Winter 2022

Patent Inconsistency

Saurabh Vishnubhakat
Texas A&M University School of Law, sv10@law.tamu.edu

Follow this and additional works at: https://www.repository.law.indiana.edu/ilj

Part of the Intellectual Property Law Commons, Litigation Commons, and the Supreme Court of the United States Commons

Recommended Citation
Available at: https://www.repository.law.indiana.edu/ilj/vol97/iss1/2

This Article is brought to you for free and open access by the Maurer Law Journals at Digital Repository @ Maurer Law. It has been accepted for inclusion in Indiana Law Journal by an authorized editor of Digital Repository @ Maurer Law. For more information, please contact kdcogswe@indiana.edu.
Patent Inconsistency

Saurabh Vishnubhatk*

Despite the promise of efficiency through the use of expert agency adjudication in U.S. patent law, administrative substitution continues to fall short. In a variety of ways, the decade-old system of Patent Office adjudication is simply an additional place to litigate rather than the robust technocratic alternative it was meant to be. These problems have arisen from important defects in the statutory design, but also from the enormous expansion and ascendancy of the Patent Office itself. Moreover, while duplicative litigation over patent validity is recognized and criticized, its scale and scope has eluded detailed empirical analysis until now. This Article documents and measures those ill effects and explains their structural origins. It evaluates varying responses from Federal Circuit and Supreme Court decisions and from regulatory experimentation in the Patent Office itself. The Article concludes with proposals for reform. The analytical framework for reform that is suggested here is especially timely given the incompleteness of judicial and regulatory responses thus far.

* Professor, Texas A&M University School of Law; Professor, Texas A&M University Dwight Look College of Engineering; Fellow, Duke Law Center for Innovation Policy. The author was formerly an advisor at the United States Patent and Trademark Office. The arguments in this writing are the author’s and should not be imputed to the USPTO or to any other organization. Sincere thanks to Colleen Chien, Stuart Graham, Sharon Israel, Dmitry Karshitek, Sapna Kumar, Megan La Belle, Mark Lemley, Alan Marco, Michael Meurer, Andrew Michaels, Tejas Narechania, Lisa Ouette, Arti Rai, Sarah Rajec, Alex Roberts, Josh Sarnoff, James Stern, Neel Sukhatme, Jason Rantanen, Greg Vetter, Chris Walker, Melissa Wasserman, and Steve Yelderman for helpful comments and suggestions.

Sincere thanks also to a great number of USPTO advisors and PTAB judges for their generous insights at the USPTO Chief Economist Visiting Speaker Series, and to the organizers and participants of the Georgetown Empirical Patent Conference, the GW Law Intellectual Property Colloquium, the 2019 Junior IP Scholars Association Workshop at GW Law, the 2019 IP Scholars Conference, the 2019 PatCon Conference, the 2019 WIP Colloquium, and the Lewis & Clark Law School Symposium on the Procedure of Intellectual Property.

This work was generously supported by a competitive grant from the Honorable Nancy F. Atlas Intellectual Property American Inn of Court in Houston, Texas, and the Institute for Intellectual Property & Information Law at the University of Houston Law Center.
INTRODUCTION .............................................................................................................. 60

I. Duplicative Litigation of Patent Validity ................................................................. 64
   A. The Promise of Efficiency from Substitution .................................................. 64
   B. Evidence of Duplication ..................................................................................... 69
      1. Prior Findings in Context .............................................................................. 70
      2. New Evidence of PTAB Duplication ............................................................ 70
   C. Collective Action and Harassment in Patent Regulation ............................... 74

II. The Arc of Duplication: Causes and Reforms ....................................................... 78
   A. The AIA’s Statutory Design ............................................................................ 79
      1. The Agency’s Weaker Burden of Proof ....................................................... 79
      2. Estoppel and Disregard for Judicial Findings .............................................. 81
      3. The Moving Target of Judicial Stays .......................................................... 84
      4. The Largely Overlooked ITC ...................................................................... 88
   B. The Patent Office’s Ascendancy ..................................................................... 90
      1. The Agency’s Unreviewable Screening of Petitions .................................... 90
      2. The Agency’s Need to Reconstruct Patents ............................................... 92
      3. The Agency’s Practice of Partial Adjudication ......................................... 96
      4. The Agency’s Self-Policed Statutory Deadline ......................................... 100
   C. Recent USPTO Attention to Duplication ......................................................... 103
      1. Agency Reaction to Internal Duplication ............................................... 103
      2. Agency Reaction to Inter-branch Duplication ......................................... 105

III. Toward More Efficient Substitution .................................................................. 107
   A. Timing and Institutional Competence ............................................................ 107
      1. Priority to the First Adjudicator ................................................................. 108
      2. Priority to a Later Adjudicator .................................................................. 111
      3. Priority to the Courts ................................................................................. 113
      4. Priority to Other Institutions ..................................................................... 117
   B. Legislative Prescriptions ................................................................................. 121
      1. Aligning the Administrative Burden of Proof .......................................... 122
      2. Making Interbranch Estoppel Bilateral ............................................... 123
      3. Codifying Court-Aligned Claim Construction ........................................ 126

CONCLUSION .............................................................................................................. 126

TABLES AND FIGURES ............................................................................................. 127

INTRODUCTION

The dramatic reform that upended patent law a decade ago has never fully come to grips with a defect at the heart of its design. When the 2011 America Invents Act (AIA) took effect amid bipartisan, bicameral consensus, its most far-reaching change was a reallocation of power away from the federal courts and into the administrative setting of the United States Patent and Trademark Office (USPTO).1 The courts had long been the target of criticism for intolerably high costs, delays, and inaccuracies that surrounded patent litigation—especially disputes over whether the patents

involved were valid at all. Patent validity is a complex inquiry frequently requiring not only familiarity with the doctrinal requirements of patentability but also the science and engineering principles underlying the patented invention. Lay judges and juries tend to be ill-equipped or at least require significant education in each case, and so the AIA aimed to have USPTO administrative patent judges with both “competent legal knowledge and scientific ability” evaluate patent validity.

The defect came from two unappreciated difficulties in the task of error-correction. One is that valid (or likely-valid) patents are not distinguishable from invalid (or likely-invalid) ones, at least not easily. Indeed, if they were, then the enterprise of error-correction would itself be largely pointless. The other difficulty follows from the first: without careful design, creating processes to invalidate “bad” patents more easily will routinely make it easier to invalidate “good” patents as well.

Yet this was the very mandate of the USPTO Patent Trial and Appeal Board (PTAB), whose expert administrative patent judges would more cheaply, more quickly, and more accurately dispense with poor-quality patents that the USPTO’s own patent examination process should never have allowed to be granted. The PTAB was to be a highly preferable substitute for the federal courts. Defendants who had been sued in U.S. district court for infringing a patent could efficiently resolve the validity of those patents in the PTAB rather than in the court. Even those who had not been sued for infringement could challenge a patent in the PTAB and seek preemptive declaratory relief.


3. Saurabh Vishnubhakat, The Field of Invention, 45 HOFSTRA L. REV. 899, 900–01 (2017) (explaining that compliance with patent validity requirements is assessed from the perspective of the “person having ordinary skill in the art (PHOSITA) to which the invention pertains” (internal quotations omitted)).


6. Id.

7. Id.


10. Strategic Decision Making, supra note 2, at 49 (discussing the standard model of defensive petitioning).

11. Id. at 48 (discussing the nonstandard model of preemptive petitioning).
This emphasis on the twin virtues of expertise and substitution brooked little sympathy for the risk of collateral damage to valid patents, especially through repeated legal attacks. One common refrain was that patent owners had no real cause for concern if their patents were valid, for they would surely survive legal challenge. Another, more systemic assurance was that substitution into the PTAB would significantly conserve judicial resources. But to the contrary, the USPTO’s expertise has not been a panacea for serial challenges against patent owners, and a significant share of PTAB litigation has duplicated the work of other tribunals.

For example, during the 2010s, Eli Lilly asserted its U.S. Patent No. 7,772,209 in U.S. district court against a number of alleged infringers. The validity of the Lilly patent, directed to “antifolate combination therapies,” was later challenged in multiple petitions for administrative review before the PTAB. The PTAB ultimately upheld the patent, but by that time, the district court litigations had already reached their own conclusions, also upholding the patent in each of five parallel cases.

In an even more stark example, Trading Technologies in 2005 and again in 2010 had asserted its U.S. Patent No. 6,766,304 in U.S. district court against multiple alleged infringers. During the same period, Trading Technologies had also been sued in U.S. district court for declaratory judgments that its patent was invalid.


and was unenforceable due to patent misuse.\footnote{16} As with Lilly, the Trading Technologies patent, entitled “Click based trading with intuitive grid display of market depth,” was later challenged in administrative review before the PTAB. The PTAB ultimately upheld the patent. In this case, though, the district court cases had not only upheld the validity of the patent before but in most cases had already reached those judgments before the PTAB challenge was even filed—a procedural posture discussed in more detail below.

These cases illustrate a troubling level of irrelevance for prior federal court decisions—even repeated decisions—upholding a patent’s validity when the PTAB is undertaking its review. This is not necessarily to the blame of the PTAB. The Lilly litigations in the court and the agency were mostly contemporaneous, and one could fairly attribute the duplication to a lack of judicial stays pending the outcome of the PTAB’s review. Yet the Trading Technologies litigations were clearly sequential, and multiple court decisions about the patent’s validity had already issued before the PTAB entered the picture. And in both situations, the resources the PTAB expended in upholding the patents were not a substitute; the courts still expended their own resources to uphold the patents as well.

The problem of resource consumption is an important one, as it puts into perspective the claim that valid patents will tend to survive review. Lilly’s patent certainly did, in court as well as in the PTAB, but not without duplication. The Trading Technologies patent also survived, but also with duplication—and the sequential timing in that case additionally suggests that PTAB review was not intended as an alternative to the federal court but was simply a result-oriented repeat attack that disregarded prior judicial conclusions to the contrary.

Moreover, not all patents fare so well even with duplication. For example, during the mid-2010s, CyWee asserted its U.S. Patent No. 8,441,438 against several high-profile technology firms including Apple, Samsung, Huawei, and Motorola.\footnote{17} These cases were also distributed across multiple patent-savvy judicial districts including the Northern District of California and the District of Delaware and not, as the increasingly common narrative goes, shoehorned into the Eastern District of Texas.\footnote{18} When Google mounted a PTAB challenge against the CyWee patent, directed to a “3D pointing device and method for compensating movement thereof,” multiple district court cases were nearing judgment and one had already upheld the patent.\footnote{19} By the time the PTAB had accepted Google’s challenge for review, additional district court decisions had upheld the CyWee patent. Nevertheless, contrary to all
prior court decisions, the PTAB eventually concluded that the patent was invalid and cancelled it.

This compounded problem, not only of duplicative litigation and resource consumption but also of inconsistent judgments, raises grave concerns about the relationship of the PTAB to the federal courts. These concerns belie the assurances and expectations of the America Invents Act that valid patents would survive legal challenges and that the PTAB would systematically be a cost-saving substitute for the federal courts.

This Article describes these ill effects using empirical evidence of opportunistic litigant behavior across the court-agency divide. It also traces these effects to important flaws in the statutory design of the PTAB. Part I describes the problem of duplicative litigation, first by explaining more fully the PTAB’s promise of efficiency from substitution and then showing that much of the PTAB’s work has instead been duplicative of other tribunals. Part II situates the causes of this duplication in important features of the PTAB’s statutory design as well as in the administrative ascendancy of the Patent Office itself. It also discusses partial corrections, both judicial and regulatory, to the problem of duplication. Part III emphasizes the need for comprehensive reform through targeted legislation, beginning with an analytical framework for more efficient substitution and concluding with specific legislative proposals.

I. DUPLICATIVE LITIGATION OF PATENT VALIDITY

For better and for worse, the current structure of post-issuance patent validity review in general, and of administrative patent revocation in particular, reflects a legislative promise at the heart of the 2011 America Invents Act. That promise was to improve the efficiency of modern patent litigation, a system that had become overburdened with pathologies of cost, delay, inaccessibility, and undesirable strategic behavior. Apart from bringing about fundamental shifts in ex ante patent examination to improve the quality of issued patents in the first instance, the AIA also created a complex suite of ex post validity review proceedings to manage the downstream problem of patents that should not have been issued but were.

The latter reform was bifurcated into two distinct approaches, one prospective and the other retrospective. This Part describes that bifurcated solution, traces its effects across the ecosystem of tribunals that are empowered to evaluate patent validity, and identifies an important problem that has persisted—and, in some cases, has been amplified—in spite of the AIA’s stated goal of improving patent litigation’s efficiency.

A. The Promise of Efficiency from Substitution

The catalog of patent litigation problems that motivated the AIA is now virtually a literature unto itself and need not be recited in full here, though a brief review of representative scholarship is helpful to understanding and evaluating the particular solutions that the AIA put in place. Scholars of law and economics have explored the

nature and origins of these problems across the patent life cycle and across the various institutional settings that patents traverse along the way.

Notable studies of patent examination have shown that patent applicants enjoy far-reaching advantages including statutory law that favors the issuance of patents,\textsuperscript{21} appellate review that disproportionately scrutinizes the denial of patents,\textsuperscript{22} and low costs to entering the patent examination process itself,\textsuperscript{23} among others. Empirical research in the same vein has revealed that the Patent Office itself suffers a number of corresponding disadvantages, including resource constraints that hamper its ability to reject patent applications with meaningful finality\textsuperscript{24} and lead the agency to allocate inadequate time for its examiners to review applications with appropriate rigor.\textsuperscript{25} Institutionally, too, patent examination has received much criticism for being unduly influenced by a Patent Office funding model that relies on maintenance fees for granted patents and thus distorts the agency’s financial viability itself toward granting patents.\textsuperscript{26}

These patent examination problems have also given rise to a variety of explanations and approaches to reform. Influential among the explanations for why Patent Office examination not only is resource-constrained but also is likely to remain that way was the rational ignorance hypothesis; that is, most patents will not prove to have been worthy of expensive and detailed examination, that it is impossible to know which applications will and which ones will not, and that it is therefore inefficient to give rigorous scrutiny to all applications alike.\textsuperscript{27} On this view, it is preferable to focus attention later, when both the value and the uncertainty of a particular patent have become salient, and only on those patents that merit this closer evaluation.\textsuperscript{28} The appetite for reform, meanwhile, frequently took issue with the anomalous place of the Patent Office as one of the few agencies in the modern administrative state that has not been delegated significant policymaking prerogatives or legal authority to promulgate substantive rules and, as a result, has

\begin{itemize}
\item \textsuperscript{21} Sean B. Seymore, \textit{The Presumption of Patentability}, 97 MINN. L. REV. 990, 1007–08 (2013).
\item \textsuperscript{27} Mark A. Lemley, \textit{Rational Ignorance at the Patent Office}, 95 NW. U. L. REV. 1495, 1497 (2001).
\item \textsuperscript{28} Id. at 1521.
\end{itemize}
not received meaningful deference from the courts for its administrative pronouncements.29

Patent litigation has received even more attention, including empirical attention, due in part to the high stakes and increasingly rich availability of relevant data. Research from industrial organization economics had long used patents as proxies for innovation30 and patent citations as proxies for knowledge flows31 as well as market value.32 In the past twenty years, this economic research turned toward the institutional and doctrinal attributes of patent litigation itself.33 What followed was a sustained empirical focus—one that continues to the present—in the legal literature on patents, including a range of contested issues such as forum shopping by patent litigants,34 the interpretation of patent claims,35 findings as to liability and as to

---


34. See generally Fromer, supra note 18; Vishnumbhat, supra note 18; see also Daniel Klerman & Greg Reilly, Forum Selling, 89 S. CAL. L. REV. 241 (2016).

remedies such as damages and attorney fee shifting, and appellate review in the Federal Circuit. This literature also shed light on problems of patent quality that would lay important groundwork for expanded post-issuance review of patent validity by interrogating the attributes of litigated patents themselves, including value, uncertainty of scope and infringement, and intrinsic as well as extrinsic correlates of selection into litigation.

The narrative that emerged from this burgeoning literature was essentially that patents were of increasingly questionable quality and were being asserted in a litigation environment that tolerated and, at times, even rewarded abuse by entities ranging from large and otherwise innovative companies to patent trolls and other rent-seekers. Expanding the ability of accused patent infringers to challenge the validity of patents was the correspondingly popular solution, and the key question for this aspect of the AIA was how to do so.

The organizing principle that Congress ultimately chose was, in a word, expertise. Although courts and, to a lesser extent, non-Article III forums for patent litigation such as the International Trade Commission had long dealt with patent validity challenges as a defensive response to infringement claims and occasionally


42. Vishnubhakat, supra note 1, at 1737–1741.
in preemptive suits for declaratory judgment, they generally lacked doctrinal expertise in patent law as well as scientific and technological expertise in the inventive subject matter of patents themselves.\textsuperscript{43} The Patent Office, meanwhile, was steeped in both forms of expertise and had already gained some institutional competence with post-issuance review,\textsuperscript{44} beginning with the creation of \textit{ex parte} reexamination and further with the addition of \textit{inter partes} reexamination.\textsuperscript{45}

This technocratic approach to reviewing already-issued patents took two forms in the AIA. The comprehensive law would already work a fundamental shift in the initial examination of future patent applications. Priority in patent rights would no longer go to the first to \textit{invent} something but rather to the first inventor to \textit{file} for patent protection over it,\textsuperscript{46} creating pressure to disclose innovations to the public more promptly or risk losing legal rights to more expeditious rivals.\textsuperscript{47} Moreover, the scope of admissible evidence about the pre-existing state of the art was expanded, pushing the quality of patents higher by raising the standard for establishing that an invention is patentable at all.\textsuperscript{48}

Still, given that many of the same examination pathologies of resource constraints, asymmetric appeal, and rational ignorance still persisted, Congress created a mechanism for immediate post-issuance review of those patents that the agency would grant even under the new AIA framework. This mechanism was post-grant review (PGR), and it applied prospectively.\textsuperscript{49} From the time of a patent’s issuance until nine months later, any person other than the patent owner could challenge the patent’s validity.\textsuperscript{50} Nearly all of the substantive statutory grounds for patentability were available, and all forms of relevant prior art evidence could form the basis for the challenge.\textsuperscript{51}

Also available prospectively were two additional forms of post-issuance review. One was \textit{inter partes} review (IPR), which would become available for AIA patents nine months after issuance—that is, once the opportunity for post-grant review had passed—or after the conclusion of a PGR on the patent if one had been initiated.\textsuperscript{52} Unlike PGR, IPR would allow only challenges to the novelty or nonobviousness of the patented invention and only on the basis of prior art evidence in the form of patents and printed publications.\textsuperscript{53} The other additional form of prospective post-issuance review was covered business method review (CBM), which was patterned after PGR in scope and prior art evidence but was limited to nontechnological patents

---

43. \textit{Id.}
45. \textit{Strategic Decision Making, supra note 2, at 55–58.}
49. \textit{Id.} § 6(d), 125 Stat. at 299–313.
50. \textit{Id.}
51. \textit{Id.}
52. \textit{Id.} § 6(a), 125 Stat. at 299–313.
53. \textit{Id.}
that claim a financial product- or service-related invention.\(^{54}\) Like IPR, CBM became available for AIA patents only after the window for PGR had passed or after the conclusion of a PGR if one had been initiated.\(^{55}\) However, unlike IPR, CBM could be used only defensively after the would-be challenger had been charged with infringement of the relevant patent.\(^{56}\)

The upshot of these prospective mechanisms was to provide prompt and vigorous scrutiny of those patents that had survived to issuance, thus supplying at least some of the incomplete information that made rational ignorance rational, followed by a less vigorous but ongoing opportunity for further scrutiny as needed. There still remained, however, the problem of all the patents issued prior to the AIA whose questionable quality had spurred the reform in the first place.

