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 BERKELEY TECH. L.J. 577 (1999) [hereinafter Merges, *Six Impossible Patents*]; Joseph S. Miller, *Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents*, 19 BERKELEY TECH. L.J. 667 (2004); Kristen Osenga, *Entrance Ramps, Tolls, and Express Lanes — Proposals for Decreasing Traffic Congestion in the Patent Office*, 33 FLA. ST. U. L. REV. 119 (2005); John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U. ILL. L. REV. 305 (2001).

2. See ADAM B. JAFFE & JOSH LERNER, *INNOVATION AND ITS DISCONTENTS* 34–35 (2004); Mark A. Lemley, *Can the Patent Office Be Fixed?*, 15 MARQ. INTELL. PROP. L. REV. 295, 297–304 (2011); Thomas, *supra* note 1, at 323–26.

3. See Kesan, *supra* note 1, at 776–77; Jay Kesan & Andres Gallo, *Why “Bad” Patents Survive in the Market and How Should We Change? — The Private and Social Costs of Patents*, 55 EMORY L.J. 61, 107–12 (2006); Merges, *Six Impossible Patents*, *supra* note 1, at 604–05; Thomas, *supra* note 1, at 326–32, 340–44. Similar reforms were recently included in the America Invents Act. See Leahy-Smith America Invents Act, Pub. L. No. 112-29, §§ 6, 8, 125 Stat. 284, 299–313, 315–16 (2011) (codified in scattered sections of 35 U.S.C.) [hereinafter AIA] (reforming post-grant review proceedings and creating avenues for third parties to submit prior art during examination).

4. But see Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1508–11 (2001) (suggesting the cost of increased scrutiny may outweigh its benefits).

form) are taken as a given, and the real work of ensuring patent quality is thought to start when one shows up at the door of the USPTO. Commentators have recognized that patent applicants sometimes possess useful information,⁵ but beyond suggesting a few mechanisms to encourage applications to pass that information along to examiners,⁶ scholars have not asked applicants to do much else. The possibility that applicants might have a more direct role to play in improving patent quality has gone largely unexplored.

This is a significant oversight. Applicants are particularly well-placed to influence and improve patent quality. Applicants, after all, are the ones who draft the claims that will later be examined by the USPTO and litigated by others. They often understand their inventions better than anyone else.⁷ They may possess the best knowledge of relevant prior art.⁸ And they often have a substantial investment in the outcome of their applications. Indeed, a conservative estimate is that patent applicants spent about \$7.5 billion pursuing patents in 2012 — dwarfing the approximately \$1.4 billion the USPTO spent examining applications the same year.⁹ To ignore applicants, then, is

5. See Kesan, *supra* note 1, at 768; R. Polk Wagner, *Reconsidering Estoppel: Patent Administration and the Failure of Festo*, 151 U. PA. L. REV. 159, 212–14 (2002); Mark A. Lemley, *Fixing the Patent Office* 10–11 (Stanford Law and Economics Olin Working Paper No. 422, 2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2023958.

6. See, e.g., Kesan, *supra* note 1, at 770, 773–77; Mark A. Lemley et al., *What to Do About Bad Patents?*, 28 REGULATION 10, 12–13 (2005); Osenga, *supra* note 1, at 141–51. In addition, several commentators have noted ways that existing patent rules fail to incentivize, or actively disincentivize, applicants to disclose relevant information to the USPTO. See Timothy R. Holbrook, *Patents, Presumptions, and Public Notice*, 86 IND. L.J. 779, 804–06 (2011); R. Polk Wagner, *Understanding Patent-Quality Mechanisms*, 157 U. PA. L. REV. 2135, 2149 (2009) [hereinafter Wagner, *Patent-Quality Mechanisms*]; Wagner, *supra* note 5, at 214–15; Michael C. Deane, Note, *Aligning the Incentives to Disclosure of Relevant Information to the USPTO: How the Jury Instruction Standards in Microsoft Corp. v. i4i Could Aid the Examination Process*, 92 TEX. L. REV. 439 (2013).

7. See Christopher A. Cotropia, *Modernizing Patent Law's Inequitable Conduct Doctrine*, 24 BERKELEY TECH. L.J. 723, 753 (2009); Ronald J. Mann & Marian Underweiser, *A New Look at Patent Quality: Relating Patent Prosecution to Validity*, 9 J. EMPIRICAL LEGAL STUD. 1, 27 (2012); Miller, *supra* note 1, at 733–34 (noting resource advantage of applicants over USPTO).

8. See Cotropia, *supra* note 7, at 751, 753–54; Kesan, *supra* note 1, at 768 (observing that valuable information about prior art is often in the hands of the patent applicant); Jay P. Kesan & Marc Banik, *Patents as Incomplete Contracts: Aligning Incentives for R&D Investment with Incentives to Disclose Prior Art*, 2 WASH. U. J.L. & POL'Y 23, 52 (2000).

9. See U.S. PATENT & TRADEMARK OFFICE, PERFORMANCE AND ACCOUNTABILITY REPORT FOR FISCAL YEAR 2013, at 80–81 (2013), available at <http://www.uspto.gov/about/stratplan/ar/USPTOFY2013PAR.pdf>. The approximation of applicants' costs is based on an estimate that a patent applicant spends approximately \$15,000 in attorneys' fees over the course of prosecution, see Lemley, *supra* note 4, at 1498–99 & n.13, and USPTO data showing that between 2010 and 2012, approximately 505,000 utility applications were filed per year, see U.S. PATENT & TRADEMARK OFFICE, *supra*, at 189. The result — \$7.575 billion — is of a similar order of magnitude to that of Lemley's estimate that \$4.33 billion were spent on patent applications in 2000. See Lemley, *supra* note 4, at 1499; see also AM. INTELLECTUAL PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY 27 (2013) (reporting that the median attorneys' fees for preparation of an original application for relatively complex patent was between \$8500 and \$10,000 in

to neglect a group that could have a significant effect on patent quality.

Perhaps one reason for this neglect is that many scholars assume that applicants want the kind of vague, overbroad, or otherwise low-quality patents that patent reformers are seeking to prevent. Understandably, some commentators regard applicants with suspicion; after all, applicants are attempting to claim private property rights instead of allowing their inventions to enter the public domain. Given applicants' inherent self-interest, it is reasonable to suppose that they will generally try to obtain the broadest patent rights possible.¹⁰ This suspicion, in turn, leads to an adversarial view of patent prosecution: applicants will seek broad, low-quality patents, and it is the role of the USPTO (and the occasional third-party intervenor) to push them back.

However, while it is true that a broader patent is generally more likely to be infringed, this benefit is subject to a critical constraint: a claim that is *too* broad runs the risk of being invalid, and may fare no better in litigation than a claim that is too narrow.¹¹ Because a successful patent infringement suit requires prevailing on both validity and infringement, an applicant may have just as much to lose from her claim being too broad as she does from her claim being too narrow.¹² Beyond a certain point, the broader the claim, the *less* valuable it becomes. And as a claim's likelihood of being valid continues to diminish, its value may become so low that the expected costs of filing it may start to outweigh its expected benefits.¹³

This insight shifts the conventional view of applicants in a subtle but important way. Applicants do not seek overbroad claims because they are inherently the most valuable claims, but rather because current prosecution and litigation rules cause the expected value of these claims to exceed their expected cost.¹⁴ Enlisting applicants' help to improve patent quality therefore does not require a complete revamping of the current *ex parte* prosecution system. Instead, improving patent quality could simply be a matter of increasing the costs and

2012); David Fagundes & Jonathan S. Masur, *Costly Intellectual Property*, 65 VAND. L. REV. 677, 689 (2012) (estimating that attorneys' fees for preparation of an initial application range from \$ 11,000 to \$ 15,000).

10. See Fagundes & Masur, *supra* note 9, at 687; Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 840 (1990). Also, other commentators have noted applicants' interest in obtaining malleable or otherwise vague claims. See Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2147–51.

11. See 35 U.S.C. §§ 282, 102, 103 (2012). Claims that are declared invalid are frequently so declared because they are too broad. See John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q. J. 185, 207–08 (1998).

12. A variety of reasons why a patentee may prefer claims that are too broad over claims that are too narrow or vice versa are discussed in Parts III.B, III.D, V.B, *infra*.

13. A variety of factors that influence the expected cost of filing a claim are discussed in Parts III.A, III.C, IV.A, IV.B, *infra*.

14. See *infra* Parts III.A, III.B.

reducing the benefits of low quality claims, so that applicants will on average file fewer of such claims in the first instance.

As this Article explains, the appeal of overbroad claims can be reduced in a variety of ways. Some of these reforms are simple — for example, changing the structure of the fees paid to the USPTO to ensure that applicants weigh the incremental costs and benefits of each claim that they file.¹⁵ Others are more elaborate — for example, flipping litigation presumptions to make narrow claims more valuable than broad ones.¹⁶ In between is a range of proposals to adjust applicant incentives in favor of narrower and higher quality claims.¹⁷

This Article proceeds in six parts, starting with this Introduction. Part II provides background regarding current problems of patent quality and the reforms that have been previously proposed to address the issue. Part III examines how current patent rules may affect applicants' incentives to seek particular kinds of claims. Part IV proposes several changes to patent law that could be expected to increase applicants' incentives to seek narrower claims and to reduce their incentives to seek broader ones. Part V raises and addresses several objections to the assumptions relied on in Parts III and IV. Part VI concludes.

II. BACKGROUND

The process of acquiring a patent starts by an inventor filing an application at the USPTO. The applicant files a description of her invention and a declaration that she in fact invented it, and pays a fee.¹⁸ She must also include one or more claims. Each claim is a technical description that provides the legal definition of what constitutes infringement of the patent.¹⁹ Although the claims contained in the same application are supposed to be generally related to each other, each is a legally distinct invention that may present independent questions of validity and infringement.²⁰

After the inventor files her application, she waits. Depending on the field of her invention, that wait could last for months or years.²¹

15. *See infra* Part IV.A.

16. *See infra* Part IV.C.

17. *See infra* Parts IV.D, IV.B.

18. *See* 35 U.S.C. § 111 (2012). In some cases an application must also include an illustration. *See id.* § 113.

19. *See id.* § 112.

20. *See id.* §§ 282, 288 (validity of claims determined individually); *Leeds & Catlin Co. v. Victor Talking Machine Co.*, 213 U.S. 301, 319 (1909) (“Claims are independent inventions. One may be infringed, others not.”). By contrast, a finding of inequitable conduct renders all the claims of a patent unenforceable. *Kingsdown Med. Consultants, Ltd. v. Hollister Inc.*, 863 F.2d 867, 877 (Fed. Cir. 1988) (*en banc*).

21. *See* U.S. PATENT & TRADEMARK OFFICE, PERFORMANCE AND ACCOUNTABILITY REPORT FOR FISCAL YEAR 2011, at 14 (2011), available at <http://www.uspto.gov/>

Eventually, the USPTO responds with an “office action” — a letter indicating which of the applicant’s proposed claims were found to meet the statutory requirements for patentability and which were not. Usually, at least some of the claims are rejected for failing to meet those standards.²² But, such a rejection is not the end of the road. Rather, in response to a rejection, an applicant has several options: she may attempt to persuade the examiner that the claim as written meets the statutory requirements for patentability, she may amend the claim in an attempt to make it satisfy those requirements, or she may (after fulfilling certain procedural prerequisites) take an appeal.²³ This cycle of rejection, argument, and amendment may go on for some time.²⁴

If at any point in this process the USPTO determines that the applicant’s claims meet the statutory requirements for patentability, the USPTO will issue a notice of allowance.²⁵ If the applicant chooses to proceed, she pays an additional fee, and the USPTO publishes her patent, mails her a formal certificate, and declares her a patentee.²⁶ Once the patent formally issues, the applicant (now a patentee) holds a time-limited right to exclude others from practicing the invention as defined by her issued claims.²⁷

Yet despite this expensive²⁸ and time-consuming²⁹ investment in pre-issuance examination, the USPTO makes mistakes. When validity is challenged in court, roughly half of litigated patents are found to contain one or more invalid claims.³⁰ And although it is unknown ex-

about/stratplan/ar/USPTOFY2011PAR.pdf (reporting average first action pendency between twenty-five and twenty-eight months and average total pendency between thirty-one and thirty-six months during 2007 through 2011).

22. See Mark A. Lemley & Bhaven Sampat, *Examining Patent Examination*, 2010 STAN. TECH. L. REV. 2 ¶ 7 (2010), available at <https://journals.law.stanford.edu/sites/default/files/stanford-technology-law-review-stlr/online/lemley-sampat-examining-patent.pdf>.

23. See 37 C.F.R. § 1.121 (2013) (rules for claim amendments); *id.* § 41.31(a) (requirements to take an appeal); Lemley & Sampat, *supra* note 22, ¶¶ 6–15 (discussing alternatives to either arguing against a rejection or amending a claim).

24. See Mark A. Lemley & Kimberly A. Moore, *Ending Abuse of Patent Continuations*, 84 B.U. L. REV. 63, 71 (2004) (finding, during period of study, that prosecution took an average of 2.47 years from earliest claimed filing date to issuance date); Lemley & Sampat, *supra* note 22, ¶¶ 19–24 (discussing use of Requests for Continued Examinations (“RCE”)s by applicants to extend debate over an application’s rejection).

25. See 35 U.S.C. § 151 (2012). See generally U.S. PATENT & TRADEMARK OFFICE, MANUAL OF PATENT EXAMINING PROCEDURE ch. 1300 (2014), <http://www.uspto.gov/web/offices/pac/mpep/mpep-1300.pdf>.

26. See 35 U.S.C. §§ 151, 153 (2012).

27. See *id.* §§ 154, 271. Even after some claims have issued, the applicant may continue to pursue other claims through a continuation application. See *id.* § 120; Lemley & Moore, *supra* note 24, at 68.

28. See Lemley, *supra* note 4, at 1499 (estimating that patent applicants spend in excess of \$ 4.3 billion a year obtaining patents).

29. See U.S. PATENT & TRADEMARK OFFICE, *supra* note 21, at 14.

30. See Allison & Lemley, *supra* note 11, at 205; Mann & Underweiser, *supra* note 7, at 7. Note that a fifty percent rate of invalidity of litigated patents is necessarily a product of the selection of disputes for litigation and does not imply that fifty percent of issued patents

actly how many of the roughly two million patents in force today are actually invalid,³¹ there is a widespread perception among practitioners and commentators that invalid claims continue to issue from the USPTO at an unacceptably high rate.³²

Poor patent quality — the tendency of issued patents to have one or more invalid claims³³ — imposes a variety of costs. A patentee can brandish an invalid claim to scare away competitors, increase would-be entrants' costs of capital, or drive away competitors' customers — even if the patentee never enforces it.³⁴ Invalid claims limit commercial freedom, deter potential innovators, and lead to wasteful legal and technical undertakings.³⁵ They increase uncertainty in developing industries,³⁶ and generally erode public and judicial confidence in the patent system.³⁷ Although these costs are difficult to quantify, one estimate puts them in the hundreds of millions of dollars annually.³⁸

A variety of reforms have been proposed to improve patent quality through better examination. Many commentators have suggested procedural and institutional changes to improve the capabilities of the USPTO.³⁹ Others have looked at ways to increase participation by

are invalid. *See generally* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984) (discussing selection bias in litigation).

31. *See* Lemley, *supra* note 4, at 1501.

32. *See* JAFFE & LERNER, *supra* note 2, at 34–35; Kesan, *supra* note 1, at 765–66; Kesan & Gallo, *supra* note 3, at 63–68; Scott Kieff, *The Case for Registering Patents and the Law and Economics of Present Patent-Obtaining Rules*, 45 B.C. L. REV. 55, 57 (2003); Lemley, *supra* note 2, at 296–97; Mann & Underweiser, *supra* note 7, at 1; Merges, *Six Impossible Patents*, *supra* note 1, at 589–91; Thomas, *supra* note 1, at 316–22; Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2144–45.

33. Although other metrics of patent quality are certainly possible, *see* Mann & Underweiser, *supra* note 7, at 2–4, this Article focuses on statutory validity as the primary proxy for patent quality.

34. *See* Christopher R. Leslie, *The Anticompetitive Effects of Unenforced Invalid Patents*, 91 MINN. L. REV. 101, 125–27 (2006). *But see* Lemley, *supra* note 4, at 1515–17 (suggesting that costs of these “in terrorem” effects have been overstated).

35. For example, an invalid claim may lead to costly litigation, unnecessary efforts to design around the claim, and reduced competition in affected markets. *See* T. Randolph Beard et al., *Quantifying the Cost of Substandard Patents: Some Preliminary Evidence*, 12 YALE J.L. & TECH. 240, 245 (2010); Farrell & Merges, *supra* note 1, at 945–46; Merges, *Six Impossible Patents*, *supra* note 1, at 592; Osenga, *supra* note 1, at 126–27.

36. *See* John R. Thomas, *The Responsibility of the Rulemaker: Comparative Approaches to Patent Administration Reform*, 17 BERKELEY TECH. L.J. 727, 731 (2002); Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2140–44.

37. *See* *Ebay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388, 397 (2006) (Kennedy, J., concurring) (noting that “potential vagueness and suspect validity of some . . . patents” may affect the availability of injunctive relief); Doug Lichtman & Mark A. Lemley, *Rethinking Patent Law's Presumption of Validity*, 60 STAN. L. REV. 45, 72 (2007); Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2141–42, 2144–45.

38. *See* Lemley, *supra* note 4, at 1507–08.

39. *See* JAFFE & LERNER, *supra* note 2, at 35; Lemley, *supra* note 2, at 297–304; Thomas, *supra* note 1, at 323–26. The AIA implemented reforms using similar mechanisms. *See* AIA §§ 6, 22 (codified in scattered sections of 35 U.S.C.) (reforming USPTO's post-grant review proceedings and requiring that patent application fees be used only for activities relating to processing patent applications).

third parties, such as hiring for-profit examiners, involving competitors in the review process, and making it easier for the public to oppose problematic applications.⁴⁰ A related line of proposals has focused on improving the flow of relevant information to the government's patent examiners with mechanisms such as increased disclosure burdens on applicants, rewards for information disclosure, and crowdsourcing of additional scientific expertise.⁴¹

However, not everyone agrees that investing additional public and private resources in examination is the answer. Some scholars have suggested that the marginal benefits of increased examination cannot justify its likely enormous marginal costs. The volume of patent applications is so massive, these commentators note, that implementing even a modest amount of increased scrutiny could easily cost more than a billion dollars annually.⁴² And, because only a tiny fraction of issued patents are ever licensed, litigated, or otherwise enforced, the herculean task of improving the quality of all patents through increased scrutiny might simply be a waste.⁴³ Along these lines, other scholars have suggested that the key to improving the patent system is not to increase examiner resources, but to better allocate them. Consequently, these commentators have proposed reforms that would help the USPTO to identify the most important patents, which would allow examiners to spend more time scrutinizing certain applications without incurring the cost of doing so for all applications.⁴⁴

All of these examination-based reforms share one thing in common: at root, they are proposals to improve the manner in which the USPTO takes in patent applications and turns out issued patents. Largely missing from the discussion, however, have been proposals to improve the inputs to that process. Indeed, with a few notable excep-

40. See Michael Abramowicz & John F. Duffy, *Ending the Patenting Monopoly*, 157 U. PA. L. REV. 1541, 1576–1600 (2009); Kesan, *supra* note 1, at 776–77; Kesan & Gallo, *supra* note 3, at 107–12; Merges, *Six Impossible Patents*, *supra* note 1, at 604–05; Thomas, *supra* note 1, at 326–32, 340–44; see also AIA § 8 (codified at 35 U.S.C. § 122 (2012)) (allowing third parties to submit prior art during examination).