To enable challenges against these pre-AIA patents, Congress gave the IPR and CBM review systems retroactive effect as well.\(^{57}\) This meant that PGR would not be available, but any pre-AIA patent that had not yet expired would immediately be subject to IPR and CBM so long as the other requirements for these proceedings were satisfied.\(^{58}\) This retroactive applicability was consistent with the overall desire of Congress to vindicate the consensus around pre-AIA patent quality, though it posed self-evident problems for the reliance interests of patent owners whose legal rights and economic interests had just shifted underfoot.\(^{59}\)

The retroactivity of IPR and CBM also created significant new problems, which were only somewhat less obvious at the time, from the potential for duplicative litigation and conflicting judgments from multiple tribunals over the same patents. The superior expertise of the Patent Office had been the watchword for substituting a system of predominantly judicial review with one that increasingly prioritized administrative review.\(^{60}\) But what if, instead of a substitute, the PTAB became yet an additional place to fight, adding greater cost to a litigation landscape already seen as inefficient?

### B. Evidence of Duplication

It is now clear from both data and doctrine that this risk of duplication and conflict has, indeed, been realized. Moreover, because of how Congress structured the post-issuance review proceedings of the AIA, the particular directions of this duplication and conflict have cut squarely against patent owners, often with little regard for the substantive merits of their patent rights. Patents whose validity has already been adjudicated are frequently relitigated, but only where the initial decision maker upheld the patent. As a result, patents that survive continue to be challenged

---

54. Id. § 18, 125 Stat. at 329–31.
55. Id.
56. Id.
57. Id. §§ 6(a), 18, 125 Stat. at 299–313, 329–31.
58. Id.
60. Vishnubhakat, supra note 1, at 1737–1741; Strategic Decision Making, supra note 2, at 53–54.
repeatedly until they are found invalid, whereas a single finding of that patent’s invalidity is conclusive.

1. Prior Findings in Context

In this regard, initial large-scale empirical research into the PTAB offers helpful guideposts. Although only 12.7% of federal-court-litigated patents are challenged in the PTAB, 86.8% of patents challenged in IPR and CBM proceedings are also involved in litigation before the federal courts. Meanwhile, a 70% majority of petitioners who challenge patents in the PTAB do so as a defensive response after being sued for infringing those patents, following the so-called standard model of PTAB substitution. The remaining 30% do so preemptively, before any litigation is initiated against them, at least as to the patents in question, following the nonstandard model. Yet these statistics answer the threshold descriptive question of what court-agency substitution looks like. They say little about the extent to which the PTAB serves as not just another forum for disputing patent rights but as an actual substitute.

For example, in the case of IPR, a standard petitioner’s challenge in the PTAB will reflect substitution only if the presiding judge in the underlying federal court litigation issues a stay pending the resolution of the IPR. The same is true of a nonstandard petitioner, who may strike first and preemptively in the PTAB but whom the patent owner might still sue for infringement subsequently. It is rare (3.01%) for an infringement action in district court, when it comes at all, to come after an IPR has been filed. In such cases, however, a large majority of IPR petitioners (81.0%) are the very ones who are subsequently named as infringement defendants in district court litigation.

Here, too, substitution requires the district judge to issue a stay pending the resolution of IPR. Otherwise, the two adjudicators proceed in parallel, and the result is simply a duplicative expenditure of resources. The tendency of judges, for their part, actually to issue stays pending IPR is highly variable across jurisdictions and over time. The variation in the rate of stays is good reason to expect that the PTAB’s workload likely includes a considerable amount of duplication.

2. New Evidence of PTAB Duplication

That expectation is borne out by the results of a new study—whose findings are reported as follows—directly examining the scope of duplication, rather than substitution, in the PTAB. The study relied on a merged set of multiple sources into a single comprehensive dataset. One source was of all patents whose validity was challenged in a PTAB proceeding from the start of AIA review in September 2012 through December 2018. A second source was of all patents that were involved in

61. Strategic Decision Making, supra note 2, at 69.
62. Id. at 73.
63. Id.
64. Id. at 70.
65. Id.
66. The principal source for this data was Lex Machina.
litigation in the U.S. district courts, the U.S. International Trade Commission, or the U.S. Court of Federal Claims from January 2005 through December 2018. A third source was of all substantive patent validity determinations by the PTAB from the start of AIA review in September 2012 through December 2018. A fourth source was of all substantive patent validity determinations by the U.S. district courts from January 2008 through December 2018.

As an initial matter, looking solely at court-PTAB duplication, of the patents that reached a final validity determination in the PTAB, 24.5% are patents that also received at least one validity determination in the U.S. district courts. This alone is cause for concern—that a quarter of the patents that come before the PTAB do not reflect a cost-saving but instead are duplicating expenditures of judicial resources in the Article III courts. What is more, the validity determinations of the courts are, for reasons further explained below, predominantly findings that uphold the validity of the patents in question. Where the PTAB invalidates the patent, any prior judicial conclusions to the contrary are necessarily disregarded. Where the PTAB now upholds the patent, any prior judicial conclusion to the same effect is wastefully replicated.

Looking more closely at the 24.5% of patents that reflected court-PTAB duplication reveals that 77.4% of this subset were not just adjudicated by the courts as well but were adjudicated by the courts first. It is this majority of duplicatively adjudicated patents for which the prior judicial conclusions upholding validity are being disregarded. The remaining 22.6% pose problems as well. For these, the administrative conclusion about validity came first while the judicial conclusion came later, indicating that the parallel court proceeding was not stayed. Where the PTAB upheld the validity of the patent, the later-acting court can only have reached the same conclusion given the divergent legal standards that govern PTAB and district court proceedings—meaning that the duplication offers no additional information at the margin. But where the PTAB struck down the patent, the later-acting court might have upheld the patent instead. The latter conclusion would carry no weight, as the single initial finding of the patent’s invalidity is conclusive, but it would nevertheless reveal arbitrage as between the divergent legal standards.

There are even situations in which a district court may act first to uphold a patent, but because the judgment itself is not entered as a formal matter, the PTAB in a parallel proceeding will consider itself free to render a conflicting judgment striking down the patent—and the Federal Circuit will disregard the first (judicial) outcome even in the same case, reasoning that the first judgment was not truly “final.”67 This unusually and controversially strict doctrine, the so-called absolute finality rule,68 further complicates the already problematic duplication that is apparently at work in the PTAB’s caseload.

In practical terms, of course, the question of which adjudicatory institution will finish first, the PTAB or the court, is not driven solely by the PTAB nor by the Federal Circuit’s treatment of competing cases that are both substantially finished. As noted above, the issuance of stays in court litigation pending the outcome of a

contemporaneous PTAB review will affect whether duplication takes place. Of the 1300 contested motions for a stay pending inter partes review filed in the U.S. district courts from September 2012 through December 2020, the top five districts together accounted for more than half of all such motions. These were the District of Delaware (176 motions), Northern District of California (160 motions), Eastern District of Texas (153 motions), Central District of California (119 motions), and Northern District of Illinois (66 motions). Not coincidentally, these are also the districts in which patent cases filings in gross are also perennially the highest.

The rates at which judges in these districts granted stays were highly variable both as compared with each other and over time when tracking the grant rates as a three-year average starting from the October 2012–September 2015 period and moving quarterly until the January 2018–December 2020 period. From the outset of the AIA trial proceedings, the Northern and Central Districts of California both began close to a grant rate of 50%. With occasional variation between them, both rose substantially to 72% and 64%, respectively.

By contrast, the Eastern District of Texas and the District of Delaware both began close to a grant rate of 30%, but the Eastern District of Texas declined to 11%, whereas the District of Delaware initially fell to 17% during the mid-2010s but rose steadily and substantially since, to 51%. Meanwhile, the Northern District of Illinois began at a grant rate of 62%, the highest among high-volume patent districts, and stayed relatively consistent until the mid-2010s. The number of motions for stays began to decline starting in 2016, and the rate at which they were granted also began to decline starting in 2017, so that the current three-year average stands at 33%. Figure 1 summarizes these trends.

There are also other tribunals to consider. Because the Article III courts have long been the predominant forum for resolving patent disputes and remain the principal competitors of the PTAB, the findings about duplication thus far are an important new contribution to the policy discussion over the design of the PTAB. However, neither the AIA’s intended substitution nor its observed duplication are limited to Article III courts alone. Decision-making about patent validity in Article I tribunals including the International Trade Commission (ITC) and the Court of Federal Claims (CFC) are also subject to the AIA and raise similar concerns. The overall share of patents that are litigated in the CFC is quite small, less than one-half of one percent, and so can be safely bracketed. The share of patents litigated in the ITC, however, is larger by an order of magnitude, about 4.9%, and is nontrivial.

Indeed, recent scholarship suggests that the role of the ITC in the patent ecosystem has grown, not diminished, in response to ongoing legal and institutional changes. For example, the Supreme Court’s 2006 decision in eBay Inc. v. MercExchange, L.L.C., which raised the difficulty of obtaining injunctive relief against accused

69. This subset of contested motions excludes motions that were mutually agreed or stipulated to by the parties.


71. See infra Tables and Figures at Figure 1.

patent infringers, drove litigants from the federal courts into the ITC. The ITC grants exclusion orders that are functionally quite similar to injunctions in equity, but the eBay framework does not apply in the ITC, initially making it an attractive avenue for mitigating or even evading the practical effects of the eBay holding. Meanwhile, more recent precedents such as the Federal Circuit’s 2008 decision in Kyocera Wireless Corp. v. International Trade Commission purporting to limit the scope of ITC remedies have done little to dampen the enthusiasm of litigants for seeking out the ITC.74

That the ITC overlaps with other forums is also empirically clear, although the relevant data on this point precedes the AIA by some years. Because of the ITC’s more factually intensive pleading standards, its more aggressive case management and trial timeline, and its greater willingness to award injunction-like relief, the ITC has certainly been a highly sought-after forum. However, its inability to award damages and its non-binding force upon district courts means that its relief is “neither complete nor final,” creating significant incentives to use the ITC for duplicative litigation rather than for substitution against other venues.77

Revisiting the ITC in light of post-AIA data on patent validity review reveals that much of the relevant activity traverses all three major settings: the Article III courts, the PTAB, and the ITC. The share of patents that are asserted in the ITC and are also involved in federal court litigation is 86.6%. The share of ITC-asserted patents that are also involved in litigation but are not challenged in the PTAB is 64.2%. Meanwhile, the share of patents that are asserted in the ITC and are also challenged in the PTAB is 24.5%. The share of ITC-asserted patents that are also challenged in the PTAB but are not involved in court litigation, however, is only 2.0%.

This means the large majority of patents from the ITC are also involved in court litigation, and over a quarter of that court-ITC overlap (25.9%) involves the PTAB as well. And although a relative minority of patents from the ITC are also challenged in the PTAB, even that modest ITC-PTAB overlap almost always (91.8%) involves court litigation as well. Among these adjudicatory institutions, the overlap between

74. See Chien & Lemley, supra note 73, at 2–3.
75. 545 F.3d 1340 (Fed. Cir. 2008).
78. Id. at 64.
79. Notably, this is nearly identical to the 86.8% of PTAB-challenged patents that are also involved in federal court litigation. See Strategic Decision Making, supra note 2, at 69, and accompanying text.
80. The remainder of ITC-court overlapped patents that are challenged in the PTAB (86.6% – 64.2% = 22.4%) represents a quarter of the overall ITC-court overlap (22.4% / 86.6% = 25.9%).
81. The remainder of ITC-PTAB overlapped patents that are in court litigation (24.5% – 2.0% = 22.5%) represents an overwhelming majority of the ITC-PTAB overlap (22.5% / 24.5% = 91.8%).
any two institutions tends to spill over into the third to a noteworthy extent. Yet until now, neither academic commentary nor public policy has taken an adequately systemwide view of substitution or of duplication in post-issuance reviews of patent validity.

C. Collective Action and Harassment in Patent Regulation

The problem of incomplete attention to substitution and duplication in PTAB proceedings also compounds inefficiencies left over from the pre-AIA world. Rather than replacing patent validity litigation in other forums, PTAB review interacts with those other forums to generate more complexity. Much of that complexity is rooted in the PTAB’s response to a collective action problem in generating patent invalidity judgments, a response that creates its own potential for harassment. There has been sparse judicial supervision (and, in some cases, none at all) over the PTAB’s own management of certain forms of harassment. All that remains is a backstop of political supervision, which creates its own ill effects.

The need for collective action against potentially invalid patents has been a problem for half a century. Since the Supreme Court decided *Blonder-Tongue Laboratories, Inc. v. University of Illinois Foundation* in 1971, a final judgment that a given patent is invalid creates defensive collateral estoppel upon the patent owner in future cases as against all parties, including nonmutual parties who were absent from the first case. In other words, the patent owner is foreclosed from relitigating that patent not only against the opponent who obtained the judgment of invalidity but also against any other parties whom the patent owner might later wish to sue for infringement. The Court’s judgment in this regard relied heavily on concerns that allowing patent owners to continue asserting their patents against others even after a court had found those patents invalid would reflect poor economic and legal policy.

Yet that decision also dramatically altered the incentives of would-be patent challengers. While the cost of invalidating a patent remains high—a fact that the *Blonder-Tongue* Court itself noted—nonmutual defensive estoppel against the patent owner means that the benefit of a successful invalidity judgment flows to everyone, not merely the challenger itself. In the language of public goods economics, a patent invalidity judgment is nonrival because each can consume it without detracting from another’s consumption of it, and the judgment is nonexcludable because it is not feasible to restrict the benefit to those who have paid for it. As a public good, something both nonrival and nonexcludable, a patent

83. 402 U.S. 313 (1971).
84. Id. at 350 (overruling Triplett v. Lowell, 297 U.S. 638 (1936)).
85. Id. at 334–48.
86. Id. at 338 (observing that “[i]n each successive suit the patentee enjoys the statutory presumption of validity, and so may easily put the alleged infringer to his expensive proof”).
invalidity judgment will tend to be undersupplied.\textsuperscript{88} Though collective action would be socially beneficial, each party’s preference to free ride on the efforts of others will often result in no action being taken at all.

The design of the PTAB, especially \textit{inter partes} review, alleviates this collective action problem in two ways. The most direct way is to reduce the cost of obtaining a patent invalidity judgment. To this end, the PTAB’s expertise and the resulting reductions in expense, delay, and error were a strong invitation for would-be patent challengers to challenge patents.\textsuperscript{89} The other way is to broaden access. \textit{Inter partes} review and post-grant review have no requirement of legal standing, so that anyone “who is not the owner of a patent” may seek agency review of a patent’s validity.\textsuperscript{90}

Naturally, there is no reason why these reduced barriers to entry should bring challenges only to invalid or questionable patents. They permit easier challenges across all patents and invite a variety of strategic behaviors. For example, after a 2013 jury verdict of $368 million against Apple, patent holder VirnetX was challenged in the PTAB by New Bay Capital.\textsuperscript{91} Prior to filing, New Bay had offered not to file its challenge if VirnetX agreed to pay it 10\% of the jury verdict, roughly $37 million.\textsuperscript{92} When VirnetX refused and New Bay challenged, VirnetX’s stock price tumbled by 25\%, reflecting a quarter-billion dollar loss.\textsuperscript{93} In fact, this stock price-based PTAB petitioning approach was systematized into an investment strategy by finance industry professionals, such as hedge fund manager Kyle Bass of Hayman Capital Management.\textsuperscript{94} Congress, in turn, responded with proposed legislation aimed at curbing financial trading based on PTAB petitions.\textsuperscript{95}

PTAB design has also invited strategic behaviors that involve larger numbers of parties and that have proven more persistent. The simplest form of this problem began as serial PTAB challenges targeting a given patent, spread across multiple petitions as well as across multiple challengers.\textsuperscript{96} For example, a majority of patents in a variety of technologies were the subject of multiple petitions, including chemical patents (60.6\%), computer and communication patents (50.9\%), and electrical patents (58.4\%).\textsuperscript{97}

\begin{thebibliography}{99}
\bibitem{89} See supra notes 41–45 and accompanying text.
\bibitem{90} 35 U.S.C. §§ 311(a), 321(a).
\bibitem{92} Id. at 932.
\bibitem{93} Id. at 933.
\bibitem{95} Innovation Act, H.R. 9, 114th Cong. § 9(b)(1)(C) (2015).
\bibitem{96} \textit{Strategic Decision Making}, supra note 2, at 68.
\bibitem{97} Id.
\end{thebibliography}
A more sophisticated form of this problem combines the incentives of district court defendants who come defensively to the PTAB with the incentives of preemptive challengers who make use of the PTAB’s lenient standing rules to attack patents before, or without ever, being sued. One way to evaluate this trend is to compare so-called standard and nonstandard petitioners as well as petitions. The “standard” model of substitution is for a PTAB petition to come as a defensive response to infringement litigation; the “nonstandard” model is for a PTAB petition to come preemptively.

Understood this way, the share of standard petitioners (individual challengers in the PTAB who have previously been sued on the patent that they now challenge) diverges systematically from the share of standard petitions (individual complaints filed in the PTAB that include at least one petitioner who was previously sued). Where the share of standard petitioners is lower than the share of standard petitions, this means that petitioners who were not previously district-court defendants are joining petitions filed by petitioners who were previously defendants. This petitioner-petition divergence is especially high among drug and medical-related patents (48.5% vs. 70.8%) and mechanical-related patents (53.1% vs. 70.2%).

While this joinder among patent challengers likely reflects beneficial collective action in a range of cases, it can also reflect harassment and delay, especially where this collective action comes in the form of subsequent PTAB petitions that are joined later to prior defendants’ PTAB petitions.