41. See Farrell & Merges, *supra* note 1, at 960–64; Kesan, *supra* note 1, at 770, 773–75; Kesan & Banik, *supra* note 8, at 26, 39–41; Lichtman & Lemley, *supra* note 37, at 61–63 (proposing process for obtaining “gold-plated” patents that would require applicants to search and evaluate relevant prior art); Mann & Underweiser, *supra* note 7, at 27; Beth Simone Noveck, “Peer to Patent”: *Collective Intelligence, Open Review, and Patent Reform*, 20 HARV. J.L. & TECH. 123, 144–51 (2006); Thomas, *supra* note 1, at 340–44; Thomas, *supra* note 36, at 757. *But see* Therasense, Inc. v. Becton, Dickinson & Co., 649 F.3d 1276, 1289 (Fed. Cir. 2011) (expressing concern that a “tidal wave of disclosure” by applicants to the USPTO “makes identifying the most relevant prior art more difficult”); Cotropia, *supra* note 7, at 777–78 (cautioning that applicant’s submission of irrelevant or duplicative information might “bury” relevant material).

42. See Lemley, *supra* note 4, at 1508–10; Mark A. Lemley & Carl Shapiro, *Probabilistic Patents*, 19 J. ECON. PERSP. 75, 84 (2005).

43. See Lemley, *supra* note 4, at 1510–11, 1523–24.

44. See Lemley & Shapiro, *supra* note 42, at 84–85; Lichtman & Lemley, *supra* note 37, at 61–63; Meehan, *supra* note 1, ¶¶ 28–43; Osenga, *supra* note 1, at 141–51.

tions discussed below, the patent literature contains hardly any analyses of how patent rules affect applicant incentives in drafting and filing patent applications.

Several commentators have identified important ways in which the current patent system may encourage or discourage certain types of patent applications. For example, David Fagundes and Jonathan Masur have observed that the cost of obtaining a patent may limit the number of applications that have little private value, which may in turn reduce the number of patents with low public value that can be exploited for nuisance or hold-up.⁴⁵ Scott Kieff and others have noted that the risk of a court later invalidating a patent may cause applicants to file narrower claims than they would if the USPTO's decision were the final word.⁴⁶ And several commentators have proposed more direct incentives to reduce over-claiming, such as fee-shifting, bounties, and penalties applied during or after examination.⁴⁷ In light of the widely recognized importance of patent quality, current understandings of applicant incentives appear seriously underdeveloped.

III. APPLICANT INCENTIVES IN SELECTING CLAIM SCOPE

This Part evaluates how current patent prosecution and enforcement rules incentivize applicants to increase or reduce their risk of filing invalid claims. Although claims can be invalid on a number of legal grounds,⁴⁸ the emphasis of the following analysis is on claims that are invalid because they are overbroad. A claim can be overbroad because it describes something that was already known,⁴⁹ because it is obvious in light of prior art,⁵⁰ or because its scope exceeds what was described and enabled by the accompanying specification.⁵¹ Other problems with validity, such as indefiniteness,⁵² lack of utility,⁵³ and improper subject matter⁵⁴ are less common⁵⁵ and typically easier for

45. See Fagundes & Masur, *supra* note 9, at 696–701; Jonathan S. Masur, *Costly Screens and Patent Examination*, 2 J. LEGAL ANALYSIS 687, 706–15 (2011) [hereinafter Masur, *Costly Screens*].

46. See Kieff, *supra* note 32, at 102–04.

47. See, e.g., Kesan, *supra* note 1, at 795–97; Miller, *supra* note 1, at 704–05; Thomas, *supra* note 1, at 340–44; Thomas, *supra* note 36, at 745–46, 758 (noting partial success of Japanese program to encourage prominent filers to improve application quality voluntarily); Wagner, *supra* note 5, at 216–26; Note, *Estopping the Madness at the PTO: Improving Patent Administration Through Prosecution History Estoppel*, 116 HARV. L. REV. 2164, 2173–75 (2003).

48. See 35 U.S.C. § 282 (2012).

49. See *id.* § 102.

50. See *id.* § 103.

51. See *id.* §§ 282, 112(a).

52. See *id.* § 112(b).

53. Cf. *id.* § 101.

54. Cf. *id.*

55. See Allison & Lemley, *supra* note 11, at 208.

the USPTO to detect.⁵⁶ Moreover, by taking some of the pressure off the USPTO in the domain of overbroad claims, improvements to applicant incentives may allow the USPTO to devote more resources to these other aspects of patent quality.

To begin to approach the problem of overbroad claims, this Part focuses on patent value that derives from the ability of a patentee to successfully assert an individual patent in an infringement suit. This Part thus assumes a hypothetical applicant seeking to maximize the enforcement value of her present patent application in future litigation.⁵⁷ Several challenges to this assumption are set aside for Part V. For example, inventors file patent applications for a variety of reasons, and some may have little or no interest in licensing or enforcement; Part V.A considers other motivations for filing patents and how they may change the current Part's analyses. It is also possible that some patentees may be able to obtain settlements that diverge from expected litigation outcomes. For example, a patent holder may be able to use the threat of even long-shot litigation to deter competitive entrants, or to extort settlements based on the high cost of defending patent suits. Part V.B evaluates how deviations between settlement and litigation outcomes could affect claiming behavior. Finally, as Polk Wagner and Gideon Parchomovsky have noted, contemporary patent licensing and litigation often take place in the context of patent portfolios.⁵⁸ Part V.C evaluates how patent value derived from a patent's role in a larger patent portfolio could affect the prior analyses.

The current Part proceeds in four sections. Section A uses several basic principles of patent law to construct a model of applicants' incentives to file various kinds of claims. Section B refines that model to account for the presumption of patent validity and explains how that presumption affects applicants' incentives in drafting certain claims. Section C explores the role that examination may play in applicants' selection of claims for filing. Section D notes how several

56. Many commentators have noted the steep odds against a patent examiner trying to find the most relevant prior art. See Ethan Katsh & Beth Noveck, *Peer to Peer Meets the World of Legal Information: Encountering a New Paradigm*, 99 LAW LIBR. J. 365, 367 (2007); Lemley, *supra* note 4, at 1500; Susan Walmsley Graf, *Improving Patent Quality Through Identification of Relevant Prior Art: Approaches to Increase Information Flow to the Patent Office*, 11 LEWIS & CLARK L. REV. 495, 502–04 (2007). Rejections on the grounds of enablement are also notoriously difficult for examiners to sustain. See Sean B. Seymore, *The Presumption of Patentability*, 97 MINN. L. REV. 990, 1020–21 (2013).

57. The following analysis also assumes that applicants (or their agents) are knowledgeable about the relevant patent laws and are able to adjust their behavior accordingly. The USPTO estimates that approximately 98.2 percent of the patent applications received in 2011 were filed by patent agents or attorneys. See *E-mail from J. Hirabayashi to R. Fernando*, July 18, 2013 (correspondence on file with author). Nonetheless, applicants and practitioners may sometimes fail to fully understand the relevant doctrines. See Mark D. Janis & Timothy R. Holbrook, *Patent Law's Audience*, 97 MINN. L. REV. 72, 73–74 (2012).

58. See Gideon Parchomovsky & R. Polk Wagner, *Patent Portfolios*, 154 U. PA. L. REV. 1, 41–42 (2005).

asymmetries in patent rules may affect applicants' preferences for broader or narrower claims.

A. A Simplified Model of Claim Drafting

The patent system has two primary mechanisms for influencing behavior at the application stage: First, there are rules that affect the value an applicant can expect the claim to have if it is issued by the patent office. Second, there are rules that affect the expected cost of obtaining a given claim. This Section evaluates how these rules interact to influence the expected costs and benefits of filing any given claim.

1. Claim Value as a Function of Scope

Asserting the exclusive rights bestowed by a patent grant is no easy undertaking. After surviving potentially years of pendency before the USPTO, the patentee still faces two distinct and open questions once she gets to court: first, whether each asserted claim is in fact valid; and second, if that claim is valid, whether it actually covers the accused infringer's products or activities.⁵⁹ Losing on either front causes the infringement action to fail. Therefore, an applicant seeking to maximize the enforcement value of her patent — whether for licensing or litigation⁶⁰ — must draft her claims in a way that takes into account both the infringement value of her claims and their likelihood of validity.

No matter how careful she is in the filing and prosecution of her patent application, an applicant can never be sure that her claims are actually valid. Any prior writing, published anywhere in the world, can be used to invalidate or deny her patent.⁶¹ Even if she conducts an exhaustive search, she can never conclusively prove the non-existence of potentially invalidating prior art. Moreover, searching is costly, and many applicants find it undesirable to perform extensive searching.⁶² Accordingly, while an applicant has the freedom to draft her claims as she sees fit, she has no choice but to do so on the basis of imperfect information.

Even though a patent applicant cannot know whether the claims she files are actually valid, she does have a straightforward means to

59. See 35 U.S.C. § 282.

60. For a discussion of how settlements may depart from expected litigation outcomes, see Part V.B, *infra*.

61. See 35 U.S.C. § 102(a).

62. See Lemley, *supra* note 4, at 1509, 1510 & n.63; Iain M. Cockburn & Rebecca Henderson, *Survey Results from the 2003 Intellectual Property Owners Association Survey on Strategic Management of Intellectual Property*, at F.8 (Oct. 2003), http://www.ipo.org/wp-content/uploads/2013/04/survey_results_revised.pdf.

increase the odds that they are: she can file narrower claims. As an applicant narrows a claim by adding additional claim elements, it becomes more difficult for challengers to find prior art anticipating that claim or to find multiple references that combine to make the narrowed claim obvious. Thus, as claim scope narrows, the probability that the claim is valid increases.⁶³ Conversely, as claim scope broadens, it becomes easier for alleged infringers and others to find invalidating prior art,⁶⁴ and the likelihood of the claim surviving litigation (assuming it even issues in the first place) becomes smaller and smaller.

If validity were an applicant's only concern, deciding which claims to file would be easy. Applicants would simply draft the narrowest claims possible, and in doing so maximize the likelihood that those claims will survive any subsequent validity challenges.⁶⁵ Of course, an enforcement-minded applicant needs claims that are more than simply valid. To successfully assert her patent, she will also need to show that the defendant actually infringes one or more of her issued claims.

When it comes to maximizing infringement value, an applicant's incentives for selecting claim scope are the opposite of those for maximizing validity: broader claims have better chances of success, while narrower claims are typically worse off. The narrower the claim, the more difficult it will be to show that the accused infringer is in fact practicing each and every element of the claim as the standard for infringement requires.⁶⁶ Thus, the broader the claim, the more likely that an alleged infringer will be found to infringe it, the greater the number of potential infringers, and the more difficult it will be for

63. Invalidity based on anticipation requires that a single reference disclose all of the elements of a claim. *In re Schreiber*, 128 F.3d 1473, 1477 (Fed. Cir. 1997). The obviousness inquiry is more nuanced, but narrower claims still stand a higher chance of surviving a challenge (all else equal), because they will tend to result in greater differences between the patent claim and the prior art. *See Graham v. John Deere*, 383 U.S. 1, 17–18, 24–25 (1966). In some cases, broader claims may also raise additional difficulties under the written description or enablement requirements. *See Boston Scientific v. Johnson & Johnson*, 647 F.3d 1353, 1362–66 (Fed. Cir. 2011); *Ariad v. Eli Lilly*, 598 F.3d 1336, 1349–51 (Fed. Cir. 2010) (en banc); *Liebel-Flarsheim Co. v. Medrad, Inc.* 481 F.3d 1371, 1378–80 (Fed. Cir. 2007).

64. *See* BURTON A. AMERNICK, *PATENT LAW FOR THE NONLAWYER* 56 (1991); DAVID PRESSMAN, *NOLO'S PATENTS FOR BEGINNERS* 232 (7th ed. 2012).

65. Of course, factors other than a claim's scope affect its likelihood of validity. For example, an applicant could hire a savvy patent attorney, file her application as quickly as possible to avoid the risk that later publications will be legally deemed prior art, and pay particular attention to filing technicalities that could later serve as a basis to invalidate her patent after the fact. However, these behaviors are largely distinct from the perceived problems of patent quality and the incentives for filing broader or narrower claims. For this reason, the following analysis assumes that these various factors are held constant.

66. *See London v. Carson Pirie Scott*, 946 F.2d 1534, 1538–39 (Fed. Cir. 1991).

those potential infringers to avoid or design around the claim.⁶⁷ While the calculus for expected infringement value is complex,⁶⁸ it generalizes to a simple trend: as long as a claim remains valid, its expected infringement value increases with scope.⁶⁹ As claim scope decreases, so does the number of potential infringers, and so does the expected infringement value of the valid claim.⁷⁰

So from the perspective of applicants, narrower claims are better for purposes of validity, while broader claims are better for purposes of infringement. Yet for any single claim, an applicant cannot have it both ways — a basic principle of patent law is that claims must be construed to have the same scope for purposes of validity as they do for purposes of infringement.⁷¹ Therefore, both the likelihood of validity and the expected infringement value of a claim are functions of scope, but they tend to move in opposite directions. As scope becomes broader, likelihood of validity decreases and expected infringement value (assuming validity) increases.

67. See Nancy T. Gallini, *Patent Policy and Costly Imitation*, 23 RAND. J. ECON. 52, 62 (1992) (discussing relationship between patent scope and cost of imitation).

68. The expected infringement value of a claim will depend not only on the likelihood that any given accused product infringes that claim, but also on the value that a finding that the product infringes will have to the patentee. For example, different infringing products may have varying volumes of total sales, ease of design-around, profit margins, and so on, all of which may affect the value of a finding of infringement. See *Lucent Techs. v. Gateway, Inc.*, 580 F.3d 1301, 1323–39 (Fed. Cir. 2009) (discussing calculation of damages for infringement); *Panduit Corp. v. Stahl Bros. Fibre Works, Inc.*, 575 F.2d 1152, 1158–65 (6th Cir. 1978) (same). And even if two products have the same sales volume and strategic relationship to the patentee, the infringing feature’s role in those products could significantly change the enforcement value of the patent. See *Apple, Inc. v. Samsung Elec. Co.*, 678 F.3d 1314, 1324 (Fed. Cir. 2012) (requiring a “nexus” between the alleged infringement and lost sales in order to obtain injunctive relief). It is also possible that the expected value of a claim may depend on the relationship between that claim and other claims in an applicant’s portfolio. See *Parchomovsky & Wagner*, *supra* note 58, at 66.

69. See SUZANNE SCOTCHMER, *INNOVATION AND INCENTIVES* 103–07 (2004). An important caveat to this general trend is that some narrow claims may have high value to an applicant because they read directly on a particular product and therefore have extremely high probability of infringement in certain strategic situations, such as when an infringer has directly copied the patent holder’s product. See John R. Allison et al., *Valuable Patents*, 92 GEO. L.J. 435, 458 (2004).

70. See Masur, *Costly Screens*, *supra* note 45, at 705. To a certain extent, an applicant may be able to substitute more narrow claims for fewer broad claims, and thereby achieve the same infringement value she would have obtained by filing broad claims. However, lacking perfect foresight, it may be difficult for an applicant to conceive of every possible narrow claim necessary to approximate the coverage that could be achieved by a single broad claim.

71. See *Lemelson v. General Mills, Inc.*, 968 F.2d 1202, 1206 n.4 (Fed. Cir. 1992) (“[T]he same interpretation of a claim must be applied to all validity and infringement issues in the case.”); *W.L. Gore & Associates, Inc. v. Garlock, Inc.*, 842 F.2d 1275, 1279 (Fed. Cir. 1988) (“Having construed the claims one way for determining their validity, it is axiomatic that the claims must be construed in the same way for infringement.”).

If losses due to invalidity and losses due to non-infringement both have the same value to the applicant,⁷² a claim's expected value will be given by

$$E(s) = V(s) \times I(s), \quad (1)$$

where $E(s)$ is the expected value of a claim as a function of its scope (s), $V(s)$ is the likelihood that a claim will be found valid as a function of its scope, and $I(s)$ is the expected infringement value of a valid claim as a function of its scope.

Because the expected value of a claim depends on both its likelihood of validity and its potential to result in infringement, claim value peaks at neither extreme of claim scope. If a claim is too narrow, it is so unlikely that it will be found infringed that validity becomes irrelevant. At the other extreme, when a claim is too broad, it is so unlikely that it will be found valid that the prospect of nearly universal infringement is similarly immaterial. The enforcement value peaks somewhere in between, where the product of the claim's probability of validity and its expected infringement value is maximized.⁷³ Figure 1 illustrates this basic trend in generalized terms.

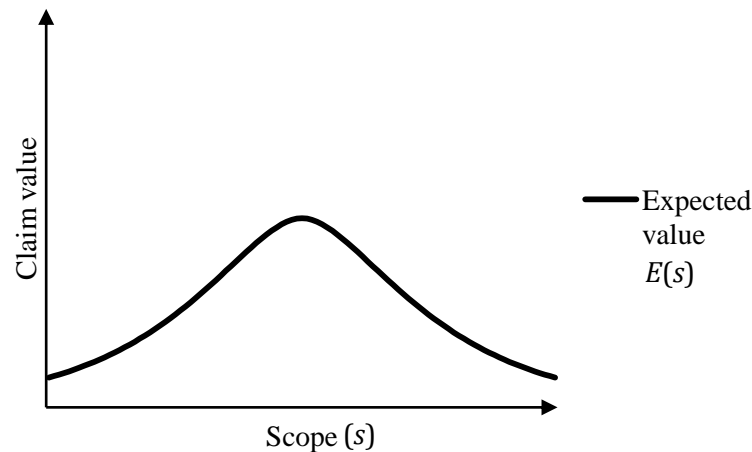


Figure 1: Expected Claim Value as a Function of Scope

72. This assumption might not always hold, however, since collateral estoppel makes a finding of non-validity more harmful to an enforcement campaign than a finding of non-infringement. For the consequences of relaxing this assumption, see Part III.D, *infra*.

73. See F. Scott Kieff, *The Case for Preferring Patent-Validity Litigation over Second-Window Review and Gold-Plated Patents: When One Size Doesn't Fit All, How Could Two Do The Trick?*, 157 U. PA. L. REV. 1937, 1961–62 (2009) (noting tension between risks of invalidity and non-infringement as a function of claim scope).

The x -axis of Figure 1 is oriented such that narrower claims are on the left and broader claims are on the right. The y -axis is oriented such that claim value increases in the positive direction.

Figure 1 depicts the general trend given in the assumptions above: claim value tends to peak at some moderate scope somewhere in the middle range of potential claim scopes.⁷⁴ At the extremes, claims that are either too narrow or too broad have their enforcement value impaired by their risk of non-infringement and invalidity, respectively.⁷⁵

These basic characteristics of claim value are widely recognized; indeed, many commentators have noted that a claim's value diminishes if it is either too broad or too narrow.⁷⁶ However, the value of claims as a function of their scope is only one part of the equation, for it does not directly answer the question most critical to improving the quality of claims through applicant incentives: given this general trend in claim value, *which* claims will applicants actually file?

2. Prosecution Cost Constraints and the Selection of Claims for Filing

At first glance, one might think that applicants will draft their claims with an eye toward hitting the peak implied above; indeed, this may be an included goal of many patent prosecution strategies. However, it would be misleading to focus exclusively on the location of the claim-value peak. After all, applicants are allowed to file a theoretically unlimited number of claims. If granted by the USPTO and asserted in court, each claim will be evaluated independently on the merits for both validity and infringement.⁷⁷ In this way, an applicant can partially hedge against the uncertainty shrouding her claim's optimal scope through claim diversification. By distributing her claims at various points along the claim-value curve, an applicant can increase her chances that one of those claims will hit near the peak. And because the quality of issued patents ultimately depends not only on the strongest claims included in a particular application but also on the weaker ones, an analysis of applicant incentives needs to consider how patent rules affect marginal claims filed as well.⁷⁸

74. None of the following analysis depends on the particular shape of the expected value function depicted by Figure 1.

75. Subject, of course, to the various assumptions stated above.

76. See, e.g., Kieff, *supra* note 32, at 102–04.

77. See *supra* note 20 and accompanying text.

78. The ability of applicants to file multiple claims introduces another complication: it is possible that there may be synergies between claims, either in the same application or across an applicant's portfolio, such that the value of controlling the full set of claims together is greater than the sum of the values of owning the claims individually. This situation might arise, for example, if three claims foreclose the three feasible ways of accomplishing some result. One limitation of the model presented in this Section is that it does not capture such interactions across claims.

Although applicants have the option of filing a theoretically unlimited number of claims, not every conceivable claim is necessarily worth filing. In reality, patent prosecution budgets are limited, and each additional claim filed will typically involve some additional filing fees, inventor time, and prosecution expenses beyond the baseline costs of preparing the patent application.⁷⁹ Given these incremental expenses, applicants will rationally seek to avoid filing claims that have an expected value that is less than the expected marginal cost to obtain them.⁸⁰ In other words, there may be a “claim-filing cutoff” — a minimum expected claim value below which an applicant will not file a claim. Figure 2 depicts one such claim-filing cutoff.⁸¹

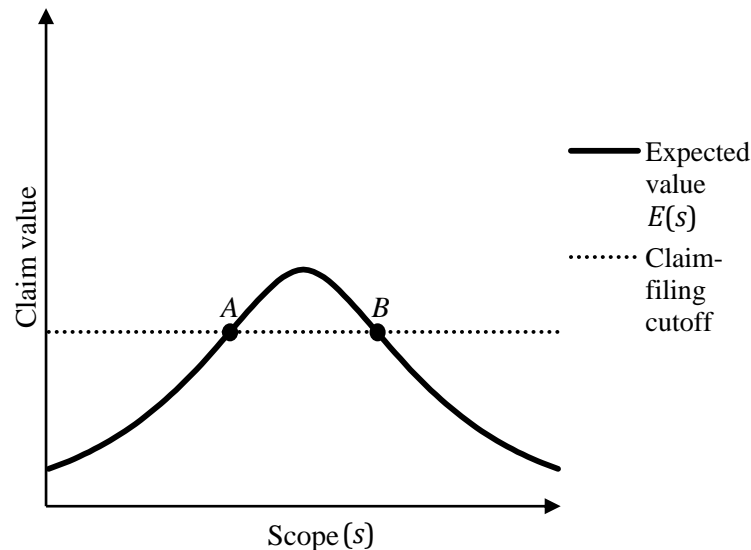


Figure 2: Expected Claim Value as a Function of Scope and a Claim-Filing Cutoff

79. The additional costs of attorney and inventor time for each claim filed are difficult to calculate at an individual level. However, it is well documented that the cost of preparing a patent application varies greatly depending on the complexity of the application. *See* AM. INTELLECTUAL PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY 27 (2013). The additional filing fees owed to the patent office for filing additional claims are discussed in detail in Part IV.A, *infra*.

80. *See* Fagundes & Masur, *supra* note 9, at 696–700; Kieff, *supra* note 32, at 102–04; Masur, *Costly Screens*, *supra* note 45, at 689–90 (arguing that one of the functions of examination costs is to select against patents having low private value); Merges, *Six Impossible Patents*, *supra* note 1, at 598 (suggesting that filing fees may discourage applicants from filing patents that are less valuable or less likely to be valid).

81. Although the claim-filing cutoff depicted in Figure 2 is flat, the cost of acquiring a claim may actually increase as a function of claim scope due to the screening function served by examination. *See infra* Part III.C.

Claims of comparatively moderate scope have an expected value above this cutoff, and so are candidates for filing. In the example illustrated in Figure 2, these net-positive-expected value claims are those with scope between points *A* and *B*.⁸² In this example, the applicant will want to avoid filing claims much narrower than those around point *A* or much broader than those around point *B*. In these outer zones, the scope of the claim renders the claim's expected value to be less than the expected cost of obtaining it, and thus the claim is not worth filing.

Although there are several ways in which the current patent system departs from this simplified model,⁸³ this basic interaction between the expected value of a claim and its expected marginal cost is the same. On average, the more claims an applicant chooses to file, the more inventor and attorney time will be necessary to prepare the application, and the greater the filing fees that will be owed to the USPTO. And although it is difficult to quantify inventor and attorney time on a per-claim basis, empirical research suggests that relatively modest changes to the filing fees imposed by the USPTO can have a measurable effect on applicant filing behavior.⁸⁴

It may be surprising that applicants would forgo filing an additional claim (potentially worth millions of dollars) to avoid incremental prosecution costs of perhaps several hundred dollars. But the potential benefit of having a valid and infringed claim must be heavily discounted to reflect the risks, uncertainty, transaction costs, and years of delay that stand between an initial patent filing and a successful litigation or settlement. In fact, given the unlikelihood of overcoming all of these obstacles, many have questioned why so many applicants file so many applications at all.⁸⁵

82. None of the following analysis depends on the particular shape of the expected value function or claim-filing cutoff depicted in Figure 2.

83. See *infra* Parts III.B, III.C, III.D, IV.A.

84. See Gaetan de Rassenfosse & Bruno van Pottelsberghe de la Potterie, *The Role of Fees in Patent Systems: Theory and Evidence*, 27 J. OF ECON. SURV. 696, 703–07 (2013); Dennis Crouch, *Evidence Based Prosecution I: Sensitivity to Claim Fee Variation*, PATENTLY-O (Sept. 21, 2006), http://patentlyo.com/patent/2006/09/patent_prosecut.html [hereinafter Crouch, *Evidence Based Prosecution: Sensitivity to Claim Fee Variation*]; Dennis Crouch, *Sensitivity to USPTO Fees*, PATENTLY-O (Oct. 1, 2008), <http://patentlyo.com/patent/2008/10/sensitivity-to.html> [hereinafter Crouch, *Sensitivity to USPTO Fees*].

85. Fewer than two percent of issued patents are asserted in court, and estimates of the percentage of patents licensed are similarly in the single digits. See Lemley, *supra* note 4, at 1501, 1507. Many prior commentators have observed and attempted to address this apparent puzzle. See, e.g., Stuart J.H. Graham & Ted Sichelman, *Why Do Start-Ups Patent?*, 23 BERKELEY TECH. L.J. 1063, 1064–70 (2008); Bronwyn H. Hall & Rosemarie Ham Ziedonis, *The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995*, 32 RAND J. ECON. 101, 102 (2001); Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 626–27 (2002); Parchomovsky & Wagner, *supra* note 58. The consequences of these various explanations are considered in Part V.A, *infra*.

To be sure, there may be some inventions that are so obviously valuable — in terms of their pure commercial value, their strategic importance, their groundbreaking novelty, or all of the above — that the applicant will file virtually any plausible claim she can think to direct at them. For applications on these types of inventions, the interaction between prosecution costs and expected benefits may be of little relevance to patent quality. However, prior work suggests that such applications are an extreme minority.⁸⁶ In the vast majority of cases, the applicant has no particular reason to believe her claims are exceptionally valuable, and so the marginal cost of additional claims may very well be enough to dissuade the applicant from filing them.⁸⁷

If applicants can at least approximately detect the outer boundaries where an additional claim is unlikely to yield much benefit,⁸⁸ there will be some zones of claim scope where applicants will try to avoid filing claims. One of these regions will be in an area that is perhaps of little interest to patent policymakers — the area where claims are too narrow to be of any use (or concern) to anyone. But the other region — where claims are too *broad* to be worthwhile — should be particularly intriguing to people interested in improving patent quality. Applicants will not typically file such broad claims, based entirely on their own self-interested prospects for enforcement.⁸⁹ If these claims are rarely filed in the first place, the public will rarely bear any cost of examining, opposing, or litigating them.

Of course, the fact some claims are too broad to be worth filing does not imply that this outer boundary is currently set in the right place or even that it secures any benefit not already provided by ex-

86. After all, estimates of the percentage of patents that produce *any* economic value hover in the single digits, *see* Lemley, *supra* note 4, at 1501, 1507, and patents that produce exceptional value that is identifiable at the application stage would necessarily be a subset of those. *See also* Allison et al., *supra* note 69, at 461 n.111 (reporting conversation with a general counsel who estimated that approximately three percent of his company's applications are so valuable that "the sky's the limit" in prosecution fees).

87. *See* de Rassenfossé & van Pottelsberghe de la Potterie, *supra* note 84, at 703–07.

88. And there is good reason to believe that they can. At a minimum, an applicant can avoid filing claims that either read directly on known prior art or are patently indistinguishable from the prior art as a matter of law. *See* Cotropia, *supra* note 7, at 760; Kesan & Banik, *supra* note 8, at 52–53. Beyond these methods, an applicant can perform a basic prior art search or literature review, if not to exhaustively rule out prior art then at least to estimate the plausibility of her broadest claims. A more advanced applicant can carefully draft her claims to avoid limitations that could easily be designed around, that are not necessary to the functioning of the invention, or that limit the claims without adding any chance of further distinguishing the invention from the prior art. And every applicant can tailor her aggressiveness to her patent prosecution budget and the expected role of the given application in her larger patent strategy. All of these techniques (with the possible exception of a proactive prior art search) are basic procedures for patent agents and attorneys. *See* RONALD D. SLUSKY, *INVENTION ANALYSIS AND CLAIMING* 27–28 (2d ed. 2012); Cotropia, *supra* note 7, at 757–62, 760 n.186.

89. Again, this proposition assumes that the enforcement value of a claim is correlated with the risk-adjusted litigation value of that claim. The effect of relaxing this assumption is discussed in Part V.B, *infra*.

amination. However, recognizing the possibility of such a region and understanding how patent rules affect it are the first steps toward leveraging this boundary to improve patent quality.

B. The Presumption of Patent Validity

Under existing law, the question of infringement and the question of validity are given substantially different treatment in litigation. On the infringement side, the plaintiff carries the burden of showing that the defendant is infringing her claim by a preponderance of the evidence.⁹⁰ Invalidity, by contrast, is an affirmative defense.⁹¹ And, under the longstanding presumption of patent validity, a defendant carries the burden of showing that the plaintiff's claim is invalid by "clear and convincing evidence."⁹² This burden is higher than a mere preponderance — making it easier, all else being equal, for claims to survive validity challenges than if there were no presumption.⁹³

Though the law regarding the presumption of patent validity is clear, the magnitude of its effect in practice has proved difficult to measure.⁹⁴ In the absence of further empirical research regarding the presumption of validity, several basic assumptions are useful to this Section's analysis.

1. Assumptions About the Presumption of Validity

In order to evaluate how the presumption of patent validity might affect claiming behavior, this Section makes three assumptions about how the presumption operates in practice. The first assumption is that the presumption of validity does not end up doing any harm to the patentee's cause. This assumption actually consists of two parts: First, no claim has its likelihood of validity reduced by the presumption. Second, and similarly, changing the standard for validity does not

90. *SmithKline Diagnostics, Inc. v. Helena Labs. Corp.*, 859 F.2d 878, 889 (Fed. Cir. 1988). Although non-infringement is listed as a "defense" in the patent code, 35 U.S.C. § 282(b)(2) (2012), it is really an argument against the plaintiff's *prima facie* case, not an affirmative defense on which the defendant must carry the burden of proof, see Roger Allan Ford, *Patent Invalidity Versus Noninfringement*, 99 CORNELL L. REV. 71, 73 n.5 (2013) (explaining this distinction).

91. 35 U.S.C. § 282(a) (2012).

92. See *Microsoft Corp. v. i4i Ltd. P'ship*, 131 S. Ct. 2238, 2242–43 (2011).

93. Recent statutory reforms may allow a challenger to proceed under the lower preponderance of the evidence standard through administrative challenges like post-grant or inter partes review. See AIA §§ 6(a), (d) (respectively codified at 35 U.S.C. §§ 316(e), 326(e) (2012)). These approaches present their own set of costs and benefits for challengers and patentees. For a thorough discussion of these procedures, see Gregory Dolin, *Dubious Patent Reform*, 56 B.C. L. REV. (forthcoming 2015).

94. See Lichtman & Lemley, *supra* note 37, at 69. For an excellent empirical study on the effect of the presumption on mock jurors, see David L. Schwartz & Christopher B. Seaman, *Standards of Proof in Civil Litigation: An Experiment from Patent Law*, 26 HARV. J.L. & TECH. 429 (2013).

reduce the odds that the claim will be found infringed. These two parts are both intuitive and correct as a matter of law. Still, it is always possible that idiosyncrasies of the jury system could produce counterintuitive (or counter-legal) results.

The second assumption is that at least some claims have their likelihood of validity improved by the presumption. This assumption would be true if, for example, a claim that would have been found valid fifty percent of the time without the presumption has a better than fifty percent chance of being found valid with the presumption. There may be a similar effect, say, for claims that previously stood a forty percent chance or a sixty percent chance of being found valid. However, none of the following analysis depends on any particular distribution of the presumption's effects. Like the first assumption, this assumption is both intuitive and correct as a matter of law. If jury instructions matter, then changing a "preponderance of the evidence" standard to a "clear and convincing evidence" standard should have some effect in cases where validity otherwise would have been uncertain.⁹⁵

The third assumption is that the presumption of patent validity is rebuttable. In other words, even with the presumption of patent validity, if a claim is too broad, it runs some non-trivial risk of being found invalid. Like the first two assumptions, this assumption is correct as a matter of law.⁹⁶ Moreover, there is direct evidence that this assumption is true — even with the presumption of validity, claims are frequently found to be invalid.⁹⁷ Although the presumption may increase the likelihood that some claims will be found valid, it hardly renders issued patents immune to validity challenges.

To summarize these assumptions: the presumption of patent validity increases the probability of validity for some claims, leaves it no worse for others, and does not change the fact that the probability of validity must decline at some point when claim scope becomes excessively broad. With these assumptions in place, it is possible to model the effects that the presumption of patent validity has on applicants' incentives when drafting claims.

95. Supporting this assumption, there is at least experimental evidence that changing the standard of proof for validity can affect the decisions of mock jurors. See Schwartz & Seaman, *supra* note 94, at 459.

96. See 35 U.S.C. § 282 (2012).

97. See Allison & Lemley, *supra* note 11, at 205; Mann & Underweiser, *supra* note 7, at 7 (approximately sixty percent of patents found invalid in Federal Circuit decisions between 2003 and 2009).

2. The Effects of the Presumption of Validity

If the assumptions discussed above are true,⁹⁸ the basic properties of the probability of validity as a function of claim scope remain the same with or without the presumption: the probability still starts high for narrow claims and falls as claim scope increases. However, the presumption of validity causes the probability of validity to remain higher for more claims as claim scope increases. Figure 3 illustrates examples of the probability of validity both with and without a legal presumption of validity.

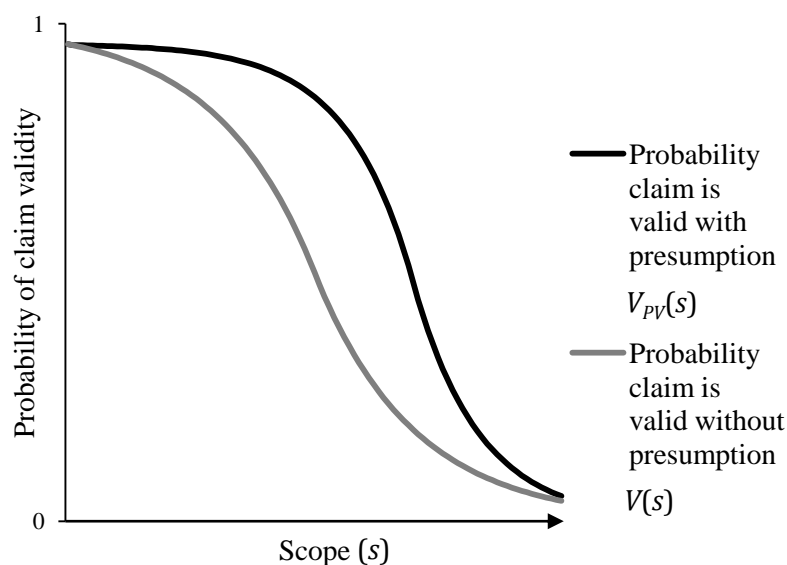


Figure 3: Probability of Claim Validity as a Function of Scope

The x -axis of Figure 3 is configured such that narrower claims are on the left and broader claims are on the right. The y -axis ranges from a probability of zero to a probability of one.

As Figure 3 illustrates, when the claim in question is extremely narrow, the probability of it being found valid is high, with or without the presumption.⁹⁹ As the claim gets broader, both probabilities decrease. But, because of the presumption of validity, the probability that the claim will be found valid is higher for some claims than it would be without the presumption. Nonetheless, when the claim in

98. See *supra* Part III.A.1.

99. None of the following analysis depends on the particular shape of the probability curves depicted by Figure 3.

question is extremely broad, the probability of it being found valid is low — again, either with or without the presumption.