Importantly, the statutory design of the PTAB contains limits to police this sort of strategic behavior, but enforcement—especially judicial enforcement—of these limits has proven elusive. For petitions that arise from an underlying district-court litigation, the AIA imposes a one-year deadline from when the district court complaint is served upon a defendant to when that defendant must seek inter partes review in the PTAB. The Federal Circuit held in 2015 that PTAB decisions about the one-year bar were immune from judicial scrutiny, but that precedent was subsequently overturned en banc in 2018, imposing judicial review on PTAB decisions. Later still, the Supreme Court in 2020 reversed the Federal Circuit’s 2018 en banc precedent, concluding that PTAB determinations about the one-year bar are judicially unreviewable and reinstating the pre-2018 status quo. The applicability of this one-year time bar to joinder has correspondingly fluctuated over time. As a result of this uneven judicial supervision, the institutional backstop for supervising the PTAB’s management of strategic behavior has, unsurprisingly, been the political leadership of the USPTO. As a matter of administrative law and policy,

98. Id. at 74.
99. Id. at 49.
100. Id. at 74.
101. Id. at 74, 102–03.
102. Id. at 74–75.
103. 35 U.S.C. § 315(b).
106. Thryv, Inc. v. Click-to-Call Techs., LP, 140 S. Ct. 1367, 1372 (2020).
107. See Facebook, Inc. v. Windy City Innovations, LLC, 953 F.3d 1313, 1319 (Fed. Cir. 2020).
this may be familiar and even appropriate. Adjudication that is conducted in agency tribunals rather than in courts has long tended to vest final decision-making authority in politically appointed agency heads, and this has continued into modern administrative practice.\footnote{Christopher J. Walker & Melissa F. Wasserman, The New World of Agency Adjudication, 107 CALIF. L. REV. 141, 143–44 (2019); Ronald M. Levin, Administrative Judges and Agency Policy Development: The Koch Way, 22 WM. & MARY BILL RTS. J. 407, 412 (2013).}

However, this consensus view of agency design generally assumes an overtly regulatory posture that does not necessarily map onto patent law. Whereas the power and remit of most agencies are defined by reference to regulating some industry, activity, or other segment of society, the Patent Office has long been charged with the altogether different task of “handing out a small dollop of state power to dispersed parties in the form of individual property rights.”\footnote{Robert P. Merges, The Hamiltonian Origins of the U.S. Patent System, and Why They Matter Today, 104 IOWA L. REV. 2559, 2561 (2019).} This posture has long left the Patent Office disempowered from making substantive political judgments.\footnote{See Vishnubhakat, supra note 1, at 1674–75.} That state of affairs, especially the agency’s lack of rulemaking authority except as to procedural matters, remains in place even today.\footnote{See id.; Melissa F. Wasserman, The Changing Guard of Patent Law: Chevron Deference for the PTO, 54 WM. & MARY L. REV. 1959, 1973 (2013) (“The PTO, however, has not historically possessed the authority to engage in formal adjudication or rule making—the two formal procedures that Mead indicates would likely warrant deference.” (citing United States v. Mead Corp., 533 U.S. 218, 229-31 (2001)).} This means that the political supervision of the PTAB is, at best, an uneasy fit with the prevailing view of patents as property rights.\footnote{Vishnubhakat, supra note 1, at 1721.} The stability and certainty that are of chief concern in a well-functioning patent regime require mediating between predictable but short-lived rules and unpredictable but durable standards.\footnote{Id. at 1721–22.} This is especially true where, as in patent law, judicially developed doctrines change so quickly that they cannot offer certainty even over the twenty-year life of a patent.\footnote{John F. Duffy, Rules and Standards on the Forefront of Patentability, 51 WM. & MARY L. REV. 609, 614–15 (2009).} Closely connected to the need for stability and certainty in patent rights is the broader concern that an unduly political valence will tend to undermine the ability of the Patent Office to make credible commitments to innovators and investors.\footnote{Vishnubhakat, supra note 1, at 1726–27.}

Indeed, the problem of credible commitment is independent of any particular agency administration and independent of the merits of any particular Patent Office policy initiative. For example, patent owners who perceived the PTAB in its early years as being overly lenient with petitioners have probably welcomed the more recent USPTO policies of limiting serial petitions under the General Plastic Indus. Co. v. Canon Kabushiki Kaisha framework\footnote{General Plastic Indus. Co. v. Canon Kabushiki Kaisha, No. IPR2016-01357 (P.T.A.B. Sept. 6, 2017) (designated by the PTAB as “precedential” on Oct. 18, 2017).} and of notice-and-comment rulemaking to tighten
the standard for construing patent claims in PTAB review.\textsuperscript{117} Yet what today’s agency leadership can give, tomorrow’s can take away. A recent USPTO request for information, which was issued during the transition from the Trump administration to the Biden administration and which has proven especially controversial, illustrates the point well. The agency during 2019 and 2020 designated two decisions, \textit{NHK Spring Co. v. Intri-Plex Technologies, Inc.} and \textit{Apple, Inc. v. Fintiv, Inc.}, as precedential and binding upon PTAB decisions.\textsuperscript{118} The resulting \textit{NHK-Fintiv} doctrine represents a policy of denying PTAB petitions where a parallel district court proceeding is so far along and so substantially similar in art and argumentation that it would be best to conserve USPTO resources rather than undertake a largely or entirely duplicative review.\textsuperscript{119} In fact, the \textit{NHK-Fintiv} rule represents one of the most clearly specified positions by the USPTO with respect to the substitutionary purpose of the PTAB.

Given the late stage of the Trump administration during which this proposal was published for comment, the political salience of the \textit{NHK-Fintiv} rule is especially high. As a result, the Biden administration now seems highly likely at least to revisit the issue and potentially to reverse course in accordance with newly prevailing policy priorities once those are announced. Thus, regardless of where one stands in the debate over the policy itself, what ultimately matters is the political reversibility and fragility of the \textit{NHK-Fintiv} rule.

In short, the experience of the PTAB has been characterized by a persistent potential for harassment and delay as a flip side of beneficial collective action, by a lack of reliable judicial supervision over court-agency substitution, and by a backstop of political supervision that is equally unstable. Even on their own, these features of the PTAB represent a new and complex set of institutional design challenges for the patent system. But what is worse, the significant court-agency duplication between the PTAB and other adjudicatory systems compound pre-AIA problems and generate more complexity, not less.

\section*{II. The Arc of Duplication: Causes and Reforms}

The problems of incomplete substitution, outright duplication, and strategic usage of PTAB proceedings for problematic purposes were no doubt unintended by Congress, but they were not unforeseeable. This Part explores the nature and causes of the problems that Part I documented and described. The first set of explanations all arise from the statutory design of the AIA itself and focus on the \textit{inter partes} review statute both because it is representative of the issues and because the body of \textit{inter partes} review proceedings accounts for the large majority of the PTAB’s operational workload and policy levers. The second set of explanations are more structural and track the institutional ascendency of the USPTO since the AIA was

\begin{itemize}
\item \textsuperscript{118} No. IPR2018-00752 (P.T.A.B. Sept. 12, 2018); No. IPR2020-00019 (P.T.A.B. Mar. 20, 2020).
\item \textsuperscript{119} See infra Section II.C.2 (discussing the \textit{NHK-Fintiv} doctrine as an agency response to inter-branch duplication).
\end{itemize}
enacted. This latter set of effects do not flow unavoidably from the AIA’s statutory
design but do reflect well-understood patterns of agency behavior, especially agency
aggrandizement.

A. The AIA’s Statutory Design

The AIA’s statutory design affects duplication in PTAB litigation in four ways.
First is through the burden of proof that the PTAB requires, which is weaker than the
burden that Article III courts require. Second is asymmetric estoppel that fosters
disregard by the USPTO of prior judicial findings about validity even of the same
patents. Third is the lack of a clear statutory structure for staying litigation in federal
courts pending the outcome of parallel proceedings in the PTAB, resulting in high
variation in the rate at which judicial stays are actually granted. Fourth is the AIA’s
inattentive view of the ITC, which leaves important gaps with respect to a significant
institutional actor in the adjudication of patents.

1. The Agency’s Weaker Burden of Proof

The most persistent structural reason for duplicative litigation in the PTAB is its
weaker burden of proof. Petitioners must prove claims of unpatentability by a
preponderance of the evidence.120 This is in contrast to the more stringent burden of
clear and convincing evidence that a litigant must meet to overcome the statutory
presumption of patent validity121 and prove unpatentability, both in the U.S. district
courts122 and in the ITC.123 The doctrinal effect of this disparity is significant, as it
permits open disregard by administrative patent judges not only to disregard the
decisions of Article III judges even with respect to the same patents but also to reach
altogether conflicting judgments of their own in the PTAB (though the conflict in
underlying standards themselves means that these contrary judgments are not
inconsistent in the legal sense). Meanwhile, the origins of this disparity in the
legislative history and in the underlying assumptions of the AIA have also been a
point of considerable confusion. It is useful to consider these issues separately.

The doctrinal effect of these divergent standards became clear immediately. After
the AIA trial proceedings went online in September 2012 and as the earliest final
written decisions began to issue in June 2013, the PTAB observed in SAP America,
Inc. v. Versata Development Group, Inc. that neither the statutory presumption of
validity nor the burden of clear and convincing evidence applied to petitioners in the
PTAB.124 The PTAB in Google, Inc. v. Whitserve LLC put it even more clearly,
explaining that “[g]iven these differing evidentiary standards, the same evidence that
the jury found insufficient could meet the burden of proof in an inter partes
review.”125 The PTAB further observed that even conclusions that the Federal Circuit

120. 35 U.S.C. § 315(e).
121. 35 U.S.C. § 282(a) ("A patent shall be presumed valid.").
reached about patent validity under the clear and convincing evidence standard would not foreclose a contrary PTAB conclusion about the same patent under the preponderance standard.\textsuperscript{126}

The Federal Circuit has largely agreed as a descriptive matter, though the implications of this divergence remain contested. The court in \textit{Novartis AG v. Noven Pharmaceuticals, Inc.} agreed that the differing burdens of proof mean “that the PTAB properly may reach a different conclusion based on the same evidence.”\textsuperscript{127} Writing separately in \textit{Smartflash LLC v. Apple Inc.}, Judge Newman further recognized that district court judgments about patent validity are also often “accompanied by the deferential standard applied to jury verdicts.”\textsuperscript{128}

The similarly deferential standard of substantial evidence review that the Federal Circuit applies to final written decisions of the PTAB\textsuperscript{129} entrenches the divergent burdens even further and leaves little likelihood that appellate correction might close the gap between conflicting judgments from the PTAB and the courts. Judge Newman also elaborated on this concern in dissenting from the denial of rehearing en banc in \textit{Merck & Cie v. Gnosis S.P.A.}\textsuperscript{130} There, she urged that the substantial evidence that can support an agency determination under the judicial review standards of the Administrative Procedure Act is inapt for the purposefully trial-like structure of AIA proceedings, especially in light of the increased risk to settled rights and future investments from, among other things, the PTAB’s more lenient burden of proof.\textsuperscript{131}

The origins of this disparity reflect the need for reconsideration and reform as well. The use of a weaker preponderance-level burden saw strong opposition throughout the years-long legislative debate that would lead to the AIA.\textsuperscript{132} Notably, this opposition was closely tied to the same concerns about duplicative litigation that now confront the PTAB.\textsuperscript{133} To this, one may fairly respond that regardless of what substantive preferences were sought during the legislative debate, the final legislative choice is what controls. Yet the problem goes deeper.

The heart of the issue is summarized well in Joe Matal’s authoritative two-volume \textit{Guide to the Legislative History of the AIA}: “there simply is no requirement in the

\textsuperscript{126} \textit{Id.} (arguing that “even if the Federal Circuit had definitively concluded [the Schrader prior art reference] does not teach [the disputed technological concept of “Internet-based data”] under the clear and convincing standard, it would not foreclose our determination that Schrader meets the preponderance standard.”).

\textsuperscript{127} 853 F.3d 1289, 1294 (Fed. Cir. 2017).

\textsuperscript{128} 621 F. App’x 995, 1006 (Fed. Cir. 2015) (Newman, J., concurring in part, dissenting in part).

\textsuperscript{129} Merck & Cie v. Gnosis S.P.A., 808 F.3d 829, 833 (Fed. Cir. 2015).

\textsuperscript{130} 820 F.3d 432, 436 (Fed. Cir. 2016) (Newman, J., dissenting).

\textsuperscript{131} \textit{Id.} at 438.

\textsuperscript{132} \textit{See, e.g., Patent Reform Act of 2007: Hearing on H.R. 1908 Before the Subcomm. On Courts, Intellectual Prop., and the Internet, 110th Cong. 56, 55 (2007) (statement of Gary L. Griswold, President and Chief Counsel of Intellectual Property, 3M Innovative Properties) (arguing that it is unfair to patent owners to allow patents to be challenged many years after the fact under only a preponderance burden of proof on the basis of uncertain events such as “public use and oral disclosures”).

\textsuperscript{133} \textit{Id.} at 55 (warning of “serial post-grant challenges” and the loss of any “right to expect quiet title at some point without facing an endless series of challenges”).
law that the USPTO give any deference or weight to a district court’s decision with respect to a patent.” 134 This point arose in the context of precluding the USPTO, as a matter of formal collateral estoppel, from reconsidering the same issues in its administrative proceedings that courts have already decided. 135 Strikingly, the reason was that administrative proceedings in the USPTO were not merely subject to different standards and burdens but were also fundamentally different in kind: unlike civil litigation, the argument went, “the examiner [or reexaminer] is not attacking the validity of the patent but is conducting a subjective examination of the claims in light of prior art.” 136

This supposed qualitative difference is starkly at odds with the prevailing view of the AIA proceedings as adjudicative in nature rather than:

One important structural change made by the present bill is that inter partes reexamination is converted into an adjudicative proceeding in which the petitioner, rather than the Office, bears the burden of showing unpatentability. . . . In the present bill, section 316(a)(4) gives the Office discretion in prescribing regulations governing the new proceeding. The Office has made clear that it will use this discretion to convert inter partes into an adjudicative proceeding. This change also is effectively compelled by new section 316(e), which assigns to the petitioner the burden of proving a proposition of unpatentability by a preponderance of the evidence. 137

In short, the AIA’s burden of proof reflects a poorly designed attempt to have it both ways. The statutory text is clear about the preponderance standard, but the policy behind it arises from a conflicted understanding about the purpose of AIA proceedings as adversarial trial-type substitutes for traditional litigation. This confusion does not merely mean that a patent that would have survived Article III court review may fall to a weaker evidentiary showing in the PTAB. It means that even a patent that actually did survive Article III review may nevertheless be relitigated in a duplicative PTAB trial on the same evidence and be struck down.

2. Estoppel and Disregard for Judicial Findings

The divergent burden of proof in the PTAB also contributes to related problems of estoppel and disregard for judicial findings. To be sure, there are reasons why these tendencies toward duplication are not necessarily surprising, but these reasons still do not adequately address the AIA’s basic aim of patent error-correction through expert administrative reevaluation.

The tendency toward duplication arises in part in the Blonder-Tongue doctrine. Ever since the Supreme Court decided Blonder-Tongue Laboratories, Inc. v.

135. Id. at 638–39.
136. Id. at 640 (citing In re Swanson, 540 F.3d 1368, 1377–78 (Fed. Cir. 2008)).
University of Illinois Foundation, a final judgment that a given patent is invalid creates defensive collateral estoppel effect upon the patent owner in future cases as against all parties, including nonmutual parties who were absent from the first case. In other words, the patent owner is foreclosed from relitigating that patent not only against the opponent who obtained the judgment of invalidity but also against any other parties whom the patent owner might later wish to sue for infringement. The reverse is not true: a patent owner who prevails on the issue of patent validity against one opposing party cannot use that finding to estop other, nonmutual parties from arguing anew that the patent is invalid. In particular, this is because issued patents are, by statute, presumed valid and must be proven invalid by clear and convincing evidence. Thus, as a formal matter, no institution ever actually finds a patent to be valid but, at most, merely not invalid. A finding that the presumption of validity was not overcome in a given case leaves open the possibility that it may be overcome in a subsequent case.

What is even more fundamental, however, than the estoppel principles that govern patent validity is that the nature and origins of post-issuance review following the AIA are themselves rooted in an intractable premise. In order to make it easier to invalidate questionable patents, Congress had no choice but to make it easier to invalidate all patents: after all, if it were already known which patents are of questionable quality and which ones have merit, then post-issuance review would be unnecessary. So, it may not be altogether surprising that the effectiveness of AIA patent validity review proceedings comes, at least in part, precisely from the greater ease with which they allow patents to be struck down. Perhaps that is the very point.

But even these accounts do not justify the current state of duplication. For one thing, if the true thrust of substituting administrative review for review by other institutions was actually to systematically weaken the survival rate of patents, then the rhetoric of Patent Office expertise that drove the AIA’s bipartisan and bicameral consensus rings hollow. For another, the background principles of common law estoppel that have been in place since Blonder-Tongue are not the only relevant forces at work in the structure of post-issuance review under the AIA.

In fact, the substitutionary choice that the AIA imposes—between the PTAB on one hand and Article I and Article III tribunals on the other hand—comprise a number of statutory provisions that, taken together, form an emergent institutional border around the PTAB. But by its own terms, the border was never absolute and left room for duplicative litigation. What is worse, a series of policy-driven encroachments by the Patent Office did much to weaken the border even further in the initial years after the AIA, especially between the PTAB and the federal courts.

139. Id. at 350 (overruling Triplett v. Lowell, 297 U.S. 638 (1936)).
140. A few scholars have discussed this scenario in analytical terms. See, e.g., Paul R. Gugliuzza, (Invalid Patents, 92 NOTRE DAME L. REV. 271, 324–325 (2016).
141. 35 U.S.C. § 282(a); Microsoft Corp. v. i4i Ltd. P’ship, 564 U.S. 91 (2011).
142. Vishnubhat, FTC Testimony, supra note 5.
144. Id. at 1080.
145. Id. at 1087–91.
Recent retrenchments by the Federal Circuit and the Supreme Court, as well as by the Patent Office itself, suggest that the time is ripe for reform. The proper focus of these reforms should be to minimize duplication, especially given the susceptibility of the PTAB to arbitrage of divergent legal standards and other strategic behavior.

The current structure of IPR proceedings offers four distinct dimensions along which to conceptualize substitution, and along which to minimize duplication:

1. the identity of the party whose action triggers a choice between the PTAB and some other institution;
2. the relevant legal action that triggers the choice;
3. the particular institution with respect to which the substitutionary choice is being triggered; and
4. the timing of the triggering action relative to the PTAB proceeding.

The IPR statute translates each of these into concrete terms. One example is (1) a would-be IPR petitioner (2) who has filed a civil action challenging the validity of a patent (3) in U.S. district court (4) before seeking IPR. Such a party is barred altogether from pursuing the IPR. Another example is (1) a patent owner (2) who has filed a lawsuit alleging infringement of the patent (3) and has done so in U.S. district court (4) before an IPR has been filed. In such a case, the accused infringer must seek IPR within one year of being served with the complaint or else cannot seek IPR at all.

Still another example is (1) a former IPR petitioner (2) who has litigated an IPR to final judgment but now wishes (3) either to return to the PTAB or to go into the U.S. district courts or into the ITC and (4) file a subsequent challenge on the same patent claim. In such a case, the former petitioner is foreclosed from relitigating that claim before the PTAB on any ground that was actually raised or reasonably could have been raised during the first IPR. The former petitioner is likewise foreclosed from relitigating before the U.S. district courts and the ITC as well.

Still, important gaps remain. Although a patent owner’s infringement suit in the district courts triggers a one-year time bar for seeking IPR, a number of PTAB precedents now hold that a patent owner’s analogous infringement action in the ITC does not trigger the one-year time bar. Likewise, although prior final judgments in the PTAB create estoppel both inside and outside the PTAB, prior proceedings in the district courts and in the ITC have no preclusive effect and do not foreclose any relitigation in the PTAB. This is the very situation that has allowed the above-discussed duplication to proliferate.