Assuming that a loss due to invalidity and a loss due to non-infringement both have the same value to the patentee,¹⁰⁰ a claim's expected value with the presumption of validity will be given by

$$E_{PV}(s) = V_{PV}(s) \times I(s), \quad (2)$$

where $E_{PV}(s)$ is the expected value of a claim entitled to a presumption of validity as a function of its scope (s), $V_{PV}(s)$ is the likelihood that a claim will be found valid given the presumption of validity as a function of its scope, and $I(s)$ is, as defined previously, the expected infringement value of a valid claim as a function of its scope.

Equation (2) for expected claim value with a presumption of validity should look familiar — it is structurally the same as Equation (1) for the expected value of a claim without a presumption of validity. Indeed, under the assumptions made above, the basic features of the expected value remain the same. As before, the expected value of an issued claim peaks at neither extreme of claim scope. If a claim is too narrow, its value is low because of the unlikelihood that it will ever be infringed. On the other hand, when a claim is too broad — so broad that the presumption of patent validity cannot salvage its overreach — the likelihood of it being valid is so low that the patent similarly has little enforcement value. Despite the presumption of validity, the expected value of a claim still falls off as the claim becomes either too narrow or too broad.

So what effect does the presumption actually have? Although the basic shape of the expected value curve remains the same, the presumption of validity can potentially change the expected value at a variety of points along the way.

To evaluate how the presumption affects claiming, consider the following relationship between the presumption's effects on likelihood of validity and its effects on expected claim value. Let $\Delta V(s)$ represent the improvement in a claim's likelihood of validity as a result of the presumption of validity as a function of its scope, such that

$$V_{PV}(s) = V(s) + \Delta V(s). \quad (3)$$

Substituting Equation (3) into Equation (2), we obtain

$$E_{PV}(s) = [V(s) + \Delta V(s)] \times I(s) \quad (4)$$

$$\equiv E_{PV}(s) = [V(s) \times I(s)] + [\Delta V(s) \times I(s)], \quad (5)$$

and substituting Equation (1) into Equation (5), we obtain

$$E_{PV}(s) = E(s) + [\Delta V(s) \times I(s)]. \quad (6)$$

100. For the consequences of relaxing this assumption, see Part III.D, *infra*.

As a result, the improvement in a claim's expected value is given by $[\Delta V(s) \times I(s)]$ — the product of the improvement in the claim's likelihood of validity multiplied by the claim's expected infringement value. In other words, the increase in a claim's value as a result of the presumption depends on *both* the improvement the presumption makes to the claim's likelihood of validity *and* the infringement value of the claim.

It may be obvious that the presumption of validity only improves the expected value of claims whose likelihood of validity is actually improved by the presumption, that is, where $\Delta V(s)$ in the equation above is greater than zero. The perhaps more subtle feature of the presumption of validity is that the degree to which it helps a given claim also depends on the expected infringement value of that claim, $I(s)$. Even significant improvements in a claim's likelihood of validity — a large $\Delta V(s)$ — will therefore be of little value to the applicant if that claim has small infringement value.

The influence of infringement value gives the presumption of validity an inherent bias toward broader claims. Provided the infringement value of a claim tends to increase as a function of claim scope,¹⁰¹ the presumption of validity will tend to offer greater improvements in expected value to broader claims than to narrower ones.

To illustrate this effect, suppose the presumption helps all claims uniformly on the question of validity, adding ten percent to the chance that any given claim will be found valid (provided that the probability can never be greater than one).¹⁰² When this change in likelihood of validity is multiplied by infringement value, the benefits of the presumption fall largely to broader claims. Figure 4 illustrates an example of expected claim value with and without a presumption of validity, assuming the presumption yields a constant improvement to the likelihood of validity across all claim scopes.

101. See *supra* Part III.A (justifying this assumption).

102. This example merely illustrates the tendency of infringement value to skew the presumption's benefits toward broader claims. There is no reason to believe that the presumption of validity has this particular distribution of benefits to claims, nor is the resulting analysis dependent on this choice of example.

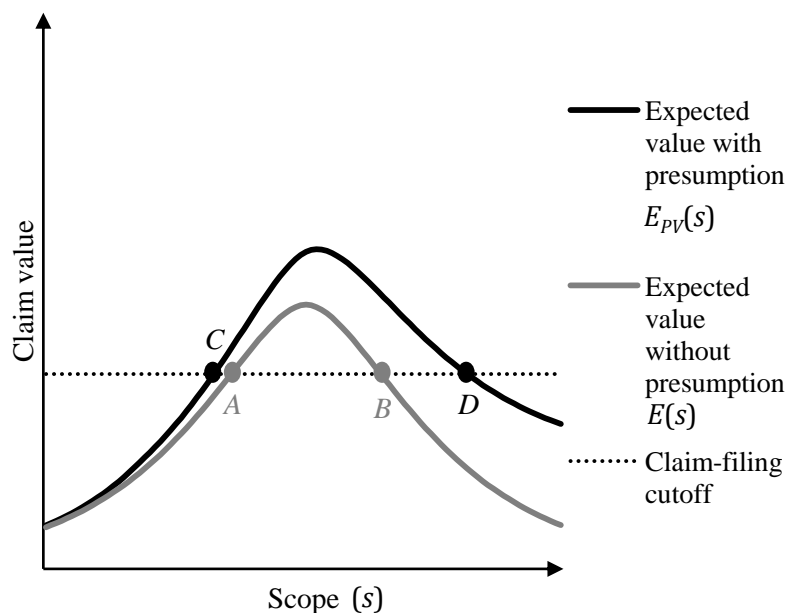


Figure 4: Expected Claim Value With and Without Presumption of Validity

As Figure 4 illustrates, even a presumption of validity that affects the validity of all claims equally has a larger effect on broader claims. When $\Delta V(s)$ is constant, the improvement in expected value becomes a direct product of the claim's infringement value. The broader the claim, the greater the infringement value and the greater the benefits afforded by the presumption of validity.¹⁰³

This improvement in expected value may cause some claims that would not have been worth filing without the presumption of validity to become profitable as a result of the presumption. In the example illustrated in Figure 4, without the presumption of validity an applicant would not want to file claims much broader than point *B* — approximately where the expected value of the claim crosses the claim-filing cutoff. But once the presumption of validity is in play, it makes sense for this same applicant to file claims as broad as point *D*. As a result of the presumption of validity, some broad claims that would not have been worth filing without the presumption of validity may become valuable enough to file.¹⁰⁴

103. Subject again to the caveats and assumptions stated in Part III.A, *supra*.

104. In some cases, narrower claims, too, may become worth filing as a result of the presumption. As Figure 4 illustrates, for example, without the presumption an applicant would

Note that the claims that become worth filing because of the presumption are not necessarily the claims getting the largest benefit from the presumption. The increase in expected value changes filing behavior only to the extent that it lifts a claim from below the claim-filing cutoff to above the claim-filing cutoff. Indeed, in the example illustrated in Figure 4, many of the overbroad claims getting the largest expected value improvement from the presumption are still not worth filing. The presumption of validity shifts, but does not eliminate, the outer boundary of claims with expected value exceeding their expected costs.

Figure 4 highlights the significance of infringement value when assessing the effects of the presumption of validity. But remember, the increase in a claim's expected value as a result of the presumption of validity depends on *both* the improvement in the likelihood of validity ($\Delta V(s)$) *and* the expected infringement value ($I(s)$) of the claim. Figure 4 is based on the somewhat conservative assumption that $\Delta V(s)$ is constant as a function of scope. But this assumption might not always be true. To the contrary, some commentators have predicted that $\Delta V(s)$ will be largest for the kind of broad claims whose probability is already in doubt.¹⁰⁵ After all, if a claim is almost certainly valid, there may be very little work left for the presumption to do. If this intuition is correct, it suggests that the presumption's benefits for expected value will be even more dramatically skewed to favor broader claims than the simple example illustrated in Figure 4 suggests. The claims most benefited by the presumption of validity will be those with a real question of validity *and* that cover a large and significant amount of potentially infringing commerce — almost certainly broader than claims that lack those attributes, and perhaps at high risk of being the very kind of overbroad claims that reformers are seeking to eliminate.

3. Assessing the Presumption of Validity

Two primary effects on applicants' incentives follow from the changes to the expected value of claims predicted above. The first effect is that applicants can expect more reward from participating in the patent system — more inventions may justify the costs of filing applications, more claims may be worth filing as a part of those patent applications, and the net expected value of those claims may be greater than they otherwise would be. Since one purpose of the patent sys-

not want to file claims narrower than point A. With the presumption, claims as narrow as point C have expected values that exceed the claim-filing cutoff.

105. See, e.g., Lichtman & Lemley, *supra* note 37, at 58 (arguing that the presumption disproportionately helps patents whose validity would otherwise be in doubt).

tem is to reward invention,¹⁰⁶ all these effects are generally considered positive — subject, of course, to some limit. Holding all else constant, future applicants can be expected to file more claims and to obtain more benefit for doing so.¹⁰⁷

The second effect is an increase in applicants' incentives to file broader claims. The claims with the greatest improvement in expected value will tend to be broader, and will very rarely be narrower. The increased likelihood that these broader claims will be found valid — coupled with the unchanged likelihood that they will be found infringed — creates pressure for an applicant to file broader claims than she would have in the absence of the presumption.

However, these conclusions are based on an evaluation of the presumption of validity in isolation. Because the presumption shifts, but does not eliminate the outer boundary of claims worth filing, it is possible that other features of the patent system may offset (or may be adjusted to offset) the potential for increased incentives to file broader claims that is created by the presumption of validity. The next Section evaluates the effects that patent examination may have on counterbalancing the presumption of validity.

C. Substantive Patent Examination

In the preceding Sections, the cost of filing a claim was modeled as being constant across all claim scopes. But it is possible that substantive patent examination may raise the expected cost of obtaining certain weaker claims. If broader claims, for example, tend to be rejected more frequently than narrower ones, higher expected prosecution costs could reduce the appeal of filing such broad claims in the first place.

106. See WARD S. BOWMAN JR., PATENT AND ANTITRUST LAW: A LEGAL AND ECONOMIC APPRAISAL 2–3 (1973); F.M. SCHERER & DAVID ROSS, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 621–24 (3d ed. 1990). Because rewarding invention is currently the dominant justification given for the patent system, the following discussion is framed around rewards. Similar trade-offs could be made with an eye to other functions of the patent system. See, e.g., Stephen Yelderman, Coordination-Focused Patent Policy (Oct. 27, 2014) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2481025 (evaluating the implications for the optimal breadth and reliability of patent rights under a coordination-focused patent system).

107. This result is intuitive. Under the assumptions above, patent law was made unambiguously better for patentees, so it should come as no surprise that the value of seeking and obtaining patents increased. In practice, of course, creating (or strengthening) a presumption of patent validity may be accompanied by imposing more rigorous examination, higher filing fees, or other additional costs that may balance out the incentive to file more patents. The effect of such prosecution costs is evaluated in Part III.C, *infra*.

1. The Effect of Substantive Examination on Expected Costs

If substantive patent examination functions as even a crude screen against overbroad claims — that is, if the likelihood of a claim getting rejected increases with claim breadth — then the expected cost of obtaining a claim will likewise increase as claim scope broadens.¹⁰⁸ Although an initial rejection is not a final determination of a claim's validity, it may nonetheless serve an important function by increasing the applicant's costs.¹⁰⁹

More rejections could result in higher costs in one of several ways. First, upon receiving a rejection notice, the applicant may undertake a potentially long and expensive battle with the USPTO in order to secure the claim as it was originally filed. In such a case, the cost of obtaining the broad claim will be substantially greater than the cost of obtaining a narrower claim that would have been granted right away.¹¹⁰ Alternatively, upon receiving a rejection notice, the applicant may undertake that same long and expensive battle with the USPTO and ultimately lose. In that case, the cost of obtaining the overbroad claim would be essentially infinite — no amount of investment in prosecution could achieve its issuance.

Applicants have other options when a claim is rejected, but they too are costly. For example, an applicant can narrow a rejected claim by amendment, potentially conceding that the claim was too broad in the first instance and retreating to a fallback position. This response exacts a penalty in claim scope — both by the reduction in the literal meaning of the claim and by the reduction of protection against equivalents as a result of prosecution history estoppel.¹¹¹ She can cancel the claim entirely, which imposes the cost of not having the claim at all. Alternatively, she can file a new claim, potentially even one that is broader. But filing a new claim imposes additional prosecution costs, and the applicant has no guarantee that the new claim will get approved either. In sum, when rejection strikes, it is costly regardless of the path the applicant chooses in response.

108. Cf. Masur, *Costly Screens*, *supra* note 45, at 700 (noting that prosecution costs increase as likelihood of invalidity increases); Fagundes & Masur, *supra* note 9, at 699–700 (same).

109. Cf. Lemley, *supra* note 4, at 1527 (noting that examination, despite its limitations, may serve a useful purpose by discouraging applicants from filing overbroad applications).

110. The cost of triggering protracted argument with the USPTO can be significant. In 2012, for example, the median reported attorneys' fees for a relatively complex reply to an office action ranged from \$ 2500 to \$ 3000. The median reported attorneys' fees for taking an appeal to the Patent Trial and Appeal Board ranged from \$ 4000 to \$ 7000. AM. INTELLECTUAL PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY 27 (2013).

111. See *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., Ltd.*, 535 U.S. 722, 733–34 (2002) (“The doctrine of equivalents allows the patentee to claim those insubstantial alterations that were not captured in drafting the original patent claim but which could be created through trivial changes. [After narrowing the claim in response to rejection,] he may not argue that the surrendered territory comprised unforeseen subject matter.”).

Applicants face expected prosecution costs that potentially increase with claim scope in a series of chunky steps. An additional claim element may be enough to earn an applicant a prompt allowance, or it might make the difference between persuading a patent examiner to allow the claim in an interview rather than requiring a lengthy appeal. In advance, however, applicants do not typically know precisely which claim expansions will trigger such a cost step.¹¹² Instead, they know only that, in some range of claim scopes, the expected cost to obtain a claim will tend to increase as a result of broader claiming. This effect, in turn, may raise the expected value that a claim must have in order for the applicant to believe it worth filing.

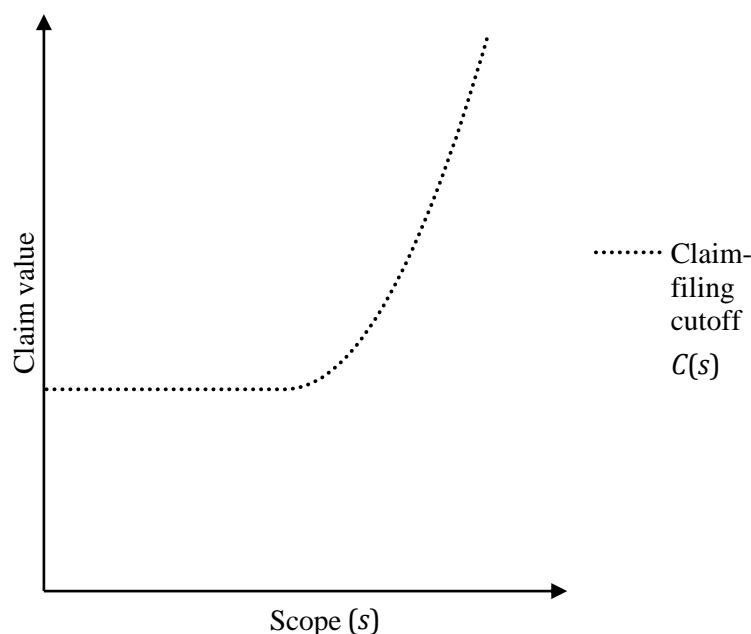


Figure 5: Claim-Filing Cutoff as a Function of Scope

Figure 5 depicts one example of how the minimum expected value for a claim to be worth filing might change as a function of its scope. Within some range of narrow claims, the likelihood of an initial allowance is high, and thus the applicant's expected prosecution costs are very near to the basic filing fees and drafting expenses. As

¹¹² There are, of course, trivial examples in which an applicant would know exactly where her prosecution costs are expected to jump — for example, when she drafts a claim that reads directly on prior art of which she is aware.

the scope of the claim gets broader, the likelihood of triggering a longer and more expensive prosecution increases as well. In the extreme, a very broad claim may have essentially infinite expected prosecution costs, since no amount of lobbying, argument, or appeal will be sufficient to obtain it. Thus, some broad claims that would have been worth filing with flat prosecution costs will fall below a claim-filing cutoff that increases with claim scope — they are not worth filing due to the increased expected prosecution costs imposed by substantive examination.

2. Counterbalancing the Presumption of Patent Validity

In the example illustrated in Figure 4 above, the presumption of patent validity increases the expected value of certain broad claims. Holding prosecution costs constant, this increase in expected value can cause some claims to be worth filing that would not have been worth filing without the presumption. In the example illustrated in Figure 4, adding a presumption of validity caused an applicant to shift from filing her broadest claim near point *B*, to filing her broadest claim near point *D* (these points are reproduced in Figure 6, below).

But suppose that the introduction of the presumption of validity is accompanied by the advent of substantive patent examination, which in turn causes the expected cost to obtain a claim to increase with claim scope. If this patent screening function were implemented properly, it could counterbalance applicants' incentives to seek broad claims as a result of the presumption of validity.

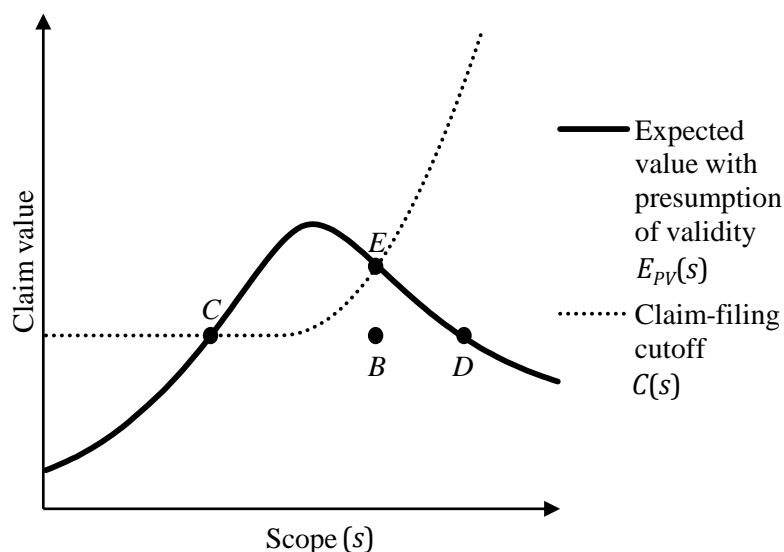


Figure 6: Expected Claim Value and Claim-Filing Cutoff as Functions of Scope

As in Figure 4, the presumption of validity increases the expected value of certain broad claims. But by adding substantive patent examination, the claim-filing cutoff for those claims is higher too. When substantive patent examination is implemented correctly, it can counterbalance the increased value of those claims as a result of the presumption of validity. In the example illustrated in Figure 6, the applicant would not want to file claims broader than point *E* — creating incentives to file her broadest claim at approximately the same scope as if there were neither substantive examination nor a presumption of validity in force.¹¹³

But the presumption of patent validity and the costly screening function of patent examination do not leave applicant incentives exactly where they started. The combination of a presumption of validity and an examination cost screen can increase the expected value of claims of moderate scope. For example, claims that have intermediate likelihoods of validity may receive some enforcement benefit from the presumption of validity, yet avoid triggering the higher prosecution costs that claims with very low likelihoods of validity would incur under a substantive examination system. If coordinated correctly, the

113. However, note that the exact relationship between the expected value of a claim and the claim-filing cutoff will depend on the value of the invention, the inventive step involved, and other factors. See *supra* Part III.A. The example depicted in Figure 6 is merely for illustration; there is no reason to believe that, in practice, substantive examination and the presumption of validity will perfectly balance their effects on the broadest claim worth filing.

combination of substantive examination and a presumption of validity may thus be able to increase total inventor rewards without pushing applicants toward broader claiming overall.