The problem of asymmetric estoppel, of course, is internally consistent with other problematic structural choices embedded in the AIA and does much to compound them. As discussed above, divergent burdens of proof mean that evidence of

146. 35 U.S.C. § 315(a).
147. Id. § 315(b).
148. Id. § 315(e)(1).
149. Id. § 315(e)(2).
invalidity may fail to be clear and convincing but still establish a preponderance.\textsuperscript{151} Similarly, as discussed below, the need for patent claims to be construed in the PTAB after their meaning has already been established in court has been an important form of duplication as well.\textsuperscript{152} Filling these gaps and eliminating these disparities between the PTAB and its adjacent institutions would do much to minimize duplication.

3. The Moving Target of Judicial Stays

Thus far, the discussion of what sources produce duplication in AIA trial proceedings has focused on statutory provisions that shape the design of the PTAB. It has also emphasized, though not assumed, a situation where the competing institution has already concluded its reevaluation of the patent, perhaps years ago. However, substitution across the court-agency boundary also requires considering the other side of that boundary, especially where the competing institution has not yet completed or perhaps even begun its own reevaluation of the patent. In other words, while it is important to question PTAB duplication of already-resolved questions of patent validity, it is also important to address the potential for duplication in concurrent proceedings. Indeed, as time goes on and older pre-AIA patents expire, an ever-increasing share of patent litigation will take place in the shadow of immediate recourse to the PTAB.\textsuperscript{153}

This means that where petitions for administrative revocation come before the PTAB while the relevant patents are also involved in co-pending litigation elsewhere, a concern for minimizing duplication must also look to how the competing institutions respond to the PTAB’s actions. Where the PTAB decides not to undertake review on the merits, duplication is not a concern, but where the PTAB does institute review, the question of judicial stays comes to the fore.\textsuperscript{154}

As described already, the rates at which stays are granted across judicial districts and over time are highly variable.\textsuperscript{155} These grant rates have ranged from steady increases reaching upwards of 72% in the Northern District of California to steady declines reaching lows of 11% in the Eastern District of Texas.\textsuperscript{156} The origins of this variation lie in two important and intertwined sources. One is the nonrandom

\begin{itemize}
  \item \textsuperscript{151} See supra Section II.A.1.
  \item \textsuperscript{152} See infra Section II.B.2.
  \item \textsuperscript{153} Empirical trends in PTAB litigation bear this out. Patents issued under the post-AIA first-inventor-to-file regime have already grown to account for more than 37.5\% of inter partes review petitions. Mixed Case, supra note 59, at 520, 535. Meanwhile, challenges against patents issued before the AIA but subject to the earlier-generation inter partes reexamination have fallen from a high of about 75\% early in the life of the PTAB to roughly 50\%. Id. Patents issued prior even to inter partes reexamination (but still susceptible to inter partes review) have fallen even further, from 35.1\% at the outset of AIA to about 10\%. Id.
  \item \textsuperscript{154} Although stays in the ITC pending PTAB review could also theoretically be a concern, the rapid pace of ITC litigation makes such stays largely unnecessary. ITC litigation data reveals that in over eight years of AIA trial proceedings, only five motions were even filed to stay a Section 337 action in the ITC pending inter partes review: three were denied, one was denied as moot, and only one was granted. See generally Docket Navigator, https://brochure.docketnavigator.com/ [https://perma.cc/74WA-AUEP].
  \item \textsuperscript{155} See supra Section I.B.2.
  \item \textsuperscript{156} See supra Section I.B.2.
\end{itemize}
choice—indeed, the highly strategic and outcome-sensitive choice—of judicial districts in which plaintiffs bring patent cases. The other is the quality of information about patent validity that courts can expect to get from the resolution of a PTAB review.

As to the choice of forum, it is not necessarily the case that judges in some districts are inherently more inclined to grant stays while judges elsewhere are inclined to deny them, though this is a possibility in certain contexts. For example, a number of commentators have argued that the Eastern District of Texas, which for a time was the leading venue for patent case filings in the United States, has been unreasonably stringent in its past treatment of motions to transfer cases to other venues, purportedly out of a desire to attract patent cases to east Texas and, once filed, to keep them there. One might expect the same premises and criticisms to be levied against that court’s treatment of stays pending PTAB review.

Alternative explanations may also be at work. One is technology-specificity. Patent cases are likely to provoke different responses to motions for stays depending on what technology is at issue. For example, the patent-to-product ratio for high-technology products such as smartphones tends to be quite high, with numerous subcomponents potentially susceptible to patent protection and so is more likely to present greater complexity as to the number of patents and claims involved in any given lawsuit. The ratio for pharmaceutical drugs, by contrast, is quite small and so is correspondingly likely to present less litigation complexity. These considerations of litigation complexity, in turn, are highly likely to affect how much of the dispute is really likely to be resolved by a co-pending administrative revocation proceeding in the PTAB, and how productive a stay would be.

Indeed, it has long been conventional wisdom that certain districts are more likely to attract certain types of patent cases based on the technologies that underlie the patents in suit, such as high-technology patents to the Northern District of California due to its geographical coverage of Silicon Valley or drug and patents to the District of Delaware due to the high number of pharmaceutical companies based in the Delaware and New Jersey region. Supervening choices may overcome those

---

157. See supra note 34 and accompanying text (discussing forum shopping in patent litigation).
158. See infra Section II.B.3.
162. See, e.g., Sag, supra note 40, at 1087; Fromer, supra note 18, at 1466 (citing Kimberly A. Moore, Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?, 79 N.C. L. REV. 889, 934 (2001)).
preferences, of course, such as the high rate at which plaintiffs chose not to assert high technology patents in the Northern District of California (perceived as a defendant-friendly venue) but rather in the Eastern District of Texas.\textsuperscript{164}

Still, empirical analysis bears out the conventional wisdom, including the rise and fall of the Eastern District of Texas. The 2017 \textit{TC Heartland} decision of the Supreme Court held that under the special venue statute for patent cases, a defendant does not “reside” simply anywhere it would be subject to personal jurisdiction, but only where it is incorporated.\textsuperscript{165} The immediate effect of the decision was to make the Eastern District of Texas an untenable venue for the great majority of patent cases filed there, especially in high technology.\textsuperscript{166}

As Figures 2–4 show, the Eastern District of Texas’s share of assertions involving high technology patents, relative to other high-volume patent districts, rose in the early 2010s and then declined—for Computers & Communications-related patents, Electrical and Electronic-related patents, and Mechanical patents.\textsuperscript{167} The competing districts that had initially lost these cases and then regained them were, indeed, the Northern District of California (where high-technology defendant companies are commonly headquartered) and the District of Delaware (where they are commonly incorporated). Meanwhile, as Figures 5–6 show, the District of Delaware has remained the predominant venue for Chemical-related patents as well as Drugs & Medical-related patents, with little impact from the Eastern District of Texas either before or after \textit{TC Heartland}. Thus, there is good reason to expect that technology-specific dynamics of patent case management—which are highly correlated with the choice of forum—will affect the practical value of staying the litigation.

Beyond forum-specific influences on the granting of stays, there is also a question of what information the court that stays its hand can expect to get from a co-pending PTAB review. As noted already in the discussion of technology-specific dynamics, considerations of litigation complexity are part of this calculus. A large and complicated lawsuit may provoke a varying range of PTAB challenges, involving a large number of \textit{inter partes} review petitions versus only a few, involving many or all of the defendants versus only one or a few, and so on. To this extent, defendant behavior can exert a fair degree of influence over the conditions that will make a stay more or less likely.

More importantly, however, a controversial practice of the PTAB itself has also played a role in the granting and denying of judicial stays. That is the practice of partial adjudication, which did much to undermine the substitutionary purpose of the PTAB.\textsuperscript{168} When the USPTO first established its procedures in 2012 for the new AIA

\begin{flushleft}
\end{flushleft}

\textsuperscript{164} \textit{See generally} Anderson, \textit{supra} note 159.


\textsuperscript{166} \textit{See generally} Anderson, \textit{supra} note 159.

\textsuperscript{167} The technology categories used here correspond to those set forth in the National Bureau of Economic Research taxonomy. \textit{See generally} Hall, Jaffe & Trajtenberg, \textit{supra} note 31.

\textsuperscript{168} \textit{See infra} Section II.B.3.
trial proceedings, it announced that after considering a petition for review, the presiding PTAB panel “may authorize the review to proceed on all or some of the challenged claims and on all or some of the grounds of unpatentability asserted for each claim.” This provoked controversy because the relevant statutory language of the AIA appeared to be binary, contemplating either a full institution or a full denial, but not a power of partial adjudication.

The operation of this practice upon judicial stays was straightforward. A full denial of review by the PTAB would signal to courts with co-pending litigation that they were free to proceed, as there was no “reasonable likelihood” that the patent would successfully be declared invalid. A full grant of review by the PTAB would signal to the courts that there was value, perhaps a good deal of it, in holding litigation in abeyance because PTAB revocations of the patents in suit might promote settlement or might even render the court proceedings irrelevant altogether. Yet only a partial grant of review, accepting some arguments as to some patent claims while rejecting others, would be different. There, “the degree to which a PTAB judgment would simplify litigation would necessarily be incremental” and thus less attractive a case for staying the litigation.

Meanwhile, the statutory text of the AIA is silent as to judicial stay in the context of inter partes review, though the statute does specify how courts must consider stays pending the outcome of covered business method reviews. In the face of statutory silence as to inter partes review, courts have resorted to a standard of simplifying the issues involved, avoiding the waste of costs that have already been sunk into the litigation, and minimizing prejudice to the nonmoving party. This highly flexible standard, in turn, is unlikely to generate consistency because it is committed to the sound discretion of the trial court and, as a result, is subject to highly deferential appellate review.

Thus, the court- and district-specific variations that already characterize patent litigation to a significant extent are compounded even further as to the practice of

170. 37 C.F.R. § 42.108(a) (2020) (emphasis added).
171. 35 U.S.C. § 314(d) (referring to a “determination by the Director whether to institute an inter partes review”) (emphasis added).
175. See Vishnubhakat, supra note 173, at 2674.
176. Id.
179. Id.; see also Vishnubhakat, supra note 173, at 2674.
granting stays, with little hope for consistency or uniformity. The result is even greater potential for court-agency duplication.

4. The Largely Overlooked ITC

Finally, beyond the statutory choices reflected in the preponderance burden of proof and of asymmetric estoppel and the statutory silence on judicial stays in *inter partes* review, the AIA gives inadequate attention to International Trade Commission proceedings in all but two important respects. One is that the statute does expressly include section 337 actions in the ITC within the scope of the estoppel that *inter partes* review,180 post-grant review,181 and covered business method review182 variously generate. The PTAB-ITC interface is not immune to the problems of asymmetry in PTAB estoppel,183 but ITC actions are treated the same as civil actions in the Article III courts. The other is that PTAB review is barred by a petitioner’s prior challenge to the validity of a claim of the patent only if that challenge was made in a prior civil action.184 The omission of prior ITC actions from this statutory bar is not from inattention but that declaratory judgments of patent invalidity are simply not available in the ITC as they are in the Article III courts.

Still, apart from these parameters, the AIA fails to address ITC-related dynamics in two other respects that are, indeed, salient to the problem of duplicative litigation. One is the trigger of *inter partes* review’s one-year time bar. A defendant who is sued for patent infringement in a U.S. district court action must seek *inter partes* review on the relevant patents within one year of being served with a complaint alleging infringement, or else must forgo *inter partes* review,185 but no comparable provision exists for the ITC, leaving the text unclear whether ITC actions trigger the one-year bar. The USPTO has interpreted the statute narrowly, holding in PTAB decisions that ITC complaints do not trigger the one-year bar.186 However, the Federal Circuit has not yet ruled one way or the other on the USPTO’s view. As a result, an important constraint on court-PTAB duplication, the one-year bar, carries only uncertain force as a constraint on potential ITC-PTAB duplication.

180. 35 U.S.C. § 315(e)(2) (providing that, after a final written decision in *inter partes* review, the petitioner, its real party in interest, or its privy are foreclosed from raising patent validity arguments that they actually raised—or reasonably could have raised—in any subsequent civil action “or in a proceeding before the International Trade Commission under section 337 of the Tariff Act of 1930.”).

181. 35 U.S.C. § 325(e)(2) (same as to post-grant review).

182. Leahy-Smith America Invents Act § 18(a)(1)(D), 125 Stat. at 330 (same as to covered business method review, except that only arguments that were actually raised before the PTAB are subject to estoppel).

183. See supra Section II.A.2.

184. 35 U.S.C. § 315(a)(1) (establishing the bar for *inter partes* review); id. § 325(a)(1) (same as to post-grant review and covered business method review).

185. 35 U.S.C. § 315(b). The prohibition also applies if the defendant is a privy or real party in interest of the would-be petitioner. Id.

The other problem is also a trigger, but one that allows PTAB review to go forward in certain situations. The AIA’s transitional program for covered business method review broadly allowed challenges based on all grounds for patentability just as post-grant review does, and also broadly covered pre-AIA patents with retroactive effect— but, unlike inter partes review, was limited to a defensive posture. “A person may not file a petition for a transitional proceeding with respect to a covered business method patent unless the person or the person’s real party in interest or privy has been sued for infringement of the patent or has been charged with infringement under that patent.” This has left gaps as yet unaddressed by the Federal Circuit. The USPTO has interpreted the standing requirement of covered business method review as requiring either (1) that the petitioner, its privy, or its real party in interest has already been sued; or (2) that the petitioner is at imminent enough risk of suit that it would have standing to bring a declaratory judgment in an Article III court. What remains unclear, however, is whether being sued in the ITC would by itself be enough to support standing for a covered business method review. It is certainly plausible that the filing of an ITC action would give rise to Article III declaratory judgment standing under the USPTO’s regulatory interpretation—and that the resulting declaratory judgment standing would be enough to justify covered business method review. However, as with the applicability of inter partes review’s one-year time bar to ITC complaints, it remains unclear whether the USPTO’s interpretation survives judicial review.

The Federal Circuit did hold in Return Mail v. USPS that the standing requirement of covered business method review is broad enough to embrace a suit against the U.S. government in the U.S. Court of Federal Claims, so that a federal agency can properly seek recourse in the PTAB. Like the ITC, the Court of Federal Claims is an Article I tribunal, suggesting an apt precedent. However, the question in Return Mail was not whether a complaint in an Article I tribunal could support covered business method review but whether the U.S. government was a “person” who was capable of being sued in the relevant fashion. And even on that question, the Supreme Court ultimately said no.

These arguments are not to say that duplication between the PTAB and ITC is equally as frequent or problematic as duplication between the PTAB and the courts. The case against court-PTAB duplication is stronger and deserves principal focus. Moreover, the recent sunset of covered business method review in September 2020 after its initially prescribed eight-year period suggests that it is a less urgent legislative priority. Still, the larger point remains that among the courts, the ITC, and the PTAB, the overlap between any two institutions tends to spill over into the third—and that any proper systemwide view of substitution and duplication in adjudicating patent validity must take greater account of the ITC.

187. Strategic Decision Making, supra note 2, at 60.
189. 37 C.F.R. § 42.302(a).
193. Id. at 1858–59.
B. The Patent Office’s Ascendancy

Turning now from statutory design effects to institutional effects from the USPTO itself, there are four additional trends to consider. The first of these is the agency’s judicially unreviewable power to grant or deny petitions for AIA trial proceedings. Second is the standard that the USPTO initially adopted for construing patent claims in PTAB trials, a standard that diverged from that of the U.S. district courts and the ITC. Third is the agency’s past practice of partial adjudication, which was partly addressed above in the context of judicial stays. Fourth is another instance of judicially unreviewable power, this time to police the one-year statutory deadline for seeking *inter partes* review.

All of these trends implicate the problem of PTAB duplication, though in some cases, their actual effects on the problem are mixed. Moreover, some of these trends have already provoked judicial and regulatory corrections, though these corrections have also been incomplete.

1. The Agency’s Unreviewable Screening of Petitions

The absolute discretion that the USPTO exercises in granting and denying petitions foregrounds much else that bears on substitution and duplication in the PTAB. As a result, this powerful threshold discretion has a highly mixed valence, at times contributing to the problem of duplication but at other times enabling significant course correction and reform. This is due in large part to the irreducibly political character of that discretion, a fact that has also received recent attention in the debate over agency aggrandizement and the growing power of the USPTO vis-à-vis the Article III courts. The discussion here is more limited, focusing on the role that agency discretion in screening petitions for PTAB review plays specifically in the analysis of substitution and duplication.

In the first instance, the discretion to grant or deny petitions for *inter partes* review, post-grant review, and covered business method review resides with the Director of the USPTO. For reasons of convenience, the Director does not exercise this power directly but has subdelegated it by regulation to the PTAB. Once a petition is granted and review is instituted, the PTAB also exercises the power to conduct the actual trials and adjudicate the cases to a final written decision, and this power resides directly with the PTAB. The agency’s current practice is that the...


196. 35 U.S.C. §§ 314(a), 315(a), 315(b) (setting the thresholds for *inter partes* review as limits on what the “Director” may or may not authorize); §§ 324(a), 324(b), 324(c), 325(c) (same as to post-grant review and covered business method review); *see also* Vishnubhat, *supra* note 1, at 1749–51 (discussing the allocation of pre-institution screening and post-institution adjudication power in the USPTO between the Director and the PTAB).

197. 37 C.F.R. § 42.4(a) (“The Board institutes the trial on behalf of the Director.”).

198. *See* 35 U.S.C. § 318(a) (delegating to “the Patent Trial and Appeal Board” the authority and obligation to issue final written decisions in *inter partes* reviews); *see also* id. § 328(a) (same as to post-grant reviews and covered business method reviews).
The availability of judicial review falls precisely along this divide between the powers of the Director and those of the PTAB. Initial determinations whether to institute review are “final and nonappealable.”\textsuperscript{200} However, assuming the PTAB does institute review, cases that reach final written decisions are appealable to the Federal Circuit.\textsuperscript{201} This dichotomy presented the first major test of USPTO power under the AIA and, since the Supreme Court’s resolution of that test, has exerted profound influence on whether safeguards against duplicative litigation will be left to the agency’s political discretion or will be vindicated in the courts. That first test came in \textit{Cuozzo Speed Technologies v. Lee},\textsuperscript{202} which presented two questions. The more operational of the two was the standard for construing patent claims in PTAB trials and is addressed later in this Section.\textsuperscript{203}

The foundational issue in \textit{Cuozzo} was how far the nonappealability provision of § 314(d) extended.\textsuperscript{204} The more modest view of this provision was that litigants could not obtain interlocutory review of the decision to accept a petition and proceed.\textsuperscript{205} On this view, the right of a patent owner not to be subjected to a questionable challenge was not legally recognized.\textsuperscript{206} Yet if and when the case reached a final written decision, review would, indeed, be available as to all issues based “on the basic administrative law principle that intermediate issues merge into an agency’s final order on the merits.”\textsuperscript{207} Where the PTAB errs in granting review, the court would not take up the question immediately and risk systemic disruption, but any such error would eventually be susceptible to appellate correction based on a full administrative record. This view did not prevail.