3. The Benefits of a Substantive Examination System

This interplay between the presumption of validity and prosecution costs that increase with claim scope is a potential reason for preferring a patent system with substantive examination and deferential post-grant review to a simple registration regime. In a registration system, patent applications are reviewed only for technical compliance with filing requirements, and not for the appropriateness of their claims.¹¹⁴ Without any substantive review of applicants' claims, the expected cost to obtain a claim will be constant as a function of claim scope. And typically, the unexamined claims that would be issued under a registration system would not be entitled to a presumption of their validity.¹¹⁵

An examination system offers a discrete set of costs and benefits, both public and private. Applicants must endure a costly examination process, but in return are afforded a presumption of validity for the claims that survive that process. That presumption may increase applicants' incentives to file broader claims, but ideally the costly screen of examination counterbalances that pressure. The net effect of this interplay is that a patent system operating under an examination model may be able to deliver greater rewards to patentees than would be available under a simple registration system, without increasing incentives for applicants to file broader claims.

This synergy between costly substantive examination and the presumption of validity could serve as an alternative to the standard, agency deference justification for the post-examination validity presumption.¹¹⁶ However, it is important to note that a substantive exam-

114. See Kieff, *supra* note 32, at 70–74 (evaluating comparative strengths and weaknesses of registration and examination systems); Lemley, *supra* note 4, at 1526–27. For a colorful account of an experiment of the young United States with a simple registration regime, see Edward C. Walterscheid, *The Winged Gudgeon — An Early Patent Controversy*, 79 J. PAT. & TRADEMARK OFF. SOC'Y 533 (1997).

115. See Keith Aoki, *Distributive and Syncretic Motives in Intellectual Property Law (with Special Reference to Coercion, Agency, and Development)*, 40 U.C. DAVIS L. REV. 717, 740 & n.68 (2007); Michael J. Harbers, *International Patent Cooperation*, 20 STAN. L. REV. 1000, 1003 (1968).

116. See *Catholic Pro. Serv. v. Am. Smelting & Ref. Co.*, 594 F.2d 499, 505 (5th Cir. 1979) (noting that the presumption of validity “is founded on the understanding that patent approval is a species of administrative determination supported by evidence”); *Chi. Rawhide Mfg. Co. v. Crane Packing Co.*, 523 F.2d 452, 458 (7th Cir. 1975) (relating the presumption of validity to “the deference due to the technical expertise possessed by the Patent Office and not generally possessed by federal judges”). Note that the presumption as it currently exists *cannot* be entirely explained by agency deference, as the presumption applies even in cases where the agency did not have the proper evidence before it. See *Microsoft Corp. v. i4i Ltd. P'ship*, 131 S. Ct. 2238, 2249–51 (2011).

ination process and a presumption of validity can each theoretically exist without the other.¹¹⁷ From the perspective of applicant incentives, each is a tool that could be used to adjust the expected cost and value of an unfiled patent claim and therefore influence the kind of claims applicants file.

D. Other Asymmetries in the Treatment of Validity and Infringement

The presumption of validity is a rather dramatic example of how the rules for determining validity may depart from the rules for determining infringement. However, there are other, more subtle differences in the treatment of these questions. Specifically, collateral estoppel principles may have different practical consequences for losses on invalidity versus losses on non-infringement, and rules for claim amendments may make it easier to narrow a claim after the fact than to broaden it.

1. Re-Litigation of Validity and Infringement

In the preceding Sections, it was assumed that patentees were indifferent as between losing an enforcement suit on the basis of invalidity and losing on the basis of non-infringement. Both outcomes were given a zero value, which is accurate when determining the value of a single round of patent litigation. Under U.S. law, however, judgments of invalidity and judgments of non-infringement often have different preclusive effects for future litigation, and so could be expected to have different values as outcomes to patentees contemplating multiple enforcement actions against a series of defendants.

On one level, invalidity and non-infringement judgments have the same preclusive effects. Collateral estoppel is a doctrine of general application — it merely prevents a party from re-litigating a question that was fully explored and finally determined against it in a prior litigation.¹¹⁸ If a question was not actually decided, or if the party prejudiced by the court's answer was not party to the prior litigation, collateral estoppel does not apply.¹¹⁹

117. See Adam Mossoff, *Who Cares What Thomas Jefferson Thought About Patents? Reevaluating the Patent "Privilege" in Historical Context*, 92 CORNELL L. REV. 953, 998–1000 (2007).

118. *Smith v. Bayer*, 131 S. Ct. 2368, 2379 (2011) (“A court’s judgment binds only the parties to a suit, subject to a handful of discrete and limited exceptions.”); 18A CHARLES ALAN WRIGHT & ARTHUR R. MILLER, *FEDERAL PRACTICE AND PROCEDURE* § 4448 (“The central proposition that a party is bound is balanced by the rule that ordinarily nonparties are not bound. Some substantial justification must be found to justify preclusion of a nonparty.”).

119. *Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found.*, 402 U.S. 313, 329 (1971); *Hansberry v. Lee*, 311 U.S. 32, 40 (1940).

However, in the context of a typical patent litigation, collateral estoppel has drastically different consequences for adjudications of validity and infringement. Collateral estoppel makes invalidity judgments particularly bad for patentees on enforcement campaigns because the same validity question is likely to be raised in each patent lawsuit. If the patentee loses on validity even once, her ability to enforce that claim is over, since any subsequent defendant will be able to invoke collateral estoppel.¹²⁰ And when the patentee *wins* on validity, that judgment in her favor will have little preclusive effect on future challengers, each of whom will have her own distinct opportunity to argue invalidity.¹²¹

The effects of collateral estoppel on determinations of infringement are more balanced. Because the infringement determination for each successive defendant will often involve a different set of factual issues, there will be many cases where collateral estoppel does not even come into question. When it does, collateral estoppel will only work to the slight advantage of defendants. As with validity, it will never work against a newcomer defendant, since she must be given one full and fair opportunity to advance her arguments. In the other direction, when a prior non-infringement ruling is invoked against a patentee, she will often have an opportunity to attempt to distinguish the factual questions decided in the earlier case from those necessary to dispose of the current one. A prior non-infringement ruling can be helpful to a defendant, but it hardly provides the same assurance as a prior invalidity ruling.¹²²

These asymmetric potential outcomes are summarized in Table 1 below.

120. *Blonder-Tongue*, 402 U.S. at 332–34, 350; *Mendenhall v. Barber-Greene Co.*, 26 F.3d 1573, 1577–78 (Fed. Cir. 1994). Moreover, losing a validity challenge will even void prior licenses and thus terminate existing royalty streams. See Michael Risch, *Patent Challenges and Royalty Inflation*, 85 IND. L.J. 1003, 1024–25 (2010).

121. *Blonder-Tongue*, 402 U.S. at 329; *Mendenhall v. Cedarapids, Inc.*, 5 F.3d 1557, 1569 (Fed. Cir. 1993). The prior ruling may be given weight, see *Stevenson v. Sears, Roebuck & Co.*, 713 F.2d 705, 711 (Fed. Cir. 1983), but a later determination that the claims are invalid is entirely possible, see, e.g., *Cedarapids*, 5 F.3d at 1569–72; *Stevenson*, 713 F.2d at 712. The earlier validity judgment “serves only to inform the . . . court . . . that caution must be taken in reaching a contrary legal conclusion.” *Cedarapids*, 5 F.3d at 1569. However, as others have noted, non-mutual issue preclusion introduces the possibility of freeriding on others’ efforts to invalidate a troublesome patent, and may make it more difficult to organize such efforts. See Miller, *supra* note 1, at 674.

122. See Miller, *supra* note 1, at 729–30. However, it is possible for findings that were necessary to a non-infringement verdict to be used against a patentee in subsequent litigation, at least in certain situations. See *Pfaff v. Wells Elecs., Inc.*, 5 F.3d 514, 518 (Fed. Cir. 1993); *Molinaro v. Fannon/Courier Corp.*, 745 F.2d 651, 655 (Fed. Cir. 1984). The strongest case, for example, would be one involving two identical products. If the first product has been found not to infringe the patent, the patentee may be estopped from asserting that the second (identical) product infringes the patent. See, e.g., *Aspex Eyewear, Inc. v. Zenni Optical Inc.*, 713 F.3d 1377, 1380–82 (Fed. Cir. 2013); *Molinaro*, 745 F.2d at 655.

Table 1: Effect of Previous Judgments

| <i>Claim Previously Found:</i> | <i>Effect for Patentee in Later Litigation:</i> |
|--------------------------------|---|
| Infringed | None / weakly positive |
| Not infringed | None / potentially negative |
| Valid | Weakly positive |
| Invalid | Fatal |

This asymmetry may have the effect of pushing applicants toward narrower, more conservative claims. Because findings related to infringement are often of limited value (either offensively or defensively) for future litigations, they put comparatively less at stake for patentees. Findings of invalidity, by contrast, have the potential to destroy all of a claim's value. When there are multiple potential infringers of a patent, the owner of that patent will anticipate at least the possibility of needing to bring multiple infringement suits. As a result, she may prefer — all else being equal — to increase the risk of losing on infringement in exchange for better odds of winning on validity.¹²³

However, patentees have an important procedural countermeasure: joinder. When a court joins infringement claims against multiple defendants into a single suit, it may make common pretrial rulings on claim scope and validity, while leaving individualized infringement determinations for each accused product or process for trial.¹²⁴ In this way, a patentee can obtain multiple infringement determinations while risking only a single determination of validity. At the limit, aggressive joinder of accused infringers could blunt the caution-inducing effects of collateral estoppel, moving patentees toward a state of indifference as to the reason for any given loss.

For reasons largely aside from claim scope, joinder of patent defendants is already a matter of controversy.¹²⁵ Until recently, there was a split among district courts on what standard applied to the permissive joinder of defendants in patent suits, with some courts presumptively allowing the joinder of unrelated defendants alleged to infringe the same patent, and other courts denying joinder unless there was some relationship among the various defendants' allegedly infringing acts.¹²⁶ In response, the AIA preempted the ordinary joinder

123. However, the costs and benefits of each potential form of loss are complex. For example, even though the typical patent holder has more to lose on the validity front, there are a variety of reasons that defendants may be reluctant to bring validity challenges, *see infra* notes 211–12, which may push applicants back in the direction of broader claims.

124. *See, e.g., In re Cruciferous Sprout Litigation*, 301 F.3d 1343, 1346 (Fed. Cir. 2002).

125. David O. Taylor, *Patent Misjoinder*, 88 N.Y.U. L. REV. 652, 689–92 (2013).

126. *See also* Taylor, *supra* note 125, at 678–89 (summarizing the split). *Compare* MyMail, Ltd. v. Am. Online, Inc., 223 F.R.D. 455, 456 (E.D. Tex. 2004), *and* Sprint Commc'ns Co. v. Theglobe.com, Inc., 233 F.R.D. 615, 617 (D. Kan. 2006), *with* Androphy v. Smith & Nephew, Inc., 31 F. Supp. 2d 620, 623 (N.D. Ill. 1998), *Philips Elecs. N. Am.*

rules applicable to accused infringers, making it more difficult for patentees to join defendants into a single infringement suit.¹²⁷ Although it does not appear that improving patent application quality was a motivation for this particular provision,¹²⁸ to the extent it increases the frequency with which patentees must withstand repeated validity challenges, it may increase incentives for applicants to claim more conservatively.¹²⁹

2. Procedural Preferences for Narrowing Amendments

The model presented above describes the incentives facing applicants as they file and prosecute their initial claims in the USPTO. However, there are also a variety of procedural options that allow for the addition and amendment of claims after a patent has initially issued. These post-grant procedures are consistently more lenient toward narrowing (as opposed to enlarging) amendments, which may add value to broader claiming in the first instance.

The rules for reissuing a patent provide a stark example. At any time during the life of a patent, the patent's owner may apply for reissuance to correct the mistake of having "claim[ed] more or less than he had a right to claim in the patent."¹³⁰ This latitude, however, is subject to an important caveat. Reissues to enlarge the scope of claims may be filed only in the first two years after a patent is granted.¹³¹ Reissues to reduce the scope of claims may be filed at any time.¹³²

This preference for narrowing claim amendments is rooted in the conventional, adversarial view of patent prosecution, in which it is presumed that applicants will seek to obtain the broadest claims possible while the USPTO tries to push them back.¹³³ Under this view, a patentee's expansion of her patent's scope will potentially cause public harms, while her reduction of the scope will presumably lead to

Corp. v. Contec Corp., 220 F.R.D. 415, 417 (D. Del. 2004), and *Rudd v. Lux Prods. Corp.*, No. 09-cv-6957, 2011 WL 148052, *2-*3 (N.D. Ill. Jan. 12, 2011).

127. See AIA § 19(d) (codified at 35 U.S.C. § 299 (2012)); Taylor, *supra* note 125, at 692–706. The Federal Circuit has since weighed in on the standard for joinder in pre-AIA suits as well. See *In re EMC Corp.*, 677 F.3d 1351, 1359–60 (Fed. Cir. 2012).

128. See Taylor, *supra* note 125, at 700–07 (summarizing legislative history).

129. Cf. Tracie L. Bryant, Note, *The America Invents Act: Slaying Patent Trolls, Limiting Joinder*, 25 HARV. J.L. & TECH. 673, 691 (2012) (observing that a stricter joinder rule will require more caution in litigation on the part of patentees).

130. 35 U.S.C. § 251(a) (2012).

131. *Id.* § 251(d).

132. See *id.*

133. See *Miller v. Brass Co.*, 104 U.S. 350, 354 (1881) (noting that when Congress passed the original reissue provision, "[i]t was probably supposed that the patentee would never err in claiming too little"). The soft "reasonableness" time limit on enlarging reissues set out in *Miller* was replaced by a strict two-year cutoff in the Patent Act of 1952. See Act of July 19, 1952, Pub. L. No. 82-593, 66 Stat. 792, 808 (codified as amended at 35 U.S.C. § 251(d) (2012)).

public benefits.¹³⁴ So the former option is time-limited, while the latter is preserved.

But from the view of applicant incentives in the USPTO, this procedural asymmetry increases the value of obtaining broader claims in the first place. The broader the claim, the greater its option value as a vehicle for future narrowing amendments. In this way, an overly broad claim is better than an overly narrow claim, since the former enables a later reissue to get scope right while the latter does not. And other vehicles for post-grant modification of claim scope are similarly biased across the board in favor of narrowing amendments.¹³⁵

While procedural preferences for post-grant narrowing amendments may create some incentives to seek broader claims in the first instance, it is important to note that patentees have other procedural tools to achieve the same result. In the case of particularly valuable inventions, for example, it is common practice to file a series of continuation applications to preserve the option of filing additional claims in the future.¹³⁶ The incentive effects of asymmetric amendment rules may therefore be blunted by the option of simply filing a continuation. However, if reforms are implemented to reduce this use of continuation practice,¹³⁷ applicants may respond by seeking broader claims in the first instance as an alternative means to preserve their claim modification options.

Based on these understandings of how existing patent rules affect incentives to file particular kinds of claims, the next Part proposes several reforms that could be expected to push applicants in the direction of filing higher quality claims.

IV. REFORMS FOR IMPROVING PATENT QUALITY

Before proceeding, it is important to note that improving the quality of the claims filed in the USPTO is not a goal to be pursued at all costs. The purpose of the patent system, after all, is to foster innovation and encourage inventive activity, not simply to ensure applicants do not obtain invalid claims. If it were the latter, solving the patent quality crisis would be easy. Policymakers could simply discourage patent filing altogether — drastically increasing filing fees, vitiating the infringement rules to disfavor patentees, and so on. As the benefits

134. See *Miller*, 104 U.S. at 355–56.

135. See 35 U.S.C. § 305 (2012) (only narrowing amendments permitted during ex parte reexamination); *id.* § 316(d)(3) (same for inter partes review); *id.* § 326(d)(3) (same for post-grant review).

136. Allison et al., *supra* note 69, at 458; Lemley & Moore, *supra* note 24, at 68.

137. See, e.g., Lemley & Moore, *supra* note 24, at 105–18 (suggesting reforms that would reduce, but not abolish, the use of continuation applications); Gary C. Ganzi, *Patent Continuation Practice and Public Notice: Can They Coexist?*, 89 J. PAT. & TRADEMARK OFF. SOC'Y 545, 600–03 (2007) (recommending reforms to continuation practice).

of filing applications decrease, the number of such applications (good or bad) could be expected to decrease as well. In the extreme, one certain way to end the current crisis in patent quality would be simply to repeal the patent statutes altogether.

The true challenge of improving patent quality is to do so in a way that preserves the rewards available for obtaining high-quality patents while reducing the allure of low-quality ones. An analysis of proposals to change applicants' incentives to file broad claims, then, should also take into account how those proposals will affect the costs and benefits of participating in the patent system in the first place.

Two policy assumptions inform the following discussion. The first policy assumption is that the patent system would be improved if patents tended to issue with narrower claims. The second policy assumption is that the rewards available to inventors should be preserved at or near their current level. In other words, the goal of these reforms is to encourage narrower claiming while leaving the total incentives to file patent applications unaffected.

Of course, these particular policy goals are neither universal nor inevitable. How much reward to provide to inventors and where on the spectrum of claim breadth that reward should be provided are both complex questions at the very core of patent policy.¹³⁸ But because analysis of these larger questions is beyond the scope of this Article, these goals are taken as assumed objectives. In any event, the following analysis of how patent rules could be used to reduce incentives to file broad claims while maintaining the total level of inventor rewards under the patent system is instructive for designing policies with different goals as well.

This Part proceeds in four Sections. Sections A and B relate to ways to improve the cost-screening function performed by filing fees and examination, respectively. Section C proposes a change to patent litigation rules that could encourage applicants to file narrower claims. Section D examines the viability of using post-examination penalties to affect applicant claim-filing behavior.

A. Rationalizing Application Fees

As discussed above,¹³⁹ the cost of filing a claim can serve as an important backstop that may screen against some of the lowest-value claims. If claims were free, an applicant could stake a claim at nearly every point of the claim value curve, hedging against miscalculation and ensuring that she files at least one claim at or near the peak.

138. See DAN L. BURK & MARK A. LEMLEY, *THE PATENT CRISIS AND HOW THE COURTS CAN SOLVE IT* 80–92 (2009) (discussing theories underlying the patent system); Merges & Nelson, *supra* note 10, at 873–74.

139. See *supra* Part III.A.2.

Without some cost constraint, there would be no reason to avoid filing extremely broad or extremely narrow claims, since there is always a chance that one of those claims might turn out to be valuable.

An important component of the cost of filing claims is the fees paid to the USPTO.¹⁴⁰ The basic filing fee allows an applicant to file a limited number of claims; after reaching that limit, each additional claim costs between \$80 and \$500 in additional fees.¹⁴¹ Although those numbers may seem small, the cost of filing additional claims can add up quickly. For example, an application may contain several embodiments of the same invention, and the applicant may seek to differentiate those embodiments from the prior art on multiple dimensions. These individually modest surcharges for extra claims, multiplied by several embodiments and multiple attempts to differentiate from the prior art, can become significant in the aggregate.¹⁴² Thus, while any single failed claim may cost the applicant only a trivial amount of marginal filing fees, a practice of filing excessive claims can dramatically increase the applicant's costs with few offsetting benefits. And, indeed, prior empirical work has observed a relationship between the magnitude of USPTO fees and the behavior of patent applicants.¹⁴³

One way to increase the screening effect of filing fees would be to raise them across the board. Such a change could be expected to have two distinct benefits: First, it would cause applicants to choose their claims more carefully, thereby reducing the number of invalid or

140. Other components of the cost of filing a claim include inventor time and attorneys' fees. However, it is difficult to predict how these costs relate to the quantity or the scope of claims filed, or how changes to patent policy could change those relationships. Therefore, the following analysis focuses on the fees paid to the USPTO for each claim filed.

141. The basic filing fee includes the right to file up to twenty claims. Each additional claim incurs a charge of \$ 80. 37 C.F.R. § 1.16(i) (2013). Moreover, the basic filing fee permits the applicant to file just three independent claims. Each independent claim in excess of three incurs a charge of \$ 420. *Id.* § 1.16(h). These charges are additive, so an applicant who files her twenty-first claim that is her fourth independent claim faces \$ 500 in additional fees. *Id.* §§ 1.16(h)–(i). Discounts ranging from fifty percent to seventy-five percent are available for small entities, universities, and independent inventors. *See id.* §§ 1.16, 1.27, 1.29 (defining categories of applicants eligible for fee reduction). For ease of exposition, the following discussion refers to the rates an applicant would pay assuming she does not qualify for these discounts.

142. Consider an applicant who pursues a strategy of filing four additional independent and eight additional dependent claims beyond the threshold of claims that are worth filing. In applications in which she has already used all of her free claims, she owes \$ 500 for each of her excess independent claims and \$ 80 for each of her excess dependent claims. The resultant increase in her filing fees — \$ 2640 — amounts to a more than twenty percent increase in prosecution costs in the context of an \$ 11,000 patent application. *See supra* note 9. For a corporation filing hundreds of applications a year, the cost of this strategy could be significant.

143. *See Crouch, Evidence Based Prosecution: Sensitivity to Claim Fee Variation, supra* note 84; Crouch, *Sensitivity to USPTO Fees, supra* note 84 (reporting fifteen percent drop in total number of claims filed after the USPTO increased filing fees for each additional claim over twenty claims from \$ 18 to \$ 50); de Rassenfosse & van Pottelsberghe de la Potterie, *supra* note 84, at 703–07.

unimportant claims cluttering the technological landscape. Second, provided the USPTO retains the funds it collects from applicants,¹⁴⁴ increased filing fees would allow the office to invest more resources in examination and therefore make fewer mistakes.

But this change would come with two downsides as well: First, higher application fees would discourage filing of both broad and narrow claims, and may on the whole actually push applicants in the direction of more aggressive claims.¹⁴⁵ Second, increased fees would reduce the net expected value of *all* claims — even the strongest, most valuable claims that will be filed despite the higher fees. Without some offsetting benefit to applicants, the expected rewards for invention are decreased, as may be overall participation in the patent system. Increasing the cost of filing each additional claim may be beneficial within some range, but it is also possible to increase fees to a point that is counterproductive to the goals of the patent system.¹⁴⁶

An alternative way to increase the screening effect of filing fees is to ensure that that screen does not have any holes. And here, the current fee schedule leaves clear room for improvement. Presently, the \$1600 basic filing fee includes the right to file up to twenty claims at no additional charge.¹⁴⁷ Not only is there no surcharge for claims two through twenty, but also there is no benefit conferred for filing any fewer than twenty claims. As a result, there may be some cases where an applicant files additional claims even though she does not expect them to have any significant value whatsoever.¹⁴⁸

Another non-linearity in the current fee schedule flows from the distinction between independent and dependent claims.¹⁴⁹ The current schedule imposes only a modest charge for dependent claims, while

144. In principle, the USPTO is entitled to keep the fees it collects, *see* 35 U.S.C. § 42 (2012), although history suggests that this financial independence may be periodically weakened during times of government austerity. *See* Michael D. Frakes & Melissa F. Wasserman, *Does Agency Funding Affect Decisionmaking?: An Empirical Assessment of the PTO's Granting Patterns*, 66 VAND. L. REV. 67, 77–78 (2013).

145. This scenario could occur, for example, if increased application fees predominantly pushed narrower claims below the claim-filing cutoff, while broader claims remained above the cutoff.

146. *See* Thomas, *supra* note 36, at 743 (noting that increasing costs of patent filing may push inventors toward the alternative of trade secret protection and reduce the rate of innovation); Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2165.

147. *See* 37 C.F.R. § 1.16 (2013) (establishing “basic fee for filing” of \$ 280, “search fee” of \$ 600, and “examination fee” of \$ 720, as well as an additional surcharge for claims in excess of twenty).

148. *See* Crouch, *Evidence Based Prosecution: Sensitivity to Claim Fee Variation*, *supra* note 84 (observing filing pattern that suggests that applicants tend to file all twenty claims included with the basic filing fee). The applicant may, however, incur additional agent or attorneys’ fees by filing each additional claim.

149. A dependent claim is one that incorporates by reference all of the limitations of another claim. *See* 35 U.S.C. § 112 (2012); U.S. PATENT & TRADEMARK OFFICE, MANUAL OF PATENT EXAMINING PROCEDURE § 608.01(n), at 600-129 to -137 (2014), <http://www.uspto.gov/web/offices/pac/mpep/mpep-0600.pdf>. Therefore, a dependent claim is always at least as narrow as the claim that it incorporates.

exacting a significant premium for an application containing more than three independent claims. After an applicant has used up all her claims included in the basic filing fee, an additional dependent claim is just \$80, compared to \$500 for an additional independent claim.¹⁵⁰ In some range, then, the marginal filing fee can be more than six times greater for some claims than it is for others.

Because every claim must either be independent or ultimately depend on an independent claim, and because independent claims are expensive, applicants have an incentive to reuse independent claims as many times as possible. This fee structure encourages applicants to draft broad, top-level independent claims to use as a base, and then to branch various dependent claim families off that base.¹⁵¹ The independent claim premium is so high that it often makes sense to use this technique even if it means filing an independent claim that would not be cost-justified on its own.

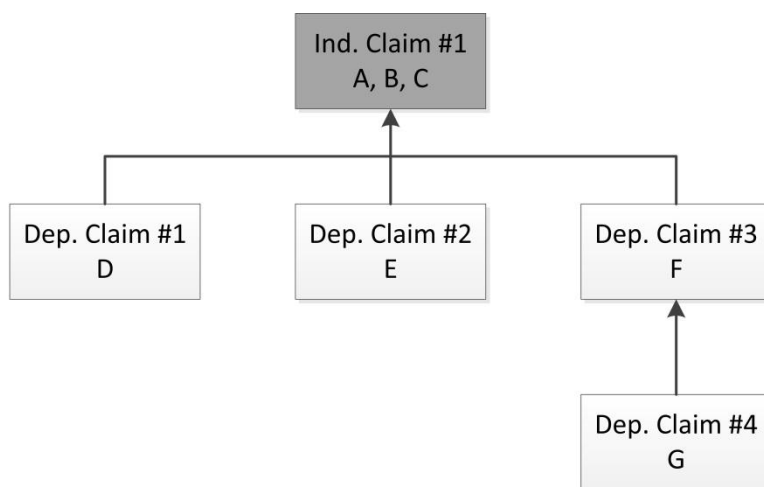


Figure 7: Claim Organization Minimizing Applicant's Fees

Figure 7 illustrates one way a patent applicant could organize her claims to take advantage of the discount for dependent claims. Independent Claim #1 (containing claim elements A, B, and C) rests at the top of the claim family tree. Three dependent claims depend directly from Independent Claim #1: Dependent Claim #1 (incorporating claim elements A, B, and C from Independent Claim #1 and adding element D); Dependent Claim #2 (incorporating claim elements A, B, and C from Independent Claim #1 and adding element E); and De-

¹⁵⁰. See *supra* note 141.

¹⁵¹. This technique is well known and commonly practiced among patent prosecutors. See, e.g., DAVID PRESSMAN, PATENT IT YOURSELF 240, 245 (15th ed. 2011).

pendent Claim #3 (incorporating claim elements A, B, and C from Independent Claim #1 and adding element F). Another dependent claim, Dependent Claim #4, depends on Dependent Claim #3 (incorporating claim elements A, B, and C from Independent Claim #1 and element F from Dependent Claim #3, and adding element G).

In the example illustrated in Figure 7, Independent Claim #1 (shaded gray) is too broad to be worth filing on its own.¹⁵² The various dependent claims (shaded white), however, are sufficiently narrow to be worth filing. With the approach illustrated in Figure 7, the applicant has one independent claim and four dependent claims. Under the current fee schedule, adding these claims to an application would cost \$820.¹⁵³

A competing claim drafting strategy is illustrated in Figure 8.

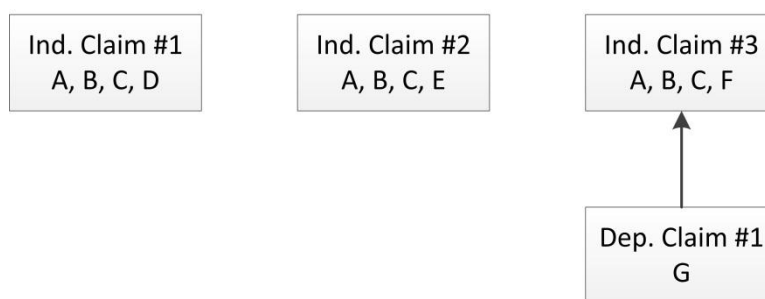


Figure 8: Claim Organization Without Fee Minimization

Figure 8 illustrates a different way the same applicant could organize her claims. Rather than relying on an overly broad independent claim as a base for the claims she would actually like to file, the applicant has split her claims into separate families of claims that are individually worth filing. Three independent claims now stand alone: Independent Claim #1 (containing claim elements A, B, C, and D); Independent Claim #2 (containing claim elements A, B, C, and E); and Independent Claim #3 (containing claim elements A, B, C, and

152. That is, the claim's expected value is so low that it would not be worth even the price of filing as a dependent claim if not for its function as a base for other dependent claims.

153. These examples illustrate the case of an applicant who has already used the three independent and twenty total claims included with the basic filing fee. The additional independent claim results in a charge of \$ 500, and each dependent claim results in a charge of \$ 80, for a total of \$ 820 in additional fees. *See supra* note 141.

F). A single dependent claim, Dependent Claim #1, depends on Independent Claim #3 (incorporating elements A, B, C, and F from Independent Claim #3 while adding element G).

In the example illustrated in Figure 8, all of the claims (shaded white) are worth filing on their own merits. The applicant has not filed any overly broad independent claims to use as a base for filing other claims. With the approach illustrated in Figure 8, the applicant has three independent claims and one dependent claim. Under the current fee schedule, adding these claims to an application would cost \$ 1580.¹⁵⁴

A comparison of the examples illustrated in Figures 7 and 8 reveals the perverse incentives potentially created by the current fee schedule. In the approach illustrated in Figure 7, the applicant has filed one more claim than she did in the example illustrated in Figure 8. And, importantly, the additional claim in the Figure 7 is overbroad and would not be worth filing if it had to stand on its own merits. Still, because of the extreme cost differences in filing independent versus dependent claims (nearly half-price fees for this aspect of her application), the applicant has a direct financial incentive to include the overbroad claim.¹⁵⁵

Both of these features of the fee schedule — the discount for dependent claims and the right to file up to twenty claims without incurring any additional charge per claim — appear to be the product of previous attempts to align filing fees with the burden that examining an application will put on the USPTO.¹⁵⁶ Because of overhead costs such as reviewing other sections of an application, identifying and retrieving the most likely sources of prior art, and other administrative tasks, it may very well take an examiner about as long to evaluate an application with several claims as it does to evaluate one with only a single claim. For their part, chains of dependent claims are easier to review because they impose a certain logical hierarchy on the application. An examiner can direct her search at the narrowest claim in a dependent claim branch. If she can identify prior art that makes that

154. See *supra* note 141.

155. Whether written as an independent or dependent claim, each claim is a legally distinct invention that may present independent questions of validity and infringement. See *supra* note 20 and accompanying text.

156. Indeed, the current framework of filing fees appears to have been implemented with examination burden in mind. When the rule that an applicant could include three independent and twenty dependent claims at no additional charge was first enacted in 1982, the House report noted that the bill included “provisions for increasing the filing fees due to increased complexities presented by certain applications, e.g., applications containing more than a specified number of claims.” H.R. REP. NO. 97-542, at 5–6 (1982), *reprinted in* 1982 U.S.C.C.A.N. 765, 769.

claim unpatentable, she can reject all the broader claims in the branch on the same basis.¹⁵⁷

While the current fee schedule may get applicant incentives right to facilitate cost-effective examination, it is inconsistent with using filing fees as a screen for improving patent quality. After all, each and every claim, whether dependent or independent, whether the only claim in an application or one of many, stands on its own for purposes of both validity and infringement.¹⁵⁸ The public interest in discouraging applicants from filing overly broad claims is the same for all claims, regardless of how they happen to fit into the structure of the larger application.

The solution to both of these problems is straightforward: impose a flat fee for each claim an applicant files, regardless of how that claim relates to other claims or the total number of claims in an application. Rather than charging a large upfront fee for filing an application with a number of claims included at no extra charge, and rather than charging a premium for independent claims, the USPTO could charge a lower upfront fee and a more significant, flat fee for each claim filed. The per-claim fee could be designed to leave the current average total filing fees paid by applicants approximately unchanged. For example, if a typical application today includes thirty claims, five of which are independent (total filing fee of \$3240),¹⁵⁹ an upfront fee of \$500 plus \$90 per claim would leave the average total filing fees nearly unchanged (total filing fee of \$3200).¹⁶⁰ Because each claim would cost exactly the same amount — without regard to the total number of claims in an application or whether the claim is dependent or independent — this reformed fee structure would render filing fees a more consistent cost screen against low-value claims. And notably, under the AIA, the USPTO now enjoys the authority to rebalance this fee schedule on its own; however, so far, the USPTO has used this authority to exaggerate these discrepancies rather than repair them.¹⁶¹

To be sure, there may be important examination efficiencies achieved by having applicants organize their claims into hierarchical

157. However, it is not clear that examiners actually give dependent claims such scrutiny. See Michael Risch, *Failure of Public Notice in Patent Prosecution*, 21 HARV. J.L. & TECH. 179, 197 (2007).

158. See *supra* note 20 and accompanying text.

159. In this example, the applicant must pay the basic filing fee (\$ 1600), see *supra* note 147 and accompanying text, plus \$ 80 for each of the ten claims in excess of twenty (\$ 800) and \$ 420 for each of the two independent claims in excess of three (\$ 840), see *supra* note 141.

160. This example merely illustrates the reformed fee structure's effect on one hypothetical application.

161. See AIA § 10. In its initial round of fee setting, the USPTO increased the basic filing fee by twenty-seven percent, the cost for each independent claim by sixty-eight percent, and the cost for each claim in excess of twenty by twenty-nine percent. See Setting and Adjusting Patent Fees, 78 Fed. Reg. 4212, 4224–25 (Jan. 18, 2013) (to be codified at 37 C.F.R. pts. 1, 41–42).

families based on one or more independent claims. Without some incentive to draft dependent claims, applicants might draft only independent ones, potentially increasing the work that must be done by patent examiners. Therefore, there may be an argument for preserving the distinction in filing fees between dependent and independent claims, though perhaps at less dramatic levels. A smaller discount for dependent claims might be enough to encourage applicants to organize their claims into families when appropriate, but without inducing as many overbroad base claims.

B. Penalizing Amendments and Abandonment

One way to discourage applicants from filing overbroad claims is to make broader claims more expensive. If the examination system causes the expected cost of obtaining a claim to increase with claim scope, broader claims will become less attractive, while narrower claims will become comparatively more appealing.¹⁶²

Under current prosecution rules, the penalty for having a claim rejected is small. Although rejection *can* lead to increased costs through the process of argument and appeal, those costs will only be incurred if the applicant persists in seeking the rejected claim. Applicants always have the option of simply cancelling any rejected claims and proceeding with the allowed ones. The maximum penalty imposed, then, is the loss of any filing and prosecution costs already incurred.¹⁶³

One way to increase the expected cost of filing overbroad claims would be to require applicants to post a completion bond for each claim filed. This bond would be returned to the applicant when the claim issues. It would be forfeited, however, when the claim is cancelled or amended.¹⁶⁴ The effect of this bond would be to make it more expensive to cancel or amend a claim, thereby increasing the expected cost of filing broad claims. The cost of a granted claim — either upon first examination or after some process of argument or appeal — would remain unchanged. By raising the cost of filing a

162. See *supra* Part III.C.

163. However, cancelled claims may impose a penalty in the form of prosecution history estoppel. See *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., Ltd.*, 535 U.S. 722, 735–41 (2002); *Honeywell Int'l Inc. v. Hamilton Sundstrand Corp.*, 370 F.3d 1131, 1134 (Fed. Cir. 2004) (en banc). However, because *Festo* estoppel only affects a patentee who ultimately asserts the doctrine of equivalents, see 535 U.S. at 735–41, the deterrent effects of this estoppel are likely small. For a proposal to deter overbroad claiming by increasing the force of prosecution history estoppel, see Wagner, *supra* note 5, at 216–26.