The more expansive view was that § 314(d) barred not only interlocutory review but all judicial review of the threshold screening function, even after the case reached final judgment.\textsuperscript{208} This was also the USPTO’s view of its own powers\textsuperscript{209} and would

\textsuperscript{199.} See Ethicon Endo-Surgery, Inc. v. Covidien LP, 812 F.3d 1023, 1028 (Fed. Cir. 2016) (noting that the Patent Office “has determined that, in the interest of efficiency, the decision to institute and the final decision should be made by the same Board panel”).

\textsuperscript{200.} 35 U.S.C. §§ 314(d), 324(e).

\textsuperscript{201.} 35 U.S.C. §§ 319, 329, 141(c).

\textsuperscript{202.} 136 S. Ct. 2131 (2016).

\textsuperscript{203.} \textit{See infra} Section II.B.2.


\textsuperscript{205.} Brief for Petitioner at 46, \textit{Cuozzo}, 136 S. Ct. 2131 (No. 15-446), 2016 WL 737452, at *46. The dissent below took this view as well. \textit{See In re Cuozzo Speed Techs., LLC}, 793 F.3d 1268, 1291 (Fed. Cir. 2015) (Newman, J., dissenting) (“The stated purpose of the ‘final and nonappealable’ provision is to control interlocutory delay and harassing filings.”).

\textsuperscript{206.} \textit{See Vishnubhatak, supra} note 1, at 1688 (observing as well that “if such a ‘right not to stand trial’ existed, then by definition it would have to be redressed up front through interlocutory review or not at all”).

\textsuperscript{207.} \textit{Id.; see also} \textit{FTC v. Standard Oil Co. of Cal.}, 449 U.S. 232, 246 (1980) (“[T]he issuance of the complaint averring reason to believe is a step toward, and will merge in, the Commission’s decision on the merits.”).

\textsuperscript{208.} \textit{See In re Cuozzo}, 793 F.3d at 1273 (holding that “§ 314(d) prohibits review of the decision to institute IPR even after a final decision.”).

\textsuperscript{209.} Brief for Intervenor-Director of the United States Patent and Trademark Office at 30–
become what the Solicitor General later advanced as the government’s position before the Court.\textsuperscript{210} This view prevailed, overcoming even the strong presumption in administrative law favoring judicial review of agency action.\textsuperscript{211}

The importance of this holding to court-agency duplication lies in the incentive that it creates for the USPTO. The threshold screening function that the PTAB performs—on behalf of the Director—is immune from judicial scrutiny, but the merits adjudication that the PTAB performs later is not. Thus, for any action as to the PTAB where the USPTO expects it has greater expertise and would be ill-served by judicial interference, the agency has a rational motivation to characterize its action as pre-institution screening rather than post-institution adjudication.\textsuperscript{212}

In fact, reframing in this way is not especially difficult. In \textit{inter partes} review, instituting review requires considering the arguments in the petition, any preliminary response filed by the patent owner, and finding “a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.”\textsuperscript{213} A likelihood-of-success standard means that the unreviewable institution and the reviewable adjudication both focus on the merits of the validity challenge.\textsuperscript{214}

The question then becomes, what safeguards against duplicative litigation might the USPTO wish to administer without judicial interference? The past decade has revealed three significant examples, each of which has eventually reached the Supreme Court. The USPTO has pressed an expansive view of its own discretion each time, winning twice and losing once.

2. The Agency’s Need to Reconstrue Patents

The earliest test case of the USPTO’s power in a context that implicated court-agency duplication was the legal standard for construing patent claims in the PTAB. From the outset, the rules that the agency promulgated to govern AIA trial proceedings all adopted a practice of giving a patent claim its “broadest reasonable construction in light of the specification of the patent in which it appears.”\textsuperscript{215} This was in stark contrast to the standard of \textit{Phillips v. AWH Corp.} used in U.S. district courts, giving claims their “ordinary meaning . . . as understood by a person of skill in the art.”\textsuperscript{216} The divergence of interpretive methods made it necessary for the PTAB to interpret patent claims anew, even when courts or the ITC had already invested time and resources to interpret the same patent claims in prior cases, creating needless and wasteful duplication. Whatever the wisdom of that view, however, the USPTO argued that its own policy judgment on the issue was conclusive.\textsuperscript{217}

\begin{flushleft}
211. \textit{Cuozzo}, 136 S. Ct. at 2140–42.
212. Vishnubhakat, supra note 1, at 1755.
214. Vishnubhakat, supra note 1, at 1692.
215. 37 C.F.R. § 42.100(b) (2012).
216. 415 F.3d 1303, 1314 (Fed. Cir. 2005) (en banc).
\end{flushleft}
As a matter of doctrine and institutional power, the USPTO prevailed on this issue easily—indeed, unanimously, in the Supreme Court. The framing of the issue made the outcome straightforward. The issue was not the wisdom of the “broadest reasonable construction” approach versus the “ordinary meaning” approach but simply who was authorized to choose. It was not Congress who chose to depart from judicial practice in claim construction. By contrast, it was, indeed, Congress who chose in the AIA to depart from the judicial standard of clear and convincing evidence in favor of a mere preponderance standard. Instead, on claim construction and on a range of other matters, Congress gave the Director authority to “prescribe regulations . . . establishing and governing inter partes review under this chapter and the relationship of such review to other proceedings under this title.”

Upon recognizing this delegation of authority to the agency, all that was left to ask was whether the USPTO’s choice of a “broadest reasonable construction” standard was unreasonable. The court concluded, sensibly, that it was not but expressly declined to decide “whether there is a better alternative as a policy matter. That is a question that Congress left to the particular expertise of the Patent Office.” That question of optimal policy, however, does have a well-reasoned answer. It is that PTAB claim constructions should follow the same interpretive standard as the Article III courts, primarily to avoid duplication, but also for other benefits such as structural fairness to patent owners and the reduction of forum-shopping.

The duplication problem lies in how using the broadest reasonable construction differs in practice from using the ordinary meaning under Phillips. Giving a patent claim its broadest reasonable construction means interpreting the claim to be as broad as the patent specification, especially including the inventor’s disclosure, can support. This usually results in patent claim scope that is broader than even the inventor intends. Correspondingly, though, this also makes it more likely that the claim, as interpreted, will impermissibly tread upon existing patents, technical literature, and other prior art that cannot be made part of the inventor’s exclusive patent right.

By contrast, the ordinary meaning approach of Phillips gives claims their “ordinary and customary meaning,” i.e., what the claims would have meant “to a person of ordinary skill in the art in question at the time of the invention.” This

218. Id. at 2142, 2146. Justice Alito’s separate opinion, joined by Justice Sotomayor, dissented from the opinion of the Court only on the question of judicial reviewability under § 314(d)—but joined the Court fully as to the claim construction standard issue. See id. at 2148–49 (Alito, J., concurring in part).
219. Id. at 2142.
220. See supra Section II.A.1.
222. Cuozzo, 136 S. Ct. at 2146.
223. Id.
interpretation is to be done in light of the patent specification, which is “the single best guide to the meaning of a disputed term and . . . acts as a dictionary when it expressly defines terms used in the claims or when it defines terms by implication.” 226 The administrative history of patent prosecution is another valuable source of intrinsic evidence about patent claim meaning and should “inform the meaning of the claim language by demonstrating how the inventor understood the invention.” 227 Expert testimony, dictionaries, and other evidence that is extrinsic to the examination and issuance of the patent is legally regarded as the least reliable of all. 228

This divergence does not merely permit duplication but essentially requires it. For a given patent claim, the broadest reasonable construction is more likely to lead to invalidation than the ordinary meaning under Phillips, primarily by exposing claims to a broader set of potentially invalidating prior art. 229 As a result, claims that have been interpreted in prior federal court or ITC proceedings must nevertheless be reinterpreted under the PTAB standard. Whether the PTAB reaches the same result or disregards a coordinate tribunal to reach a contrary result, the problem of duplication manifests either way. In this regard, usage of the broadest reasonable construction standard in the PTAB presents problems quite similar to the application of divergent burdens of proof, which also are more lenient in the PTAB and are more skewed accordingly toward findings of patent invalidity. 230 The only saving grace of the claim construction problem is that it is not ossified by statute.

The related concern about structural fairness to patent owners goes to why the USPTO chose the broadest reasonable construction standard for PTAB trials at all. That interpretive method is certainly well-suited for patent examination ex ante because it pragmatically assumes that applicants will draft claims broadly and, accordingly, forces them to narrow those claims by adding specificity and detail until it is clear that they have sufficiently drafted around the prior art. 231 In this way, the broadest reasonable construction standard acts as an information-forcing penalty default rule. 232 In fact, the same interpretive method has also long been used in reexamination ex post for the same reason. 233 Assuming a broad remit for already-

226. Id. at 1321.
227. Id. at 1317.
228. Id. at 1318–19.
229. See, e.g., Calvin M. Brien, An Empirical Analysis of Patent Validity in Inter Parties Reviews Through the Lens of KSR, 46 AIPLA Q.J. 413, 426 n.82 (2018) (noting that “[t]he broadest reasonable interpretation standard arguably results in a broader claim construction than the so-called Phillips standard, which could explain the high likelihood of invalidation at the PTAB”); see also Clearlamp, LLC v. LKQ Corp., No. 12 C 2533, 2016 WL 4734389, at *12 (N.D. Ill. Mar. 18, 2016) (characterizing the court’s own claim construction standard of ordinary and customary meaning as “more stringent” than the broadest reasonable construction employed by the PTAB).
230. See supra Section II.A.1.
issued patent claims subjects patent owners to the same pressure, similarly forcing
them to narrow those claims as needed to avoid treading on prior art.

However, this pressure is unnecessary in AIA trial proceedings, where patent
owners cannot enlarge the scope of their already-issued claims, but can only define
or narrow the scope of their claims as originally issued.\textsuperscript{234} Moreover, the right to
amend is not nearly as robust in PTAB trials as it is in examination or reexamination,
where amendment is central to the entire process. The AIA grants only a right to
make a motion to amend,\textsuperscript{235} and the frequency with which the PTAB fully grants
such motions has been consistently low from the start, at rates at or lower than ten
percent.\textsuperscript{236} Until the en banc Federal Circuit decided \textit{Aqua Products, Inc. v. Matal} in
October 2017, the USPTO had improperly given patent owners the burden of proof
in motions to amend as well, making amendments still harder to come by.\textsuperscript{237}

Recognizing these problems and responding in particular to concerns about
duplication and the waste of PTAB resources, the USPTO recently reversed its policy
on claim construction.\textsuperscript{238} Following notice-and-comment rulemaking, the USPTO
amended its regulations to provide that patent claims in the PTAB “shall be construed
using the same claim construction standard that would be used to construe the claim
in a civil action under 35 U.S.C. 282(b).”\textsuperscript{239} The rule specifically addressed \textit{Phillips}
by adopting “the ordinary and customary meaning of such claim as understood by
one of ordinary skill in the art and the prosecution history pertaining to the patent.”\textsuperscript{240}
The rule also provided that the PTAB would consider “[a]ny prior claim construction
determination concerning a term of the claim in a civil action, or a proceeding before
the [ITC], that is timely made of record.”\textsuperscript{241}

The USPTO’s concern with avoiding duplication was central to this regulatory
reform, and the final rule articulated it. The high level of empirically observed
overlap between the PTAB and other tribunals, explained the agency, “favors using

\begin{itemize}
\item \textsuperscript{234} 35 U.S.C. § 316(d)(3); \textit{see also} Vishnubhakat et al., \textit{supra} note 224, at 3.
\item \textsuperscript{235} Even this is limited by statute to one motion as a matter of right, with additional
motions allowed only by mutual agreement or in the regulatory discretion of the USPTO. 35
\item \textsuperscript{236} Vishnubhakat, \textit{supra} note 59, at 519, 528 (citing a series of the PTAB’s own internal
studies confirming these findings).
\item \textsuperscript{237} 872 F.3d 1290 (Fed. Cir. 2017) (en banc). Following \textit{Aqua Products}, the chief judge
of the PTAB in November 2017 issued a guidance memorandum on implementing the new
Following a 2019 notice of proposed rulemaking, the USPTO formalized its post-\textit{Aqua
Products} guidance in December 2020. \textit{See} Rules of Practice to Allocate the Burden of
Persuasion on Motions to Amend in Trial Proceedings Before the Patent Trial and Appeal
\item \textsuperscript{238} Changes to the Claim Construction Standard for Interpreting Claims in Trial
Proceedings Before the Patent Trial and Appeal Board, 83 Fed. Reg. 21,221 (May 9, 2018) (to
\item \textsuperscript{239} Changes to the Claim Construction Standard for Interpreting Claims in Trial
Proceedings Before the Patent Trial and Appeal Board, 83 Fed. Reg. 51,340, 51,345 (Oct. 11,
\end{itemize}
a claim construction standard in AIA proceedings that is the same as the standard used by federal courts and the ITC.” The agency also agreed that “[u]sing the same claim construction standard across the fora would increase efficiency, as well [as] reduce cost and burden because parties would only need to focus their resources to develop a single set of claim[] construction arguments.”

3. The Agency’s Practice of Partial Adjudication

Despite this initial victory, the next major test of the USPTO’s power proved to be a retrenchment, and it implicated the problem of PTAB duplication even more than divergent claim construction had. This was the agency’s practice of only partly adjudicating AIA trial petitions. The practice was rooted in one view of agency efficiency, though it was an unduly narrow view. Its principal impacts upon court-agency duplication were twofold—the dilution of judicial stay practice as briefly discussed already and, more directly, the weakened application of the AIA’s estoppel provisions. Partial adjudication also gave rise for a time to the controversy over redundant grounds by which the PTAB itself tried understandably to manage duplicative arguments within petitions but did so in a way that created larger concerns about unchecked agency discretion and the constitutional separation of powers.

Partial adjudication began with partial institution. Under original AIA trial practice rules, a PTAB panel reviewing a petition could “authorize the review to proceed on all or some of the challenged claims and on all or some of the grounds of unpatentability asserted for each claim.” Where only certain portions of a petition were instituted, only those portions would be litigated on the merits and eventually adjudicated in the PTAB’s final written decision.

Under this arrangement, estoppel effects are an immediately salient question. All estoppel provisions across all AIA trial proceedings first require a final written decision. This already creates a possibility that the scope of estoppel might be limited in accordance with the scope of the adjudication. Even the broad language that estoppel applies to “any ground that the petitioner raised or reasonably could have raised” is cabined to arguments that were available to the petitioner “during...
that *inter partes* review” (or post-grant review or covered business method review).\(^\text{250}\)

If the time *during* a proceeding was deemed to begin with the filing of the petition, then estoppel could foreclose even arguments that the petitioner raised in its petition but that the PTAB never instituted and thus never gave a chance to litigate. This would create obvious due process concerns by depriving petitioners of both an initial opportunity for a PTAB review and any later opportunity due to estoppel.\(^\text{251}\) Those due process concerns would carry far less force, of course, where the petitioner failed to include the relevant arguments in the petition at all.\(^\text{252}\)

But if the time *during* a proceeding was deemed to begin only if and when the PTAB granted institution, then the problem of duplication would replace that of due process. Only arguments that were raised or that reasonably could have been raised after institution would be subject to estoppel, but all of the petitioner’s other arguments would remain available for relitigation even after a final written decision by the PTAB. Indeed, this was the very position that the USPTO took as to the estoppel effects of its partial institution practice.\(^\text{253}\)

The USPTO probably chose correctly between its available options: undermine estoppel and tolerate duplication, or else maintain estoppel at the risk of offending due process. Yet it was the agency’s own flawed practice of partial adjudication that forced the choice. Similarly, the negative effect of partial institution on judicial stays and, by extension, on the problem of court-agency duplication as already discussed\(^\text{254}\) was also a result of a policy tradeoff by the USPTO. In that regard, accepting some arguments for review and rejecting others reduced the clarity and quality of information available to judges who were considering stays. Meanwhile, the PTAB was understandably looking to its own institutional benefit by reducing its workload and streamlining the issues before it for adjudication.\(^\text{255}\)

In this conflict between the supervening legislative policy interest in promoting PTAB review as a substitute rather than a wasteful duplicate for other forums of litigation and the USPTO’s own past policy choices in favor of its own insular interests, the resolution eventually came from the AIA’s text. When the Supreme Court in 2018 rejected the validity both of partial institution and of partial adjudication,\(^\text{256}\) its grounds were chiefly textual. The error in the Federal Circuit’s contrary holding below was also textual in notable respects.

---

\(^{250}\) *Id.* (emphasis added).


\(^{253}\) Shaw Indus. Grp., Inc. v. Automated Creel Sys., Inc., 817 F.3d 1293, 1300 (Fed. Cir. 2016) (noting the USPTO argument that “the denied ground never became part of the IPR”).

\(^{254}\) See supra Section II.A.3.

\(^{255}\) Vishnubhakat, supra note 173, at 2645 (discussing partial institution as “both a policy lever and a case management lever” in the PTAB’s control over its own docket).

In creating partial institution, the USPTO’s rationale that the rule comported with the AIA was little more than a recitation of its power to “prescribe regulations . . . establishing and governing inter partes review.”257 To this might have been added the logical principle, common in the law, that the greater includes the lesser.258 Thus, the grant of full institution power might also grant a lesser included power of partial institution; so, too, for denial.259

During litigation, the USPTO turned more specifically textual. The relation between the institution decision and the final written decision was of key importance. The statute that governs institution, § 314(a), refers specifically to “claims challenged in the petition.”260 By contrast, the statute that governs final written decisions, § 318(a), refers instead to “any patent claim challenged by the petitioner.”261 The Federal Circuit agreed that “the differing language implies a distinction between the two subsections such that § 318(a) does not foreclose the claim-by-claim approach” of the PTAB in adjudicating only the claims on which it had previously granted partial institution.262

A majority of the Supreme Court squarely rejected that reading. The grant in § 314 of authority to determine “whether to institute” was a binary grant.263 The “party-directed, adversarial process” that Congress created in AIA trial proceedings reflected “structural choices” that were no less “deserving of judicial respect” than the text itself.264 The claim-by-claim approach failed as well. The USPTO had argued that the requirement in § 314(a) of the petitioner’s likelihood of success as to “at least 1 of the claims challenged in the petition” commanded a claim-by-claim approach that carried through to the § 318(a) requirements for a final written decision.265 The Court noted instead that § 314(a) specifically does not require evaluating every claim but only the likelihood of success on “at least 1” claim:

Once that single claim threshold is satisfied, it doesn't matter whether the petitioner is likely to prevail on any additional claims; the Director need not even consider any other claim before instituting review. Rather than contemplate claim-by-claim institution, then, the language anticipates a

258. See, e.g., Sheldon v. Sill, 49 U.S. 441 (1850) (holding that the power of Congress under Article I to create the lower federal courts includes the lesser power to define their jurisdiction); United States v. Wilson, 26 F.3d 142, 150 (D.C. Cir. 1994) (finding that the Ethics Act’s grant of plenary power to the Office of Independent Counsel also grants the lesser included power to “participat[e] in court proceedings and engaging in any litigation . . . that such independent counsel considers necessary”).
259. Vishnubhatk, supra note 173, at 2652; see also Vishnubhatk, supra note 245, at 793–94.
260. 35 U.S.C. § 314(a) (requiring “a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition”).
261. 35 U.S.C. § 318(a) (requiring the final written decision to address “the patentability of any patent claim challenged by the petitioner”).
264. Id. at 1355.
265. Id. at 1356.
regime where a reasonable prospect of success on a single claim justifies review of all.\textsuperscript{266}

Though the Court’s opinion did not address it, the Federal Circuit’s reasoning on differing language was questionable as well. It is true that § 314(a) refers to “claims challenged in the petition” while § 318(a) refers to “claim[s] challenged by the petitioner.” However, making this comparison in isolation ignores the remaining text of § 318(a), that the final written decision must address “the patentability of any patent claim challenged by the petitioner and any new claim added under section 316(d).”\textsuperscript{267} When it comes to the petitioner and their petition, the requirements of the final written decision statute do not diverge at all from the requirements of the institution statute. There is merely the further requirement that the final written decision must also address new claims introduced by the patent owner.