164. Given the pressure to file quickly under the new first-to-file system, see AIA § 3 (codified in scattered sections of 35 U.S.C.), it might make sense to offer an initial grace period before the completion bond penalties kick in. For example, applicants could be allowed to cancel or amend their claims without forfeiting their bonds during the first year of their applications' pendency, or perhaps until substantive examination of the claims has begun. The benefits of this flexibility for applicants would need to be weighed against the costs imposed by its additional administrative complexity.

claim that should rightfully be rejected, a completion bond requirement could be expected to push applicants in the direction of filing higher quality claims.¹⁶⁵

Another potential consequence of such a bond might be that applicants will persist in arguing for debatable claims longer than they would in the absence of a bond. As long as there is a chance of persuading the examiner (or the Patent Trial and Appeal Board) of the claim's merits, the expected cost of pursuing such arguments would be reduced by the possibility of earning back the bond. Therefore, in addition to producing higher quality claim filings, a completion bond requirement might also produce more protracted argument and review of claims closest to the line of patentability.¹⁶⁶

While more frequent argument and appeals may strike some as expensive and undesirable, they may actually be an additional benefit of a completion bond requirement. Reduced incentives to file broad claims means fewer examiner resources spent challenging plainly unpatentable claims. More frequent arguments and appeals regarding rejected claims means more intense scrutiny of those claims that are close to the line. Indeed, a common call for patent reform involves triaging the work of the USPTO, providing different levels of review depending on the nature of the claim.¹⁶⁷ A completion bond requirement — which could lead to higher quality claim filings from the beginning and more careful review of those that are more questionable — could be one mechanism for organically transitioning to such a system.

C. Replacing the Presumption of Validity with a Presumption of Infringement

As discussed above,¹⁶⁸ the presumption of validity has two primary effects on applicant incentives: First, it increases the rewards inventors can expect to obtain by participating in the patent system. Second, it provides those rewards at least in part by encouraging applicants to file broader claims than they otherwise would.

One way to improve the quality of claims filed would be to simply repeal the presumption of validity. This repeal could be expected to have mixed effects. On the one hand, it would reduce the expected

165. Other scholars, too, have proposed penalties for claim amendments. See Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2167–68. This concept is in some ways similar to the examination-stage bounty proposed by John R. Thomas. See Thomas, *supra* note 1, at 340–44.

166. Cf. Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2167.

167. See, e.g., Lemley, *supra* note 5, at 10–11; Lemley & Shapiro, *supra* note 42, at 84–85; Lichtman & Lemley, *supra* note 37, at 61–63; Meehan, *supra* note 1, ¶¶ 28–43; Osenga, *supra* note 1, at 141–51.

168. See *supra* Part III.B.

value of broader claims. On the other hand, it would also reduce the appeal of participating in the patent system altogether. Under the policy goals assumed above, the former would be a benefit, the latter an unintended consequence.

However, there is another alternative. Rather than simple repeal, the presumption of validity could be replaced with a presumption of infringement. Such a presumption of infringement could have two distinct benefits: First, it could increase the value of narrow claims as compared to broad ones. Second, it could make up some of the inventor rewards lost from repealing the presumption of validity.

A presumption of patent infringement would be similar to the presumption of validity, in that it would shift the burden of proof on one of the central questions arising in a typical patent litigation. Rather than carrying the burden to prove invalidity, patent defendants would carry the burden to prove their own non-infringement. Of course, infringement is not (and would not be) evaluated by the USPTO in advance.¹⁶⁹ In order to prevent potential abuse, patentees would likely need to make some threshold showing (for example, a “substantial likelihood”¹⁷⁰) that a specific product infringes a specific claim to trigger the presumption.¹⁷¹ Upon this showing, non-infringement would become an affirmative defense like invalidity is today, to be proven by the accused infringer by clear and convincing evidence. The presumption of infringement would be rebuttable, and defendants might be able to regularly overcome this presumption, just as current defendants regularly overcome the presumption of validity.¹⁷²

So what effect would a presumption of infringement have on applicant incentives? Assuming that losses due to invalidity and losses due to non-infringement are of equal value,¹⁷³ a claim’s expected value will now be given by

$$E_{PI}(s) = V(s) \times I_{PI}(s), \quad (7)$$

169. See U.S. PATENT & TRADEMARK OFFICE, MANUAL OF PATENT EXAMINING PROCEDURE § 701, at 700-6 (2014), <http://www.uspto.gov/web/offices/pac/mpep/mpep-0700.pdf> (noting that patents are examined for validity under 35 U.S.C. §§ 101, 102, 103, and 112).

170. Cf. 35 U.S.C. § 295 (2012).

171. This preliminary showing could also address a concern recently articulated by the Supreme Court that a patent holder is in a better position to point out, at least initially, which product infringes which claim of the patent. See *Medtronic, Inc. v. Mirowski Family Ventures, LLC*, 134 S. Ct. 843 (2014) (declining to shift burden of infringement in declaratory judgment context).

172. For the following analysis, it is assumed that the presumption of infringement could be implemented in a way such that it (1) increases the likelihood of infringement for at least some claims and does not decrease the likelihood of infringement for any claims, (2) does not affect the likelihood (one way or the other) that any claim will be found valid, and (3) does not prevent at least some claims from being found non-infringed despite the presumption. Cf. *supra* Part III.B.1.

173. See *supra* Part III.D.

where $E_{PI}(s)$ is the expected value of a claim entitled to a presumption of infringement as a function of its scope (s), $V(s)$ is the likelihood that a claim will be found valid as a function of its scope, and $I_{PI}(s)$ is the infringement value of a valid claim entitled to a presumption of infringement as a function of its scope.

Like the presumption of validity, the presumption of infringement changes the expected value of a variety of claims.¹⁷⁴ Let $\Delta I(s)$ represent the increase in a valid claim's infringement value as a result of the presumption of infringement as a function of its scope, such that

$$I_{PI}(s) = I(s) + \Delta I(s). \quad (8)$$

Substituting Equation (8) into Equation (7), we obtain

$$E_{PI}(s) = V(s) \times [I(s) + \Delta I(s)] \quad (9)$$

$$\equiv E_{PI}(s) = [V(s) \times I(s)] + [V(s) \times \Delta I(s)], \quad (10)$$

and substituting Equation (1) into Equation (10), we obtain

$$E_{PI}(s) = E(s) + [V(s) \times \Delta I(s)]. \quad (11)$$

As a result, the improvement in a claim's expected value is given by $[V(s) \times \Delta I(s)]$ — the product of the claim's likelihood of validity multiplied by the increase in the claim's infringement value. In other words, the increase in a claim's value as a result of the presumption of infringement would depend on both the improvement that the presumption makes to the claim's infringement value and the claim's existing probability of validity.

Like the presumption of validity, a presumption of infringement could be expected to increase the total expected benefits available to an inventor participating in the patent system. But the presumption of infringement benefits a different kind of claim than does the presumption of validity. The claims that are most benefited by a presumption of infringement will have two characteristics: First, there will be some doubt as to whether they are infringed (that is, $\Delta I(s)$ is large). Second, they will be narrow enough to have a significant likelihood of being valid (that is, $V(s)$ is large). While it may be difficult to predict how the improvement to infringement value will be distributed, claims with high probabilities of validity are by definition high quality claims.¹⁷⁵ So while the presumption of validity tends to skew its benefits toward broader claims of dubious validity,¹⁷⁶ a presumption of infringement favors higher quality claims.¹⁷⁷ If applicants care about

174. The logic here is the converse of that described in Part III.B.2, *supra*.

175. See *supra* note 33 and accompanying text.

176. See Figure 4 and accompanying text.

177. This argument is a generalization and is admittedly vulnerable to counterexamples. For example, there may be certain broad claims with large improvements to infringement value (large $\Delta I(s)$) that will enjoy a large increase in their expected value despite having low likelihoods of validity (small $V(s)$). The claim is only that behind a veil of ignorance about the distribution of the improvement to infringement value, the presumption of in-

enforcement outcomes,¹⁷⁸ this shift in presumptions can be expected to lead to narrower, higher quality claims being filed in the first instance.¹⁷⁹

This result is intuitive. Under the presumption of validity, patentees have protection against over-claiming — to prevail on an argument that an asserted claim is too broad, a patent defendant must show invalidity with clear and convincing evidence. Patentees are much more vulnerable to under-claiming, since the burden is on them to show that the defendants actually infringe the claim.¹⁸⁰ Replacing the presumption of validity with a presumption of infringement would invert this logic. Instead, patentees would be protected against under-claiming, but comparatively more vulnerable to over-claiming. Anticipating the possibility of litigation, applicants will play to their strengths and away from their weaknesses — in this case, by investing more of their prosecution resources in narrow claims as opposed to broad ones.

One potential risk of a presumption of infringement is that it may tend to broaden the effective scope of a claim beyond what it would have been without the presumption. In this way, the effective broadening caused by the presumption itself may offset any success this policy has in moving applicants toward facially narrower claims. However, careful design of the threshold for triggering the presumption could reduce (and perhaps eliminate) this potential for effective broadening. If patentees are required to meet a sufficiently high likelihood-of-infringement threshold to trigger the presumption, the presumption's effects could be limited to cases where there was already a colorable infringement claim without the presumption. If the presumption has an only weak effect on dubious infringement arguments, then the degree of effective re-broadening may be negligible — though admittedly, this risk raises questions in need of further study.

To be sure, introducing an entirely new litigation presumption has many potential costs and benefits,¹⁸¹ and a full treatment of the desirability of such a presumption will be the subject of a subsequent pro-

infringement can be expected to have a larger effect on claims with higher likelihoods of validity than on claims with lower likelihoods of validity. The actual distribution of the improvement to infringement value will greatly depend on the design of the presumption, and thus is the subject of later work.

178. For a discussion of applications with no interest in enforcement, see Part V.A, *infra*.

179. However, replacing the presumption of validity with a presumption of infringement will likely affect inventor rewards differently for different kinds of inventions. *Cf.* Ford, *supra* note 90, at 105–07 (noting differences in information asymmetries depending on the form of infringement).

180. See *supra* Part III.B.2.

181. For example, it is possible that a presumption of infringement could lead to increased rates of patent litigation. On the other hand, evidence relevant to the question of infringement is more likely to be found in the hands of defendants, so shifting the burden of production may have the offsetting benefit of reducing the cost of patent litigation overall.

ject. But it is worth noting that a limited presumption of patent infringement is already available under existing law. By statute, a patentee alleging infringement of a process patent based on the importation, sale, or use of a product made by that process¹⁸² must only establish a “substantial likelihood” of such infringement. Upon this showing, the burden shifts to the defendant to show that the product was *not* made by the infringing process.¹⁸³ So presumptions of infringement are not an entirely untested concept, and could have the benefit of reducing applicants’ incentives to file broad claims while increasing their incentives to file narrow ones.

D. Penalizing Invalid Claims After Examination

Section B introduced the possibility of using application-stage penalties to influence the kinds of claims applicants file. For completeness, it is worth noting that additional deterrence of overbroad claiming could be achieved through monetary penalties for claims that are found to be invalid at some point *after* examination.¹⁸⁴

Imposing a penalty for a claim that is invalidated after issuance changes the model presented above in an important way. Rather than a claim that is found to be invalid having no value, it instead has a negative value. The magnitude of that negative value is equal to the invalidity penalty. Therefore, a claim’s expected value is reduced to account for the possibility that the claim will require the applicant to pay the penalty. This adjusted expected value is given by

$$E_F(s) = [V(s) \times I(s)] - [\bar{V}(s) \times F], \quad (12)$$

where $E_F(s)$ is the expected value of a claim that might require the applicant to pay a penalty as a function of its scope (s), $\bar{V}(s)$ is the probability of a claim being invalid as a function of its scope,¹⁸⁵ and F is the penalty a patentee must pay if she is caught with an invalid claim.¹⁸⁶

182. See 35 U.S.C. § 271(g) (2012).

183. *Id.* § 295. Additionally, to invoke the presumption the patentee must overcome certain procedural obstacles. Specifically, she must show that she “has made a reasonable effort to determine the process actually used in the production of the product and was unable to so determine.” *Id.*; see also *Creative Compounds, LLC v. Starmark Labs.*, 651 F.3d 1303, 1314–15 (Fed. Cir. 2011) (shifting the burden of infringement where the requirements of 35 U.S.C. § 295 had been met).

184. See *Kesan*, *supra* note 1, at 795–97; *Miller*, *supra* note 1, at 705; *Wagner*, *Patent-Quality Mechanisms*, *supra* note 6, at 2170–72.

185. The likelihood of a claim’s invalidity is simply the complement of the likelihood of its validity, that is, $[1 - V(s)]$.

186. This relationship is true provided that the claim actually issues. To focus on the effects of the penalty, the following discussion assumes no substantive examination or that the risk of any claim being permanently rejected is small. Under a functioning examination system, of course, the application-stage deterrent effect of the penalty would need to be discounted by the possibility that an applicant will be “saved” from paying the penalty by

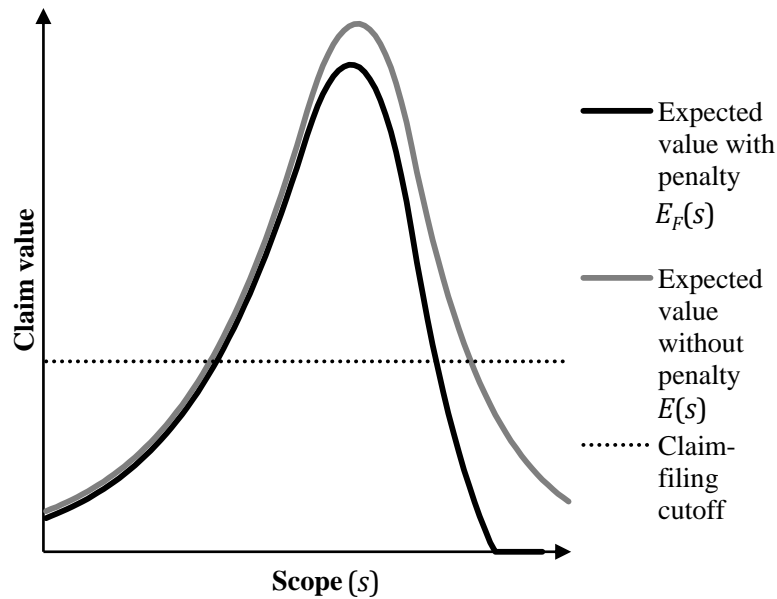


Figure 9: Expected Claim Value With and Without Invalidity Penalty

As Figure 9 illustrates, the expected values of all claims are reduced to some degree by the imposition of an invalidity penalty. Whenever the risk of invalidity is greater than zero, the chance of the penalty will reduce the expected value of a claim by some amount $[\bar{V}(s) \times F]$. At narrow claim scopes — where the likelihood of a claim being invalid is low — this reduction in expected value will be small. However, as claim scope increases — and the likelihood of invalidity increases — the reduction in expected value as a result of the penalty becomes more significant. Holding the claim-filing cutoff constant, applicants will file fewer broader claims than they would if not for the penalty. In this way, an invalidity penalty operates like other changes in legal rules that reduce the expected value of broader claims.

However, there is one feature of the invalidity penalty that is different from any of the legal rules so far discussed: its ability to cause the expected value of a claim to actually become negative. Without penalties for invalidity, any issued claim with *some* chance of validity (no matter how small) will have an expected value greater than zero.

receiving a rejection from the USPTO. This result is one of the paradoxes of a post-examination penalty — its application-stage deterrent effect is actually reduced as examination improves.

That creates the possibility that either a hole in the cost screen¹⁸⁷ or a particularly valuable invention could cause applicants to file claims even if they were very unlikely to be valid. When the expected value of a claim becomes negative, however, this cost screen is no longer necessary to prevent the filing of low-value claims. Beyond that point, the risk of having to pay the penalty is sufficient to deter applicants from filing such broad claims.

An invalidity penalty does not just cause a claim's expected value to become negative; it causes that expected value to cross zero at a specific likelihood of validity that a patent policymaker can control. Recall from Equation (12) that the expected value of a claim subject to an invalidity penalty F as a function of its scope is given by

$$E_F(s) = [V(s) \times I(s)] - [\bar{V}(s) \times F],$$

where $\bar{V}(s)$ is the probability of a claim being invalid as a function of its scope. Replacing $\bar{V}(s)$ in Equation (12) with the equivalent $[1 - V(s)]$, we obtain

$$E_F(s) = [V(s) \times I(s)] - [[1 - V(s)] \times F] \quad (13)$$

$$\equiv E_F(s) = [V(s) \times I(s)] - F + [V(s) \times F]. \quad (14)$$

Setting $E_F(s) = 0$ in Equation (14) and rearranging, we see that the expected value of a claim is zero when

$$F = [V(s) \times I(s)] + [V(s) \times F] \quad (15)$$

$$\equiv F = V(s) \times [I(s) + F], \quad (16)$$

which is satisfied when

$$V(s) = \frac{F}{I(s) + F}. \quad (17)$$

In other words, the likelihood of validity at which the expected value of a claim becomes negative is determined by the relationship between the infringement value of the claim and the penalty for invalidity. For example, if a policymaker wants to ensure that no claims with an expected likelihood of validity of less than fifty percent are filed, she should set the invalidity penalty to equal the enforcement value of the claim against all potential infringers. In that case, setting $F = I(s)$ in Equation (17), the expected value of a claim crosses zero when $V(s) = 0.5$, and applicants would be expected to avoid filing claims with likelihoods of validity below that threshold. Or, as another example, if a slightly more forgiving policymaker wants to set the zero crossing at claims with a likelihood of validity of less than thirty-three percent, she should set the invalidity penalty to be half the enforcement value of the claim against all potential infringers. In that case, setting $F = 0.5 \times I(s)$ in Equation (17), the expected value of a

187. See *supra* Part IV.A.

claim will cross zero when $V(s) = 0.33$, and applicants would be expected to avoid filing claims with likelihoods of validity below that threshold.

These penalties are large — potentially hundreds of millions of dollars charged to the patentee after she loses a validity challenge.¹⁸⁸ However, it is not surprising that such large penalties would be necessary to turn a patent from a net asset to a net liability at some point in claim scope. For the risk of an invalidity penalty to singlehandedly offset the upside potential of a successful infringement suit, the penalty would need to be approximately the same order of magnitude as the infringement value. So, while penalties can theoretically create an independent screen against low-value claims, this feature is not enabled until those penalties become quite large.

Moreover, another obstacle may further limit the usefulness of penalties for adjusting applicant incentives in the USPTO. If litigation-stage penalties are to create *application-stage* deterrence against overbroad claiming, there must be some limits on how applicants may cancel or amend their claims down the road. Otherwise, an applicant could simply file overbroad claims in the USPTO and cancel them if invalidating prior art is discovered at some point in the future. Similarly, if the penalty only affects patentees whose claims are found invalid in litigation commenced by the patentee, the threat of the penalty can be mooted by simply not asserting the claim.

As long as applicants and patent holders have the option of cancelling, amending, or at least not litigating their claims at some point between their initial application and the courthouse steps, the possibility of a litigation-stage penalty is unlikely to serve much deterrent effect at the application stage. These tools may be useful for influencing behavior at later stages,¹⁸⁹ but they are not promising candidates for adjusting applicants' incentives in the USPTO.

V. REFINING THE MODEL

The previous Parts of this Article presented a model of the incentives facing patent applicants. That model worked on the basis of several assumptions, and as a result may not describe the full range of behavior demonstrated by applicants in the USPTO. This Part presents and responds to several potential objections to the arguments and assumptions made above. Specifically, Part A analyzes the incentives facing applicants who file for reasons other than licensing or

188. However, prior commentators have proposed bounties of potentially similar magnitude. See Miller, *supra* note 1, at 711–20 (proposing bounty equal to patentee's past profits from practicing the invention).

189. For example, they may be useful in discouraging patentees from maintaining or asserting questionably valid claims post-issuance.

enforcing their issued patents. Part B considers how deviations between settlement and expected litigation outcomes could affect claiming behavior. Finally, Part C evaluates how value derived from a patent's role in a larger patent portfolio could affect the prior analyses.

A. Some Applicants Have No Intention of Enforcing Their Patents

The previous analysis operated on the assumption that applicants seek to maximize the future enforcement value of the claims they are filing in the USPTO. This assumption is certainly true for some applicants, but it is also certainly not true for all applicants.

Legally, a patent grants its holder the right to exclude;¹⁹⁰ yet, it is difficult to explain the sheer volume of patent applications by legal rights alone. Fewer than two percent of issued patents are ever mentioned in a federal district court, and only a tiny fraction of those are actually litigated to judgment.¹⁹¹ Of course, those numbers understate the total percentage of patents that are privately asserted, since some patents are licensed without a complaint being filed. No one really knows how many patent disputes are resolved this way, but a leading estimate is that the number is approximately five percent.¹⁹² Even with generous rounding of these estimates, it appears safe to say that the great majority of issued patents are never litigated, licensed, or otherwise asserted.

This puzzle — why do so many people file so many patents that are never brought to bear on potential infringers? — has been the subject of discussion among commentators. A simple answer is that while it is unlikely that any single patent will turn out to be both valid and significantly infringed, the payout from obtaining just one such patent justifies many failed attempts.¹⁹³ This phenomenon has been dubbed the “lottery ticket” explanation of patent filing — inventors file large numbers of patents in hopes of striking gold on a few of them.¹⁹⁴ So while the odds of enforcing any single patent are low, inventors nonetheless file frequently in hopes that some will turn out to be worthwhile.

An alternative explanation for why firms file patents they, in all likelihood, will never enforce is rooted in the maxim that “the best defense is a good offense.”¹⁹⁵ Firms that are operating in an area

190. See 35 U.S.C. §§ 154, 271 (2012); *Bloomer v. McQuewan*, 55 U.S. 539, 549 (1852).

191. See Lemley, *supra* note 4, at 1501.

192. See *id.* at 1507 & n.53.

193. See F.M. Scherer, *The Innovation Lottery*, in *EXPANDING THE BOUNDARIES OF INTELLECTUAL PROPERTY 3* (Rochelle Dreyfuss et al. eds., 2001).

194. See Lemley & Shapiro, *supra* note 42, at 81.

195. For discussions of defensive patenting behavior, see Graham & Sichelman, *supra* note 85, at 1065; Lemley, *supra* note 4, at 1504–05.

where they will likely encounter competitors' patents may find that having patents of their own can help secure their own freedom to operate. For example, if Company A holds a patent (or portfolio of patents) that can foreclose a particular product market, the only way Company B may be able to continue to participate is to secure a patent (or portfolio of patents) that is equivalently capable of foreclosing Company A from the market. Facing the prospect that both companies will be locked out of the market, the theory goes, Company A and Company B will enter a cross-licensing agreement, giving each the ability to compete in the market despite the other's foreclosing patents.¹⁹⁶ For applicants like these, patent litigation might only be considered as a last resort. Provided that other industry participants are cross-licensing and otherwise going along peacefully, they would be quite content — and might well prefer — to let their own patents live out their issued life without ever seeing the inside of a courtroom.¹⁹⁷

However, both of these reasons for patenting share a common intermediate goal: to preserve the option of enforcing the issued patent should circumstances evolve to make enforcement desirable. And so even applicants like these — those who may not expect or even hope to someday litigate their issued patents — will have an eye on maximizing the potential enforcement value of their claims. The “lottery ticket” applicant and the “freedom-to-operate” applicant both maximize the strategic value of their applications when they obtain claims with the highest expected enforcement value.¹⁹⁸

Yet as other scholars have observed, there is a limit to how much patenting behavior can be explained by the possibility of enforcement. For example, Clarissa Long has noted that patents may perform an important signaling function.¹⁹⁹ At an individual level, issued patents may reflect the significance of an engineer's accomplishments, her value to the organization, and her potential usefulness to competitors. Indeed, it is common for inventors to list their issued patents on their resumes, and many corporations reward inventors with cash bonuses, public honors, and eligibility for certain promotions.²⁰⁰ Even without

196. See Colleen V. Chien, *From Arms Race to Marketplace: The Complex Patent Ecosystem and Its Implications for the Patent System*, 62 HASTINGS L.J. 297, 299, 308–10 (2010).

197. *But see id.* at 333–34 (noting the ineffectiveness of defensive patenting as protection against non-practicing entities).

198. It is possible that the “freedom-to-operate” patentee may be more narrowly focused on claims that her competitors — or at least her patent-holding competitors — will be likely to infringe. But given the difficulty of predicting exactly what any given firm will be doing years down the road, broad claims are still likely to maximize the leverage that any given patent can provide in negotiations with a patent-holding competitor.

199. See Long, *supra* note 85, at 636–37.

200. See Richard S. Gruner, *Corporate Patents: Optimizing Organizational Responses to Innovation Opportunities and Invention Discoveries*, 10 MARQ. INTELL. PROP. L. REV. 1, 30 (2006); Lemley, *supra* note 4, at 1506; Jessica Silbey, *Patent Variation: Discerning Diversity Among Patent Functions*, 45 LOY. U. CHI. L.J. 441, 454–56 (2013).

such material rewards, many inventors may simply want the satisfaction of seeing their name on an embossed, issued patent.²⁰¹ And at a corporate level, issued (or even pending) patents may signal that a company invests in research and development and that it is protecting its intellectual property.²⁰²

If issued patents are ends unto themselves, then some applicants may be indifferent to the enforcement value of their patents.²⁰³ An engineer seeking prestige, a company wishing to attract investors, an inventor hoping to call herself a patentee — all are drawn to the recognition carried by a patent, rather than its attendant legal rights. As such, they may not care whether their patents are actually valid, provided their patents look good enough to the group for whom these patent signals are intended.²⁰⁴

But with indifference to validity likely comes indifference to claim scope as well. There is no obvious reason why an applicant seeking only a trophy for her wall or a number for an investor report would value a claim with one scope over another — and particularly no obvious reason why she would prefer broad claims over narrow ones. To the applicant lacking any interest in enforcement, the value of an issued claim may very well be constant as a function of scope.

It is not clear how an applicant who is truly indifferent to enforcement will draft her claims. One possibility is that she will draft them essentially randomly, hoping that at least one of them will be granted by the USPTO and result in an issued patent. But another possibility is that she will seek to minimize the time and cost spent in prosecution. In this case, even a very weak examination screen may be sufficient to push this type of applicant toward narrow claims.²⁰⁵ If the cost of obtaining an issued claim increases with scope, but the benefits to the applicant do not, then such an applicant can be expected to draft her claims narrowly — minimizing prosecution costs and increasing the chances of obtaining a patent as swiftly as possible.

201. See William Hubbard, *Inventing Norms*, 44 CONN. L. REV. 369, 373 (2011); Lemley, *supra* note 4, at 1506.

202. See Lemley, *supra* note 4, at 1505–06; Long, *supra* note 85, at 636–37, 647–55.

203. See Lemley, *supra* note 4, at 1514.

204. See *id.* at 1514–15.

205. Cf. *id.* at 1527 (noting that even imperfect examination may serve a socially useful function by discouraging applicants from filing overbroad claims).

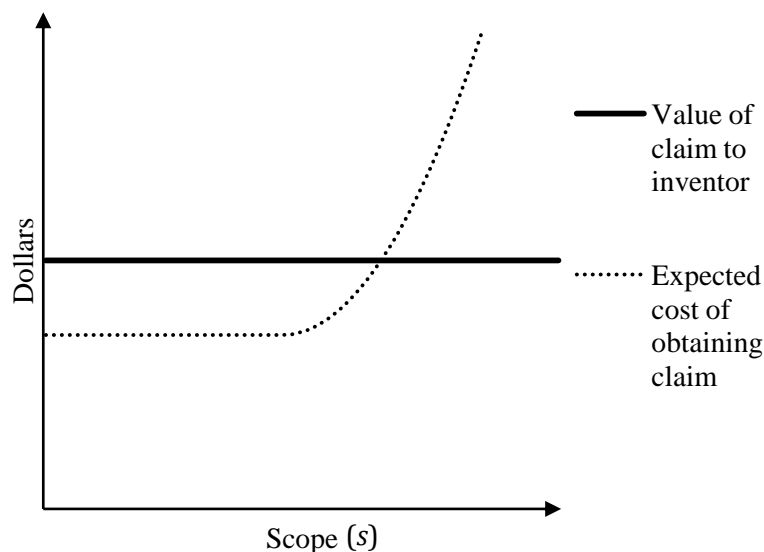


Figure 10: Expected Cost of Obtaining a Claim as a Function of Scope and Scope-Neutral Value

Figure 10 illustrates an example of a claim-value function that is invariant across all claim scopes, along with an expected cost to obtain a claim that increases with claim scope. At narrow claim scopes, the benefit to the applicant of obtaining a claim far exceeds the cost of obtaining that claim. As claim scope increases, the expected cost to obtain the claim rises as well, but the benefits to the inventor do not. Therefore, this type of applicant could be expected to file claims on the narrower side of the spectrum, where the gap between the value of having a patent and the cost of obtaining it is the greatest. Assuming examination plays even a weak screening function, narrow claims will tend to offer the fastest, cheapest route to an issued patent.²⁰⁶

Of course, some inventors and start-ups may want to signal their inventive prowess not only by their number of issued patents, but also by some metric of those patents' values. In other words, applicants with no intention of enforcing their patents may still have an interest in obtaining enforceable patents simply because their employers, investors, or potential acquirers do. In that case, their behavior may be more like the behavior of the applicants described in Parts III and IV, *supra*.

206. However, current patent rules may be failing to screen overbroad claims filed by these types of applicants. The basic filing fee includes the right to file up to twenty claims, and there is no penalty for canceling a rejected claim. *See supra* note 79; *supra* note 147 and accompanying text; *supra* note 163 and accompanying text.

Finally, it is possible that some applicants have no intention of enforcing their patents, but desire claims that are facially as broad as possible. This might be the case if, for example, broader claims — enforceable or not — were thought to signal greater inventive achievement. Of course, pursuing overbroad claims as a signal of prowess risks damaging the inventor's own credibility,²⁰⁷ but it is possible that some inventors are willing to take this risk. This particular group of applicants — to the extent that they become a concern for policymakers — will likely respond more strongly to increases in the expected cost to obtain broad claims than to reductions in their expected enforcement value.²⁰⁸

B. Enforcement Outcomes May Diverge from Expected Litigation Outcomes

The previous analysis assumed that the enforcement value of a patent was based on the likelihood and value of potential litigation outcomes. However, there is a growing concern that many patent settlements today do not accurately reflect expected litigation outcomes.²⁰⁹ Therefore, it is necessary to consider whether (and how) a divergence between settlement value and litigation value could affect applicants' incentives in drafting claims.

The classic example of patent settlements diverging from expected litigation outcomes is the nuisance suit. In this paradigm, a patentee brings an extremely weak or even frivolous case against an accused infringer. The accused infringer knows that she has an excellent chance of prevailing at trial, but because of the high costs of patent litigation, she pays a small settlement to avoid the expense of doing so.²¹⁰ Thus paid off, the patentee moves along the line to the next dubiously accused infringer.

It is not immediately clear how this divergence between patent settlements and expected litigation outcomes affects the values of individual claims to patentees. For example, if a claim's nuisance value is independent of its scope, then the possibility of nuisance suits may

207. See Long, *supra* note 85, at 649–50.

208. Presently, it is unclear whether there exists a significant number of these signaling-focused applicants who desire broad claims without regard to their enforceability. The frequency with which applicants seek patents with truly *no* interest in enforcement, and the behavior of such applicants during the course of prosecution, are ripe fields for further study.

209. See Michael J. Meurer, *Controlling Opportunistic and Anti-Competitive Intellectual Property Litigation*, 44 B.C. L. REV. 509, 516–20 (2003) (describing the rise of opportunistic patent litigation premised upon raising claims that defendants know they would win on the merits, but nevertheless settle because litigating them would be more costly); Ranganath Sudarshan, *Nuisance-Value Patent Suits: An Economic Model and Proposal*, 25 SANTA CLARA COMPUTER & HIGH TECH. L.J. 159, 161–62, 171 (2009).

210. See Fagundes & Masur, *supra* note 9, at 697–98; Sudarshan, *supra* note 209, at 161–64.

not have any effect on an applicant's incentives to file a marginally broader or narrower claim. Instead, a more general topic for consideration is whether this divergence affects some kinds of claims differently than it affects other kinds.

One theory proposes that the divergence between settlements and expected litigation outcomes is greater for asserted claims that are overbroad than for asserted claims that are not infringed. This theory relies on the assumption that it costs a defendant more to defend on grounds of invalidity than it does to defend on grounds of non-infringement.²¹¹ If this assumption is true, then a defendant facing a claim that is very likely invalid may pay more to avoid trial than a defendant facing a claim that is very likely not infringed would pay. And, therefore — all else equal — applicants may prefer broader claims to narrower ones to exploit this larger divergence between settlement values and expected litigation values for overbroad claims.

To be sure, comparing the costs of invalidity defenses with the costs of non-infringement defenses is difficult, and some commentators have observed reasons for which a defendant may actually prefer to make arguments of invalidity over arguments of non-infringement.²¹² Additionally, one must consider the costs of countering each defense from the perspective of the patentee. It is possible that the risk of losing a claim to an invalidity judgment counterbalances any cost advantage that patentees may have when it comes to litigating that issue.²¹³

But, if the divergence between settlements and expected litigation outcomes tends to be greater for broader claims than for narrower ones, then this divergence may be an additional factor that pushes applicants toward filing broader claims. In that case, stronger examination or other cost-based reforms may be necessary to counterbalance these incentives to file more broadly.

211. One reason that invalidity may be the more expensive defense is that defendants typically have better and more timely access to the evidence that is determinative of infringement. *See Ford, supra* note 90, at 104–07.

212. For example, infringement may involve a much larger volume of evidence, and is not typically amenable to cost-sharing through joint defense agreements. *See Sudarshan, supra* note 209, at 185–86. The comparison is further complicated by the fact that invalidity judgments may lend an unearned advantage to the defendant's competitors in a way that non-infringement judgments do not. *See Farrell & Merges, supra* note 1, at 952. For a thorough discussion of the desirability of each defense from the perspective of defendants, *see Ford, supra* note 90.

213. As discussed above, judgments of invalidity typically have a more harmful preclusive effect for patentees than do judgments of non-infringement. *See supra* Part III.D.1.

C. Even Invalid Patents Have Portfolio Value

The previous analysis was based on a particular patent's value for purposes of either signaling²¹⁴ or enforcement.²¹⁵ However, there is another possibility: a large portfolio of patents could have significant signaling or enforcement value, even if some (or many) of those patents are not likely to have much of either type of value on their own.²¹⁶ If the high cost of litigating even a single patent can cause settlement outcomes to diverge from expected litigation outcomes, then the threat of asserting hundreds or thousands of patents could cause settlement outcomes to diverge even more significantly. An applicant pursuing a portfolio enforcement strategy might therefore be more interested in obtaining a large number of patents, rather than maximizing the enforcement value of those patents individually.²¹⁷

The ability of large portfolio owners to extract undeserved royalties may signal problems with current patent rules. However, it is not clear why a portfolio-based licensing strategy would lead to a preference for overbroad patents over narrow patents with a low likelihood of infringement. In the portfolio-based licensing paradigm, a portfolio derives its value from the holder's ability to bring a long tail of patent infringement suits, even if only a few (if any) of those suits turn out to be winners. Winning is obviously better for the portfolio holder, but even losses are not so bad given a large enough portfolio. As between losing on grounds of invalidity and losing on grounds of non-infringement, portfolio holders would seem to prefer to lose on non-infringement, since a loss on non-infringement preserves the holder's ability to assert the patent against another potential licensee in the future. By contrast, a finding of invalidity effectively takes the patent out of the assertable portfolio.²¹⁸

Moreover, a portfolio-focused prosecution strategy may actually lead applicants to seek narrower claims than they would if they were evaluating the value of their patents individually. A shift toward portfolio-based licensing and enforcement creates pressure for applicants to seek a large number of low-cost patents.²¹⁹ But this shift does not have to mean pressure to obtain low-quality patents. With an examination screen working properly, the cheapest way to obtain lots of issued patents may be to simply file lots of narrow claims. In this way, portfolio-building applicants may look a lot like signaling-focused applicants, pursuing strategies that will yield the greatest number of

214. See *supra* Part V.A.

215. See *supra* Parts III, IV.

216. See Parchomovsky & Wagner, *supra* note 58, at 66.

217. See *id.* at 42; Wagner, *Patent-Quality Mechanisms*, *supra* note 6, at 2155. See generally Michael Risch, *Patent Portfolios as Securities*, 63 DUKE L.J. 89 (2013).

218. See *supra* Part III.D.1.

219. See Chien, *supra* note 196, at 338.

issued patents for their prosecution dollar.²²⁰ Thus, increasing the effectiveness of the examination cost screen may push portfolio-building applicants in the direction of narrower claims that are more likely to be valid. A functioning examination cost screen could help transform market pressures to pursue a large number of cheap patents into applicant incentives to seek narrow ones.

VI. CONCLUSION

Patent applications do not file themselves. Behind each application is an applicant who has invested substantial time and resources into seeking patent protection. The decisions of that applicant can have a significant effect on the costs that must be spent examining and opposing that application, as well as on the quality of the patent that ultimately issues.

This Article has noted several ways in which existing patent rules affect applicants' incentives to file narrow or broad claims. It has also suggested several changes to those rules that could push applicants in the direction of filing higher quality applications. Undoubtedly, there are other ways that existing rules affect claim drafting and other reforms that could further align applicants' incentives with those of the public. A greater emphasis on the role of applicants in improving patent quality could lead to better patents without requiring additional investment in examination or intervention by third parties.

220. See *supra* Part V.A.