Moreover, though the opinion of the Court declined to engage with the policy wisdom of partial institution and partial adjudication, the respective dissents of Justices Ginsburg and Breyer diverged from the majority’s statutory reading and its impacts on PTAB case management.\textsuperscript{268} Justice Ginsburg’s dissent was especially trenchant, noting that even under the binary choice now remaining, the PTAB could simply issue full denials with explanations of those arguments that merited review and those that did not.\textsuperscript{269} By these decisional roadmaps, petitioners could simply refile “new or amended petitions shorn of challenges the Board finds unworthy of inter partes review”—leading back to the same streamlining that partial institution had more promptly created.\textsuperscript{270} On this view, the whole exercise was a waste of the PTAB’s time and was, if anything, a recipe of duplicative filing in the PTAB itself.\textsuperscript{271}

But this was not so. Justice Ginsburg’s mock proposal was, in fact, a concise statement of precisely why the Court’s rejection of partial adjudication actually fosters a policy of internal efficiency\textsuperscript{272}—and, as an added benefit, does away with the problems of estoppel and judicial stays that past USPTO practice had fostered. In the post-\textit{SAS Institute} world of binary institution, petitioners must now “pare down their challenges in order to balance the hope of full institution against the risk of full denial.”\textsuperscript{273} The PTAB’s decisions will, indeed, be roadmaps as Justice Ginsburg noted, but not only for unsuccessful petitioners to refile without undue guesswork.\textsuperscript{274} These more stringent decisions on binary institution will also channel future petitioners to “file challenges that are more likely in the first place to be fully granted than fully denied.”\textsuperscript{275}

Another new study, whose findings are reported as follows, seems to bear out this expectation of greater dynamic efficiency. Briefly put, filing a PTAB trial petition

\begin{itemize}
  \item \textsuperscript{266} \textit{Id.} (emphasis in original).
  \item \textsuperscript{267} 35 U.S.C. § 318(a) (emphasis added).
  \item \textsuperscript{268} \textit{SAS Inst.}, 138 S. Ct. at 1360 (Ginsburg, J., dissenting); \textit{id.} (Breyer, J., dissenting).
  \item \textsuperscript{269} \textit{Id.} (Ginsburg, J., dissenting).
  \item \textsuperscript{270} \textit{Id.} (emphasis added).
  \item \textsuperscript{271} \textit{Id.}
  \item \textsuperscript{272} See Vishnubhakat, supra note 173, at 2666.
  \item \textsuperscript{273} \textit{Id.} at 2667.
  \item \textsuperscript{274} \textit{Id.}
  \item \textsuperscript{275} \textit{Id.}
\end{itemize}
now carries a new downside risk. Sufficiently meritorious petitions will continue to be instituted fully, and sufficiently defective ones will continue to be denied fully. Yet a petition with a more even mix of winning and losing arguments might be fully rejected based on the PTAB’s discretion. As a result, the overall risk of filing a PTAB petition has grown after *SAS Institute*. Accordingly, the expected result should be a decline in the number of litigants who seek PTAB review. Conditional on seeking PTAB review, though, there should be a rise in the number of petitions relative to the number of disputes, as petitioners distribute their challenges across multiple petitions as a hedge against full denial.

Indeed, this is what came to pass based on PTAB trial filings from the months before and after the *SAS Institute* decision. The number of disputes (where a dispute is defined as a set of petitions between the same petitioner and patent owner) declined by 6.3% from the four-month period just before *SAS Institute* to the four-month period just after. The number of individual petitions, meanwhile, rose by 8.8% across the same period. Looking only at *inter partes* review, the number of disputes fell by 2.2% while the number of petitions rose by 14.9%. Figure 9 summarizes these findings.276

Thus, though the USPTO may not have recognized it and though the Supreme Court majority did not even address the issue in *SAS Institute*, the end of partial institution and partial adjudication has been a net gain in disciplining extravagant petitioning and in eliminating the ill effects on PTAB estoppel and judicial stays, resulting in a measure of progress toward reducing duplication.

4. The Agency’s Self-Policed Statutory Deadline

The most recent test of the USPTO’s power has been oversight of the statutory deadline for seeking *inter partes* review. This enforcement issue has also been the duplication-related doctrine that has seen the most judicial reversal and uneven enforcement.277 Unlike claim construction and partial adjudication, it is currently an open problem. Compared with AIA trial proceedings, the timing for filing an *inter partes* review is more flexible.278 If the petitioner has not yet been sued for infringing the relevant patent, the *inter partes* review petition can be filed at any time. If the petitioner has been sued, then the petition must come within one year from the date when the petitioner was served (or when its privy or real party in interest was served) with the court complaint.279

The bulwark of this one-year deadline against court-agency duplication lies in the choice it forces upon district court defendants. Before substantial time and resources have been invested in the Article III proceeding, they must either commit to seeking administrative review in the PTAB or else forgo it altogether.280

276. See infra Tables and Figures at Figure 9.
277. See supra notes 104–107 and accompanying text.
279. 35 U.S.C. § 315(b).
280. Vishnubhakat, supra note 143, at 1088.
Prior empirical results also confirm that the one-year deadline acts as a strong and meaningful constraint on standard petitioners (those who have previously been sued on the patents they now challenge), and that the defensive posture of standard petitioning represents the primary usage of *inter partes* review. As discussed already in the context of PTAB joinder and the delicate balance it reflects between beneficial collective action and harassment, nonstandard petitioners in *inter partes* review frequently join the challenges brought by standard petitioners.

Closely related to this petitioner strategy is the time lag between when patents are asserted in the courts and when they are first challenged in the PTAB. There are two measures of this lag. One is what may be termed the “first court/first PTAB” lag, measured from the first *inter partes* reviews of patents back to their earliest respective court litigations. The other is the “last court/first PTAB” lag, measured from the first *inter partes* review back to their most recent pre-PTAB court litigations. The distributions of these lags offer clues about how frequently *inter partes* review is being used defensively—and, thus, how much meaningful work the one-year deadline is doing to guard against court-agency duplication.

For the first court/first PTAB lag, nearly 23.4% of petitions exceed the one-year deadline, reaching upwards of three years in some cases. This is to be expected, especially for repeatedly asserted patents, because any of the various defendants may be the one to bring the first PTAB challenge. Yet most petitions come within the one-year deadline, distributed symmetrically at a median of six months and with a modal spike at the one-year mark.

The last court/first PTAB lag is more tightly distributed with only 11.4% of petitions exceeding the one-year deadline and with a sharper modal spike at the one-year mark. This is also to be expected, as earlier lawsuits against others are likely to have revealed useful information about the patent owner’s enforcement strategy, requiring less time to decide whether and how to mount a PTAB challenge. Meanwhile, the reduced share of 11.4% of petitions that comes after one year is virtually all nonstandard. Figures 10–11 show these distributions.

The upshot is that *inter partes* review is, indeed, chiefly a defensive response to existing litigation and does substantially implicate the potential for court-agency duplication. The one-year time bar for bringing *inter partes* review is a correspondingly important structural pillar in the substitutionary boundary between the PTAB and other tribunals.

281. See *supra* notes 62–63 and accompanying text.
282. See *supra* notes 96–102 and accompanying text.
283. *Strategic Decision Making, supra* note 2, at 75.
284. *Id.*
285. *Id.*
286. *Id.* at 76.
287. *Id.* at 75.
288. *Id.* at 76.
289. *Id.*
290. *Id.*
291. *Id.* Indeed, the only way this conclusion would not hold true is if there were systematic nonenforcement of the one-year deadline.
distributions indicate especially that would-be petitioners work to the statutory deadline and will take as much time as agency enforcement actually tolerates in practice.\textsuperscript{293}

Yet for all its importance, concerns about lax enforcement of the one-year bar were long underappreciated. The USPTO argued that PTAB determinations about the one-year deadline were judicially unreviewable under § 314(d), just like its decisions whether to institute and its practice of partial institution.\textsuperscript{294} The Federal Circuit initially accepted this argument in its 2015 panel opinion in \textit{Achates Reference Publishing, Inc. v. Apple Inc.}\textsuperscript{295} However, pressure for judicial intervention continued to mount, and in 2018 the Federal Circuit reversed itself en banc in \textit{Wi-Fi One, LLC v. Broadcom Corp.}\textsuperscript{296} The court in \textit{Wi-Fi One} expressly recognized the boundary-enforcing function of the one-year deadline and agreed it was “fundamentally different” from the routine institution decisions that the Supreme Court had immunized from judicial scrutiny in \textit{Cuozzo}.\textsuperscript{297}

Given the persistent controversy over this highly important safeguard of court-agency substitution and the high-profile turnabout of the Federal Circuit, it is helpful to clarify the stakes of the USPTO’s position. The reason for skepticism at USPTO self-enforcement of the one-year deadline is not that the PTAB is disproportionately error prone or that internal oversight by the agency’s political leadership is systematically inferior, let alone that the PTAB would engage in willful disregard of statutory constraints.\textsuperscript{298} The problem of judicial non-supervision is more subtle.

The one-year deadline itself is triggered not only by serving a complaint upon a defendant who goes on to seek inter partes review. It also constrains court defendants that may turn out to be privies or real parties in interest as to other entities that go on to seek inter partes review. Deciding what circumstances satisfy these relational criteria is a highly case-specific enterprise. Thus, the real value of judicial oversight, especially in the early years of AIA trials, would have been prompt appellate-course correction and the development of uniform guidance for the PTAB. \textit{Wi-Fi One} did bring en banc resolution, but the early error of \textit{Achates} still squandered a valuable opportunity.

Unfortunately, even that en banc resolution did not hold. In 2020, the Supreme Court held in \textit{Thryv, Inc. v. Click-to-Call Technologies, LP} that the presumption favoring judicial review of agency action—which carried the day in \textit{Wi-Fi One}—was sufficiently rebutted as it was in \textit{Cuozzo}.\textsuperscript{299} By its holding in \textit{Thryv}, the Court has once again restored the pre-2018 status quo of \textit{Achates}, leaving PTAB decisions as to the one-year time bar judicially unreviewable and entirely self-policed by the USPTO.

\textsuperscript{293} See \textit{Strategic Decision Making}, supra note 2, at 76.

\textsuperscript{294} Vishnubhakat, supra note 143, at 1084–86; Vishnubhakat, supra note 173, at 2651–52; Vishnubhakat, supra note 245, at 791.

\textsuperscript{295} 803 F.3d 652 (Fed. Cir. 2015).

\textsuperscript{296} 878 F.3d 1364 (Fed. Cir. 2018) (en banc).

\textsuperscript{297} \textit{Id.} at 1374; see also Vishnubhakat, supra note 143, at 1087–88.

\textsuperscript{298} Vishnubhakat, supra note 143, at 1082 (noting that, “[t]o the contrary, the Patent Office can be quite politically responsive, and stakeholder complaints about disfavored agency policies can lead to dramatic reversals”).

\textsuperscript{299} \textit{Thryv, Inc. v. Click-to-Call Techs., LP}, 140 S. Ct. 1367, 1373–74 (2020).
As a result, Thryv also represents the current state of the USPTO’s ascendance as an institutional power center in the patent system. The Agency’s discretion over institution decisions is broad, and its attempts to characterize many of the PTAB’s structural features as institution-related have been successful more often than not. The result is that the duplication observed in PTAB trials with respect to the work of other tribunals, is not solely a result of the AIA’s statutory design. In significant part, that duplication is also a result of the USPTO’s expansive views about its own power and primacy.

C. Recent USPTO Attention to Duplication

It is essential, therefore, that any serious assessment of reform must begin with the USPTO itself. Indeed, in the PTAB’s most recent years, as problems of duplication have grown to the scale documented here, the Agency itself has paid correspondingly greater attention to calls for reconsideration and reform. In this effort, the leadership of the PTAB and of the USPTO as a whole have made moves to address internal as well as external duplication.

1. Agency Reaction to Internal Duplication

A discussion of internal duplication within the PTAB may seem odd. The duplication problem, as defined thus far, has focused on the interaction of the PTAB with other tribunals: Article III courts and the ITC. Likewise, the foregoing discussions of shortcomings in AIA trials has situated the PTAB as a substitute for other tribunals, both in statutory design and institutional design. Internal duplication (i.e., relitigating in the PTAB those cases which the PTAB itself has already decided) has simply not been part of the discussion.

There are good reasons for this. Confined by its nature to the USPTO, internal duplication is unlikely to frustrate the particular efficiency aims of the AIA. More basically, it is also less likely to take place. As a matter of operational control, the USPTO’s workload management incentives make it quite unlikely that the Agency will tolerate the same cases being litigated repeatedly in the PTAB.

As to the functional mechanisms for avoiding duplication, divergences between the PTAB and other tribunals do not—indeed, cannot—exist internally. Unlike court-agency estoppel, agency-agency estoppel is bilateral. Standards for claim construction and the burdens of proof have always been consistent from one PTAB case to the next. Finally, as a systemic matter, any duplication that does arise internally does not present concerns about inter-branch comity and the separation of powers.

Nevertheless, the reason to consider internal duplication now is that the mechanism that the USPTO has used for managing it is the same mechanism that the USPTO has quite recently begun to use for grappling with the problem of external duplication. That mechanism is the Agency’s broad discretion at the pre-institution

300. See supra Section I.B.
301. See supra Section II.A.
302. See supra Section II.B.
screening phase of AIA trials. The substantial rise of the USPTO’s administrative and adjudicative power in the past decade is connected closely with its unreviewable discretion to select its cases. By now, the courts have repeatedly reaffirmed the breadth of this discretion, making it a reliably powerful policy lever for the Agency to guard against duplicative litigant behavior.

The PTAB’s framework for applying its discretion in this way comes from its decision in General Plastic Industrial Co. v. Canon Kabushiki Kaisha. Decided in September 2017, the case was designated precedential the following month. General Plastic offers a multifactor analysis for determining whether a petition for review that might otherwise merit institution should nevertheless be rejected in order to preserve “the efficiency of the inter partes review process and the fundamental fairness of the process for all parties.” The seven factors that the opinion specifies reflect all the same concerns as to duplication rather than substitution and as to harassment rather than beneficial collective action. Those factors are as follow:

1. whether the same petitioner previously filed a petition directed to the same claims of the same patent;
2. whether at the time of filing of the first petition the petitioner knew of the prior art asserted in the second petition or should have known of it;
3. whether at the time of filing of the second petition the petitioner already received the patent owner’s preliminary response to the first petition or received the Board’s decision on whether to institute review in the first petition;
4. the length of time that elapsed between the time the petitioner learned of the prior art asserted in the second petition and the filing of the second petition;
5. whether the petitioner provides adequate explanation for the time elapsed between the filings of multiple petitions directed to the same claims of the same patent;
6. the finite resources of the Board; and
7. the requirement under 35 U.S.C. § 316(a)(11) to issue a final determination not later than 1 year after the date on which the Director notices institution of review.

303. See supra Section II.B; Vishnubhat, supra note 1, at 1676–1704.
305. Opinions designated “precedential” bind all future PTAB cases unless and until later binding authority supersedes them. Lesser designations such as “informative” or “representative” leave an opinion nonbinding but persuasive in various respects. See generally Saurabh Vishnubhat, Precedent and Process in the Patent Trial and Appeal Board, Patently-O (May 10, 2016), https://patentlyo.com/patent/2016/05/precedent-process-patent.html [https://perma.cc/D2EB-H77T].
307. Id. at 16.
In making this framework binding upon future PTAB decisions, the USPTO specifically adopted a policy view that “[a]llowing similar, serial challenges to the same patent, by the same petitioner, risks harassment of patent owners and frustration of Congress’s intent” in the AIA.\footnote{Id. at 17 n.14.} Notably, factor one takes aim at multiple petitions by the same petitioner, and factors two through five seem to reflect the same assumption consistent with the facts of General Plastic.

Yet over time, the PTAB has also applied the framework to multiple petitions as to the same patent but filed by different petitioners. The April 2019 decision in Valve Corp. v. Electronic Scripting, Products Inc. squarely extended General Plastic in this way, and Valve Corp. itself was designated precedential a month later.\footnote{Nos. IPR2019-00062, IPR2019-00063, IPR2019-00084 (P.T.A.B. Apr. 2, 2019) (designated precedential May 7, 2019).} Though controversial in some circles, it seems clear that the use of the General Plastic framework against nonmutual petitioners is consistent with reducing internally duplicative litigation\footnote{See, e.g., Steven C. Carlson, Weaponizing IPRs, LANDSLIDE, Sept.–Oct. 2019, at 36, 39–40 (framing Valve Corp. as a “tightening” of the General Plastic standard and situating the use of the PTAB discretion through these cases as a response to the “harm posed by duplicative IPR petitions”).} and is within the unreviewable, institution-related discretion of the USPTO.\footnote{See Koninklijke Philips N.V. v. Iancu, 829 F. App’x 967, 967–68 (Fed. Cir. 2020) (per curiam).}

2. Agency Reaction to Inter-Branch Duplication

What has been more controversial and, indeed, remains a live issue even now, is the USPTO’s still further expansion of this framework to manage external duplication.\footnote{Much of the discussion that follows was included in a public commentary on an APA lawsuit against the USPTO. See Saurabh Vishnubhakat, Silicon Valley’s APA Challenge to PTAB Discretion, PATENTLY-O (Sept. 2, 2020), https://patentlyo.com/patent/2020/09/silicon-challenge-discretion.html [https://perma.cc/TC5S-J7U5].} The basis for that expansion first came in NHK Spring Co. v. Intri-Plex Technologies, Inc.\footnote{No. IPR2018-00752 (P.T.A.B. Sept. 12, 2018) (designated precedential May 7, 2019).} and was subsequently elaborated in Apple, Inc. v. Fintiv, Inc.\footnote{No. IPR2020-00019 (P.T.A.B. Mar. 20, 2020) (designated precedential May 5, 2020).} These decisions together form the NHK-Fintiv doctrine, a policy of discretionarily denying institution where a parallel district court proceeding is far enough along and similar enough in evidence and argument that undertaking review would be largely or entirely duplicative.

The decision in NHK was a discretionary denial of NHK Spring’s petition against an Intri-Plex patent.\footnote{NHK, No. IPR2018-00752, at 2.} The reason for the denial was not duplicative litigation in the PTAB itself, either by Intri-Plex or by other petitioners, but instead was an infringement lawsuit involving the same patents and parties that was co-pending in the Northern District of California.\footnote{See id. (identifying the co-pending case as Intri-Plex Techs., Inc. v. NHK Int’l Corp., No. 3:17-cv-01097-EMC (N.D. Cal.).}

308. Id. at 17 n.14.
310. See, e.g., Steven C. Carlson, Weaponizing IPRs, LANDSLIDE, Sept.–Oct. 2019, at 36, 39–40 (framing Valve Corp. as a “tightening” of the General Plastic standard and situating the use of the PTAB discretion through these cases as a response to the “harm posed by duplicative IPR petitions”).
316. See id. (identifying the co-pending case as Intri-Plex Techs., Inc. v. NHK Int’l Corp., No. 3:17-cv-01097-EMC (N.D. Cal.).)
the facts at hand did not satisfy the traditional *General Plastic* framework for managing internal duplication. However, the PTAB did accept the argument “that the advanced state of the district court proceeding is an additional factor that weighs in favor of denying the Petition under § 314(a).” The parallel litigation was nearing conclusion, and the PTAB explained that instituting review would not be consistent with the AIA’s substitutionary purpose. In short, *NHK* did not apply *General Plastic* as such but did add an eighth factor, expressly looking outward to parallel proceedings, to inform PTAB discretion.

To this, *Fintiv* provided a number of factors for inquiring more specifically into the state of a parallel proceeding for purposes of a denial under *NHK*. As with the internally focused questions of *General Plastic*, these additional *Fintiv* factors also reflect concerns of duplication rather than substitution, especially with an eye toward sunk costs of adjudication:

1. whether the court granted a stay or evidence exists that one may be granted if a proceeding is instituted;
2. proximity of the court’s trial date to the Board’s projected statutory deadline for a final written decision;
3. investment in the parallel proceeding by the court and the parties;
4. overlap between issues raised in the petition and in the parallel proceeding;
5. whether the petitioner and the defendant in the parallel proceeding are the same party; and
6. other circumstances that impact the Board’s exercise of discretion, including the merit.

The USPTO designated *NHK* as precedential on May 7, 2019 (the same as it did *Valve Corp.*) and *Fintiv* a year later on May 5, 2020.

Even more recently, the USPTO designated two additional decisions in the *NHK-Fintiv* framework as precedential in December 2020. These were *Sotera Wireless, Inc. v. Masimo Corp.* and *Snap, Inc. v. SRK Technology, LLC*. Both were decisions to grant institution. In *Sotera*, it was because the petitioner stipulated broadly that if the *inter partes* review is instituted, the petitioner “will not pursue in [the District Court] the specific grounds [asserted in the *inter partes* review], or on any other ground . . . that was raised or could have been reasonably raised in an IPR
In Snap, it was because the district court proceeding was stayed pursuant to a jointly stipulated motion by the parties. These further precedential and binding refinements to the PTAB’s NHK-Fintiv policy suggest that the Agency is committed to using its discretion broadly, as it is continuing to invest in decisions that clarify how that discretion will be used in the future. The NHK and Fintiv cases themselves offer clear doctrinal signposts for discretionary denials, and now the Sotera and Snap cases offer comparable signposts in the other direction for discretionary grants.

Nevertheless, while the influence and importance of PTAB discretion requires an assessment of reform to begin in the USPTO, the discussion must not end there. Even with the General Plastic and NHK-Fintiv frameworks in place, much of what remains to be clarified or corrected with duplicative litigation in the PTAB is embedded in politically inflected decision-making that can only be restored by statute to a regime of adequate judicial supervision. And of course, much that remains to be corrected is simply codified into statute in the first place and so can be reformed only by legislation. A plan for reform that adequately addresses all of the foregoing problems and assessments of their causes will require a coherent theoretical foundation.

III. Toward More Efficient Substitution

This Section offers that theoretical foundation and from it derives a series of proposed legislative corrections. The discussion to this point has reflected a preference for maximizing the use of PTAB review in its intended manner as a substitute, for minimizing the duplication of effort over the same patents as between the PTAB and other tribunals, and for disfavoring the strategic use of divergent legal standards to reach certain outcomes. One may reasonably ask, however, what makes these goals particularly desirable. To make that assessment, it is helpful first to consider the roles that timing and institutional competence play in resolving legal questions.

A. Timing and Institutional Competence

As to timing, while there are reasonable arguments for giving priority to a later adjudicator, those favoring the first adjudicator are ultimately more persuasive. Meanwhile, there are sound arguments for favoring either the Article III courts or other institutions such as Article I or Article II tribunals. However, though an expert institution like the PTAB may be well suited to carry out post-issuance patent validity review, some current features of PTAB review lack support in the justification from expertise and merit reform.

325. Snap, No. IPR2020-00820, at 10.
1. Priority to the First Adjudicator

In general, the more familiar stance in U.S. law is that priority should go to the first institution that decides a case. In its stronger form, this preference is embodied in res judicata doctrines like claim preclusion[^326] and issue preclusion[^327], which if they do apply to a given matter, conclusively foreclose the matter from further dispute. In a somewhat weaker form, this preference can be reflected in the jurisprudential norm of stare decisis[^328], which fosters stability and continuity[^329], just as res judicata does but which also yields adaptively to fairness concerns that precedent, unlike preclusion, can affect a great many people who were not before the court that issued it[^329]. As applied to the patent system in particular, the virtue of finality that underlies these preferences for first deciders implicates four major values.

The first is repose, widely understood not only as a basic value of adversarial civil dispute resolution but also among the most important[^330]. The pathologies of cost[^331], delay, inaccessibility, and strategic behavior in patent litigation—all of which were important spurs to the AIA itself[^331]—map quite well onto that “expensive and harrowing ordeal of litigation” from which litigants should be entitled to “a final result that not only precludes a second ordeal but also achieves the independent values of repose.”[^332]

In fact, the salience of repose to patent litigation is so peculiarly strong that, since well before the U.S. economy reached the pre-AIA tipping point during the mid to late 2000s, courts have found ways to take specific account of prior decisions regarding patent validity[^333]. For example, in 1975, just two years after Blonder-Tongue gave nonmutual defensive collateral estoppel effect to findings of patent invalidity[^334], the Northern District of Illinois held that a prior decision upholding a patent’s validity—though not entitled to collateral estoppel[^335]—could be admitted as evidence against a nonparty without violating due process[^334]. The opinion in Columbia Broadcasting v. Zenith Radio[^336] explained that admitting the prior judgment into evidence afforded the opposing party enough opportunity to rebut the conclusion of the prior judgment (e.g., by impeaching or by offering its own contrary evidence[^335]). Other cases from the same period have drawn the connection in terms...

[^326]: See generally Restatement (Second) of Judgments §§ 18–20 (Am. L. Inst. 1982).
[^327]: See generally id. §§ 27–29.
[^331]: See supra Section 1.A.
[^332]: 18 Wright, Miller & Cooper, supra note 330, at 22.
[^335]: Id. at 786.
that are even more starkly grounded in promoting repose for patent litigants. In all, these cases reflect a strong preference that the decision of the first adjudicator should receive priority.

The second major value at stake is reliance. Like repose, reliance is frequently at the core of arguments in favor of strong principles of finality in adjudication and adherence to stare decisis in jurisprudence. The parties to a litigation should, the argument goes, be able to rely conclusively on a final judgment, and a judgment that appropriately binds them is, by definition, preclusive. Meanwhile, if third parties can also rely on the judgment because it will also bind their conduct in some meaningful way and will oblige them to order their affairs accordingly, then such a judgment can “earn the label of precedent” and receive stare decisis effect. In both contexts, what drives the ability of people—parties and nonparties alike—to rely on the judgment is the expectation that it will be final in some conclusive sense.

The particular role of reliance interests in the patent system is considerable, perhaps even more in some ways than that of repose. As a system made primarily of incentives to innovate and to commercialize innovation, patent rights are foremost a means for appropriating value from investments in those activities. Such investments are often made on long time horizons that reflect greater uncertainty and thus respond more favorably to reliable appropriability mechanisms such as patents. The relatively fixed twenty-year term of U.S. patents, measured from the date when the application was filed, offers some hedge against that uncertainty. Still, the imperfect evaluation that goes into patent examination means that even an issued patent will have some uncertainty of being found invalid or noninfringed later. The legal resolution of that uncertainty is valuable, sometimes with as much

336. *E.g.*, Pachmayr Gun Works, Inc. v. Olin Mathieson Chem. Corp., 502 F.2d 802 (9th Cir. 1974). Notably, the court in *Pachmayr* found the relevant rationales of repose equally applicable between patent and trademark law. *Id.* at 805–06, 806 n.1.


340. *Id.*


344. However, the high variation among product life cycles and across industries also means that the twenty-year patent may overprotect in fast-moving contexts like high technology, and underprotect in slower-moving and more highly regulated contexts like pharmaceutical drugs. *See Fed. Trade Comm’n, The Evolving IP Marketplace: Aligning Patent Notice and Remedies with Competition* 92, 109 (2011).

value in publicly traded markets as the initial patent grant itself. However, this value is derived entirely from—and only to the extent that—the legal decision can be relied upon and not undone at another date even further in the future. In this way, the promotion of reliance interests also reflects a strong preference for the decisions of the first adjudicator.

The third major value is the mitigation of intervening changes in the law, and this may be regarded as a more generalized variation of reliance. The individual legal decisions on which litigants and third parties rely are, of course, rarely mechanical applications of determinate law to clear facts. Instead, they often reflect revisions and updates to the law. But any drift in the content of law that may occur through this incremental and interstitial process is different from the structural shock that comes from major legislative reforms like the AIA. In the context of finality, it is the major structural shock that poses the greater problem, and a particularly apt way to manage its ill effects is to give greater priority to the first adjudicator (whose decision preceded the intervening legal change) than to the second adjudicator (who decided after).

The fourth and final major value at stake is accuracy, especially by avoiding the risk of hindsight bias. Unlike the relatively universal jurisprudential values of repose and reliance and the particularly patent-salient problem of intervening legal changes, hindsight bias is quite specific to the patent system. It is most familiar in the analysis of whether an invention is, in fact, truly inventive—a requirement that is mediated primarily through the doctrine of nonobviousness. That is to say, to be a meaningful advance beyond the existing state of the art (and therefore be deserving of patent protection), an invention must not merely be new in the technical sense of being different from what has gone before. The invention must be sufficiently different—nontrivially different.

When undertaking this analysis, it is important not to be derailed by hindsight bias, the tendency to make factual findings about what happened at some time in the past by relying inappropriately (though perhaps unavoidably) on information that was not available at the time. Hindsight bias generally disfavors patent owners, as even inventions that actually were difficult to conceive may later seem simple or even trivial in retrospect, all the more if the success of the invention has made it familiar and commonplace.

346. *Id.* at 129.

347. Indeed, the same is also true of major judicial decisions such as those the Supreme Court has issued during its current period of renewed interest in patent law, especially in the area of patent-eligible subject matter. *See* Duffy, *supra* note 114, at 638–47.


Similarly, the risk of hindsight bias can also infect underlying assumptions about the technological field or discipline itself. Because inventions in patent law are defined by reference to the existing state of the art, both ex ante and ex post evaluation of an invention must limit itself only to what knowledge those in the field would consider relevant, and potentially to any analogous fields where they may reasonably be expected to look when searching for a solution to a technical problem. However, technical fields are not static but rather they grow and evolve, just as scientific research itself progresses. This propensity for change over time creates its own risk of hindsight bias. Indeed, given how foundational technological classification is to the process of evaluating patent validity, this sort of hindsight bias can quickly reach the point that it can make an outcome-determinative difference in a range of questions from the patentability of an invention, to the scope of the patent right, to whether a different good or service infringes the patent, and so on.

In both forms of hindsight bias, the ultimate implication for decisional design is the same. The closer that one can be to the point in time being evaluated, the less risk that hindsight bias will pose and the more accurate the decision will be. This, too, counsels in favor of giving priority to the first adjudicator, who is always closest in time to the relevant point of reference.

2. Priority to a Later Adjudicator

Although the reasons for preferring the first adjudicator are compelling, they are not conclusive. Both in litigation generally and the patent system specifically, there are reasons why a second or subsequent adjudicator may sometimes be better situated to produce sound decisions.

The more general set of reasons is connected to selection effects in litigation itself. These effects shape the incentives of parties and the outcomes of cases in ways that may accurately reflect the private value of resolving a given dispute over a given patent but not necessarily the social value of the resulting decision. The starting point for this intuition is that disputes enter the costly, protracted, and often inefficient fray of federal court only when two main conditions exist. First, there is sufficient uncertainty about the outcome of the case such that both parties are relatively optimistic about their respective chances for success. Second, the parties possess relatively good information about the respective costs of various outcomes and the direct transaction costs from litigation itself. Where these criteria are met,
both parties will tend to litigate so long as their risk-adjusted costs from losing or settling are lower than their risk-adjusted benefits from prevailing.

In patent litigation, this translates primarily into uncertainty over the validity or scope of the patent itself, with an overall cost-benefit calculation driven by the value of the patented technology relative to any available non-infringing market alternatives. The net effect is that two main forms of selection take place. Litigants enter the process nonrandomly either as plaintiffs (usually patent owners) or as defendants (usually accused infringers), and disputes persist to final judgment nonrandomly with most exiting through settlement along the way.

The impact of these selections can be profound where patents that are socially worthless or even harmful, at least from the perspective of some, continue to persist and proliferate because the parties who have the incentive to challenge them may not have the ability to do so, and vice versa. One particularly high-profile example from recent years is that of the Myriad Genetics patents on the BRCA1 and BRCA2 genes, which formed the basis for Myriad’s genetic diagnostic test for hereditary breast and ovarian cancer. The Patent Office had granted such patents on isolated gene sequences for many years, and companies in the industry had sued and defended a great many infringement litigations on those patents. However, because they all worked in the same space, such companies had no incentive to make the broad categorical argument that these inventions were ineligible for patent protection. That argument did not come before the courts until 2009, when a group of physicians and patients represented by the ACLU sought a declaratory judgment to that effect against Myriad’s patents, continuing without settlement until the Supreme Court finally concluded the case in 2013.

What these selection effects reveal is that a patent that has survived prior challenges may have survived because it is, in fact, a strong and meritorious patent—or it may have survived because of the nature of the arguments that were leveled against it and the incentives of the parties advancing those arguments. These may not be representative of all potential arguments and all potential challengers, and so there may be value in prioritizing the subsequent adjudicator over the earlier one. This allows subsequent adjudicators to benefit from contemporary conditions that may have been unavailable before due to earlier selection effects, but it does require leaving open a continuous opportunity to litigate and relitigate the same matter.

Another, more patent-specific reason to prioritize later adjudicators is that earlier decisions do not always arise from full information. This is certainly true of patent examination itself, which is theoretically an evaluation of inventions with respect to all relevant statutory requirements for patentability and all relevant evidence about the pre-existing state of the art. However, as the literature on examination’s shortcomings as well as the rational ignorance thesis concludes, resource constraints

---

360. See generally Marco & Vishnubhakat, supra note 39; Ford, supra note 39.
361. See generally Allison, Lemley, Moore, & Trunkey, supra note 38.
364. Dreyfuss, supra note 362, at 237.
365. Id. at 292–93.
and the potential for distortion in the incentives of the Patent Office mean that patents are not likely to emerge from a fully informed evaluation. Recent empirical research on the evidence asserted against patents in litigation bears out this conclusion. In particular, district courts that find patents invalid rely to a significant extent on prior art evidence that was not only absent from the patent examiner’s initial evaluation of the invention but would have been difficult or even impossible to find through the examiner search process.

Similar issues arise when comparing the first adjudicator to a later adjudicator of patents that have already been issued. Where relevant information about prior art, the technological history of the problem that the patented invention purported to solve, and the commercial success of the invention, may not be available to litigants in an initial dispute but may become available in a subsequent dispute, it would be more appropriate to give priority to the second adjudicator than to the first. However, though empirical research shows an information gap between ex ante patent examination and ex post patent validity review, there does not appear to be comparable support for an information gap between initial and subsequent ex post patent validity reviews. Thus, the case for prioritizing the first adjudicator’s decision about a patent’s validity is, on balance, stronger.

3. Priority to the Courts

As with the familiar preference for first adjudicators, the decisions of courts tend to enjoy primacy in U.S. law in general. In part, this is a longstanding touchstone of American legal culture itself, expressed as being “emphatically the province and duty of the judicial department to say what the law is” and the like. The argument from historical practice, however, has now also been under sustained criticism for quite some time—indeed, in terms that reflect the legislative motivations for the AIA itself: the judicial process is costly and lengthy, frequently inaccessible, and often inexpert where the subject matter of litigation is particularly complex or technical.

It is certainly true that the more common approach in contemporary institutional design is to allocate decision-making authority to other institutions, such as administrative agencies or legislative tribunals, while retaining eventual appellate review in the Article III courts as a backstop. Yet this is an answer to a different question of which institution should adjudicate an issue in the first instance due to its greater relevant competence. The question at hand is which institution we should listen to when multiple competing institutions all speak on the same issue.

In practical terms and for present purposes, there are two main arguments that favor choosing judicial decisions about patent validity over the decisions of other

366. See supra notes 24–29 and accompanying text.
369. See generally Or Bassok, The Court Cannot Hold, 30 J.L. & Pol. 1 (2014) (distinguishing between domains in which courts, especially the Supreme Court, might reasonably be thought to exercise expertise—such as constitutional interpretation—from expertise in the technical sense).
tribunals. One is the avoidance of distortion in individual case decisions through bias or capture in tribunals that sit within the political branches and are directly answerable to politically responsive forces. Another, related to the first, is the preservation of stability in the law and the continued ability of the patent system to offer credible commitments to would-be innovators and commercializers. The first is likely not a significant problem in the current system of ex post patent validity review, but the second does present important concerns.

The application of regulatory capture theory to patent law is a relatively recent trend in the literature, but a few salient conclusions have already emerged. One is that the Patent Office itself is at risk of capture, if not already captured, by industry forces, political actors, or both.\textsuperscript{371} The origins of this risk include the agency’s technical and epistemic barriers to entry by non-technologists\textsuperscript{372} and its resulting legal insularity from a range of other trans-substantive issues such as jurisdiction,\textsuperscript{373} jury review,\textsuperscript{374} judicial deference,\textsuperscript{375} and pleading procedure.\textsuperscript{376} Another is that the agency’s risk for capture is not especially greater than that of other agencies and that meaningful safeguards are both feasible and available.\textsuperscript{377} Still another is that courts themselves are as susceptible, or even more susceptible, to capture than the Patent Office,\textsuperscript{378} and the Federal Circuit is a case in point.\textsuperscript{379}

These conclusions are primarily from the context of ex ante patent examination, though some post-AIA analysis has extended the capture concern to administrative ex post review as well.\textsuperscript{380} Still, there does not yet appear to be significant analytical or empirical reason to think that the PTAB is captured in the sense of resolving

\begin{itemize}
  \item \textsuperscript{372} See generally Peter Lee, \textit{Patent Law and the Two Cultures}, 120 YALE L.J. 2 (2010).
  \item \textsuperscript{373} See generally Paul R. Gugliuzza, \textit{The Federal Circuit as a Federal Court}, 54 WM. & MARY L. REV. 1791 (2013).
  \item \textsuperscript{378} J. Jonas Anderson, \textit{Court Capture}, 59 B.C. L. REV. 1543 (2018).
  \item \textsuperscript{380} See generally Sherkow, supra note 371.
\end{itemize}
individual cases in systematic accordance with the preferences of a well-defined regulated industry or set of industries. The same is true of the ITC and other Article I tribunals.

The problems of stability and credible commitments are different. Even though the political-branch tribunals that adjudicate the validity of already-issued patents may not be captured, their modes of decision-making may nevertheless present concerns not merely that the substantive content of patent law is changing in one direction or another but that its rate of change is itself becoming unmanageably volatile. Achieving legal stability in the patent system is already fraught with difficulties. These difficulties include the familiar tension in patent law between predictable rules and flexible standards, the lack of durable principles—whether they take the form of rules or standards—in modern patent jurisprudence, the numerous problems of patent notice, and even simply of multifocal decision-making that requires users of the patent system to satisfy more than one legal authority before reaching certainty.

The structures of both the PTAB and the ITC compound these problems. Apart from technical expertise, which is written into the AIA as a requirement that PTAB judges “shall be persons of competent legal knowledge and scientific ability,” the political valence of PTAB decision-making has also recently come into greater focus. In a series of cases during the mid-2010s, the Patent Office leadership disapproved of PTAB panel decisions and reconfigured the panels with additional administrative patent judges to rehear the case and reach different outcomes that were more in line with the leadership’s preferences. The agency subsequently defended this “panel stacking” practice in multiple cases before the Federal Circuit, explaining that it was a mechanism for implementing the agency’s legal policy more uniformly, that the agency considered it permissible under Federal Circuit precedent, and that the way in which panels were actually stacked in practice did not injure the decisional independence of the PTAB. High-profile instances of panel stacking as well as skepticism from the Federal Circuit in oral arguments and published opinions, however, undermined the agency’s position.

The ITC, too, is a tribunal with considerable expertise and experience in resolving patent disputes. But this expertise notwithstanding, the ITC remains susceptible to political reversals of its decisions because it is subject to oversight by the President, whose authority in this regard is currently delegated to the United States Trade

381. See generally Vishnubhakat, supra note 1.
382. Id. at 1721–22.
383. Id. (citing Duffy, supra note 114).
384. Id. at 1722.
385. Id. at 1723.
386. 35 U.S.C. § 6(a).
387. Vishnubhakat, supra note 1, at 1676–77.
388. Id. at 1678–79.
389. Id. at 1679–80.
390. Id. at 1682–83.
391. Id. at 1683–84.
Representative (USTR). 392 USTR reversals of the ITC are rare, 393 but they are not unprecedented, and they reflect calculations that are explicitly political. 394

These institutional incentives to inject political priorities into questions of patent validity, which are otherwise defined primarily by technological fact and legal doctrine, do much to injure the ability of the patent system to offer a credible commitment that innovators and commercializers will reliably be able to recover their investments in the future. 395 The theory of credible commitments, derived from new institutional economics, holds that when an institution makes promises in order to induce others to behave and invest in certain ways, those promises are valuable, and thus credible, only to the extent that the institution cannot later change its decision in arbitrary or self-serving ways. 396

There is certainly strong and specific reason to expect that recent politically inflected practices of the Patent Office, such as PTAB panel stacking, are likely to injure the agency’s ability to make credible commitments. 397 The case that ITC decision-making similarly puts that body at the risk of weakened credibility is less specific but generally follows the same logic as that of the PTAB. To the extent that the ITC’s decisions are bound by stable legal precedent and the ITC itself is answerable to judicial bodies like the Federal Circuit, its decisional ability is constrained from changing in arbitrary or self-serving ways. But to the extent that its decisions are beholden to a more regularly changing cast of political actors such as the USTR or the President, the credibility of the ITC’s commitments may also fairly be called into question.

The sum of these institutional considerations is that actual regulatory capture of the PTAB or the ITC is not likely to present significant problems, but the disruptive influence of political decision-making in both settings, especially the PTAB, does raise serious concerns about legal stability in patent rights. That disruptive influence also raises concerns about the ability of patent institutions to make credible commitments and induce private actors to make choices about resource allocation in accordance with policy makers’ preferences for social welfare. Those concerns counsel in favor of giving priority to judicial decisions instead, which are relatively more insulated from political influence and thus relatively more credible in their commitments.

397. Vishnubhakat, supra note 1, at 1729–33.
4. Priority to Other Institutions

An important additional point regarding the comparison of Article III courts to non-Article III tribunals is that, prior to the AIA and historically, judicial process was the baseline of *ex post* patent validity review. Thus, the widely subscribed premise of the AIA in favor of transferring even more adjudicatory power out of the courts and into more expert and more accessible tribunals merits even greater credit in light of the entrenched status quo that had to be overcome. To put the point another way, even if the proponents and drafters of the AIA valued expertise and access highly and allocated power accordingly, it could not have been enough that the Patent Office was an expert agency, for that had always been the case. Rather, the promise and expected payoff of Patent Office expertise had to be so substantial as to produce a broad bipartisan and bicameral consensus.

That expertise is well documented. In the case of the Patent Office prior to the AIA’s enactment, the agency’s expertise had generally been manifested in three activities: classifying technologies to enable patent examination, denying patents, and granting patents. Classification reflects the least controversial—though, for a long time, also the most easily overlooked—of these. All major requirements for patentability as well as several doctrines pertaining to infringement liability and remedies rely on antecedent definitions of the relevant field of technology in which the invention sits.\(^{398}\) The Patent Office, for its part, has extensive and longstanding expertise in developing, maintaining, and periodically updating a comprehensive taxonomy of technological fields and of classifying individual inventions within that taxonomy.\(^{399}\) The source of this expertise, that Congress expressly delegated the authority to the Patent Office in 1898, is also its strongest justification.\(^{400}\)

The agency also exercises its expertise in actually examining patents, naturally, and the greater portion of its expertise is likely to be reflected in denying patents, more than it is in granting patents, because the statutory presumption of examination is that the patent should issue.\(^{401}\) Because the agency bears the burden of disproving patentability, its denials are more likely to reflect the substance of the agency’s technical expertise than are its grants,\(^{402}\) which may at times reflect little more than the agency’s resource-constrained inability to push back adequately against a persistent applicant.\(^{403}\) Moreover, on the question of which institution’s expertise is best exercised at which stage of decision-making, cases in which the complexity to

---

399. *Id.* at 938.
400. *Id.* at 939 (citing 35 U.S.C. § 6 (1946)).
401. 35 U.S.C. § 102(a) (providing that the applicant “shall be entitled to a patent unless” the applicant fails to overcome various hurdles set forth in the statute).
402. Arti K. Rai, *Allocating Power over Fact-Finding in the Patent System*, 19 *BERKELEY TECH. L.J.* 907, 912 (2004) (arguing that “when the USPTO denies a patent, the fact-finding associated with the USPTO’s analysis is much more likely to be accurate. This is because the patent examiner has the burden of demonstrating the unpatentability of the applicant’s assertions, and thus the examiner is required to assemble evidence supporting the rejection”).
403. *See supra* notes 20–25 and accompanying text.
be resolved is in the technological facts themselves are usually best left to the expert fact-finder in the first instance, i.e., the agency.404

Meanwhile, the ITC is also an expert agency of a sort.405 Unlike administrative patent judges of the PTAB, administrative law judges and commissioners of the ITC are not required to have scientific training or doctrinal familiarity with patent law.406 Nonetheless, they have tended to preside over a high volume of patent-related cases.407 In fact, that volume has only risen over time as the ITC’s prominence has grown in the wake of the eBay v. MercExchange decision.408

These extensive indicia of expertise on the part of non-Article III tribunals prior to the AIA were good reason to expect that the same expertise would offer substantial benefits in the ex post patent validity review proceedings that the AIA envisioned. In fact, the well-rehearsed pathologies of patent litigation arose largely from the lack of expertise among lay judges and juries. If litigation in the federal courts was expensive and protracted, it was because attorneys had to invest significant amounts of money and time in teaching generalist decision makers about the relevant technology and the relevant law or else risk inaccurate or otherwise distorted outcomes.409

Apart from the familiar and well-supported invocation of expertise, preferring other institutions than federal courts also offered a means for broadening access to patent validity challenges410 and for reducing the collective action problem that had long plagued traditional judicial review.411 The problem of access stems from the need for Article III standing, which permits the invocation of federal judicial power only in genuine cases or controversies.412 In this regard, a generalized grievance that a patent was, in the view of the challenger, wrongly granted is not likely to confer standing to sue whereas a more specific showing that the wrongly granted patent poses actual or imminent injury to the challenger probably is.413

Where the line of demarcation lies for showing an injury-in-fact sufficient for Article III standing is, of course, contested and difficult to predict as with any other flexible standard of a constitutional dimension. Where there is at least some agreement, however, is that the Federal Circuit’s precedents on standing in disputes involving patent law have been unduly restrictive and that even the Supreme Court’s 2007 intervention in MedImmune v. Genentech did little to dislodge the Federal Circuit’s hidebound jurisprudence.414 These patent-specific standing rules may fall

405. Kumar, supra note 379, at 1547.
406. Id. at 1590.
407. Id.
408. Chien & Lemley, supra note 73, at 2–3.
409. Vishnubhatkat, supra note 1, at 1671; Strategic Decision Making, supra note 2, at 51–55.
411. Id. at 54.
especially hard on economically disempowered segments of the market, such as individual consumers. By contrast, non-Article III tribunals such as the PTAB and the ITC are not subject to the case-or-controversy requirement and need not impose any requirement of standing upon parties seeking relief. In the PTAB, no standing is required at all: anyone “who is not the owner of a patent” may seek administrative review challenging the validity of a patent. In the ITC, meanwhile, the level of standing necessary is rather minimal: a patent owner may bring suit only if a domestic industry that is related to the patented invention “exists or is in the process of being established.”

The result is substantially greater access for litigants in these alternative tribunals as compared to their corresponding access in the federal courts. In fact, this is true even though subsequent appeals to the Federal Circuit still require Article III standing, which a litigant at the margin would not have enjoyed in the first instance. For appeals from the PTAB, the Federal Circuit thus far has rejected arguments that a petitioner who otherwise lacks all Article III standing can acquire a sufficient injury-in-fact simply by losing in the administrative proceeding below. Instead, the court has held that the AIA creates only a procedural right to seek review, a right that is satisfied regardless of whether the challenge is ultimately successful or not, and that denying a patent validity challenge by a party that has no personal stake in making, selling, or using the patented invention does not invade any legal right.

On first impression, it may seem that the Federal Circuit’s case law on appellate standing dissipates whatever benefit might have accrued from the fully permissive access that the AIA confers in the PTAB below. But this is not so. Patent owners, who come before the PTAB in a defensive posture to argue against the revocation of their patent rights, can satisfy the Consumer Watchdog standard—and, indeed, the more general injury-in-fact jurisprudence—quite easily by pointing to any adverse effects on their patent rights. It is only unsuccessful patent challengers who may not be able to appeal, Ironically disadvantaging “those very parties who benefit the most at the margins from having access to administrative review—precisely because they could not otherwise have established standing to seek declaratory relief in federal court.”

---

420. Id. at 1261–62.
422. Id.
judgment from the PTAB, the successful challenger can safely rely on the appellant-patent owner’s standing to enter the Federal Circuit. Thus, of all the non-Article III litigants who can enter the system through the PTAB, only a subset (those who lose) are currently excluded from continued access. This is still an improvement over the status quo ante and is another reason to prioritize non-Article III courts.

The problem of access is closely connected to the problem of collective action in mounting challenges to patent validity: even those who do possess Article III standing are frequently at a disincentive to use it. While litigation is expensive, the investment of an individual challenger in revoking a patent is nearly impossible to recoup because of *Blonder-Tongue Laboratories, Inc. v. University of Illinois Foundation.* The estoppel effect of a patent invalidation makes a successful patent challenge a public good: the challenge is nonrival and nonexcludable from all others, including those who played no part in supporting the patent validity challenge.

Such goods are generally undersupplied, reflecting a collective action problem that makes would-be challengers unwilling to undertake *ex post* patent validity review for fear of inviting others to “free ride from the efforts of the former patentee and the opponent, employing the teachings of the invalidated patent to practice the invention without compensation to anyone.” PTAB review, given its lower cost, does not remove the collective action problem but likely does reduce its scale. This, too, is a further reason to give priority to non-Article III courts.

Yet the comparative institutional arguments advanced here—which weigh the deeper expertise and broader access of agency tribunals against the greater legal stability and more credible commitments of federal courts—still prove too little about the actual design of the AIA. Suppose that Congress was right to prioritize agency expertise because it leads to cheaper, faster, and more accurate outcomes. And suppose that Congress was also right to remove any standing requirement because questionable patents have no valid claim to survival merely by limiting the pool of people who can challenge them. These departures from the pathologies of litigation both address the quality of decision-making about patent validity vis-à-vis who decides and what arguments and evidence are brought forth.

These emphases on decisional quality are legitimate because they are content-neutral as to the patents being evaluated. The aforementioned intractable premise of the AIA, that easing the invalidation of questionable patents requires easing the invalidation of all patents, can be managed at least to this extent. Though it is unavoidable that all patents will now be easier to revoke as a practical matter, the reason why they are easier to revoke is still a valuable policy lever to help differentiate patent quality reform that is genuinely concerned with weeding out poor-quality patents while leaving meritorious patents alone. However, if the

424. See supra notes 138–139 and accompanying text.
425. Miller, supra note 82, at 688–89.
427. Vishnubhakat, supra note 1, at 1696–1697; Strategic Decision Making, supra note 2, at 54–55.
428. See supra note 142 and accompanying text.
objective is a more indiscriminate attack on patent rights with relative indifference to leaving meritorious patents intact, then the particular reasons why all patents are easier to invalidate are far less important.

The current structure of *ex post* patent validity review in the PTAB suggests the latter. In addition to deeper expertise from agency adjudicators and broader access for properly incentivized challengers, the greater risk of invalidation in PTAB review has also arisen, until recently, from giving patent claims their broadest reasonable interpretation.\(^ {429}\) And the greater risk of invalidation continues to arise even now from failing to give patents a presumption of validity and failing to require that invalidity be proven by clear and convincing evidence.\(^ {430}\) These are further deviations from the pre-AIA baseline of federal court litigation with no basis either in the expertise of the decision-maker or in the proper incentives of would-be challengers. They are not content-neutral with respect to the patents being challenged and are equally as likely to adversely affect a meritorious patent as they are to weed out an unmeritorious one.

**B. Legislative Prescriptions**

By decoupling and separately considering the design parameters of timing and institutional competence, the foregoing discussion revealed two important insights. One is that timing, standing alone, tends to favor the first adjudicator. This might provoke concern that large swaths of outcomes from PTAB review, especially *inter partes* review, have been the result of poor institutional design. But that conclusion does not follow from timing alone. The discussion of institutional competence, after all, makes a fairly strong case for favoring courts but recognizes the significant advantages of agency adjudication as well, readily accommodating the AIA’s actual preference for the expert agency. The second insight, that even expertise and access in non-Article III adjudication leave important features of the PTAB lacking in justification, is what calls the current design of the AIA into serious question—while also narrowing the scope of concern.

To grapple with that question, this subsection offers specific legislative reform proposals, informed in part by lessons from trademark law. Patents in PTAB reevaluation should enjoy the same statutory presumption of validity and the same requirement for invalidity to be proven by clear and convincing evidence as in the courts.\(^ {431}\) Statutory estoppel between the courts and the PTAB should be bilateral, rather than binding the courts while leaving the agency free to disregard judicial decisions. And patent claim construction in the PTAB should be subject, by statute, to the standard set forth in *Phillips v. AWH Corp.*\(^ {432}\)

\(^{429}\) See supra Section II.B.2.\(^ {430}\) See supra Section II.A.1.\(^ {431}\) See supra Section II.A.1.\(^ {432}\) 415 F.3d 1303 (Fed. Cir. 2005).
1. Aligning the Administrative Burden of Proof

Similarly, and with greater reason for statutory amendment, the presumption of validity, as well as the burden and standard of rebutting that presumption by clear and convincing evidence, should also be the same in the agency as they have long been in the courts and those presumptions of validity were recently reaffirmed by the Supreme Court in *Microsoft Corp. v. i4i Limited Partnership*. In enacting the AIA, Congress dealt specifically with the standard of proof and chose a mere preponderance of the evidence as the standard. This was a mistake of policy by Congress for the same reason that the initial use of broadest reasonable interpretations was a mistake of policy by the Patent Office. Although the agency did, and does, have significant virtues of superior expertise and access relative to the federal courts, those virtues still do not justify the use of divergent substantive standards for evaluating patentability. Not only does this disparity allow patents that would otherwise have survived judicial review to be struck down, but it also tolerates and even encourages numerous instances in which patents actually did survive prior judicial review only to fall in the PTAB as the result of legal arbitrage.

It is no reply, moreover, to say that if aligned standards of proof are all that matter, then the courts, as well as the Patent Office (and, for that matter, the ITC), could all turn instead to a preponderance standard. Before the enactment of the AIA, this very question came before the Court in *Microsoft v. i4i*. In that case, i4i had sued Microsoft for allegedly infringing i4i’s patent on an improved method for computer document editing, and Microsoft responded by asserting that i4i’s patent was invalid. There was no doubt that the patent examiner during examination had never considered the prior art that Microsoft now offered as evidence that i4i’s patent was invalid, but there was vigorous dispute over whether that prior art would actually have any invalidating effect on the i4i patent. Microsoft argued that it should not be required in court to prove invalidity by clear and convincing evidence but rather that a mere preponderance should be enough where the evidence offered was evidence that had not been before the examiner.

The Supreme Court disagreed, explaining that the common law background against which Congress had codified the presumption of validity in the 1952 Patent Act included an understanding that the presumption was “not to be overthrown except by clear and cogent evidence.” That background understanding, in turn, reflected “nearly a century of case law from [the Supreme Court] and others” supporting a near-universal consensus “that a preponderance standard of proof was too ‘dubious’ a basis to deem a patent invalid.” For the same reason, it would be prudent now to revise the PTAB’s statutory standard of proof in *ex post* patent

435. 564 U.S. at 92.
436. Id. at 98.
437. Id.
438. Id. at 98–99.
439. Id. at 101 (citing Radio Corp. of Am. v. Radio Eng’g Lab’ys., Inc., 293 U.S. 1 (1934)).
440. Id. at 101–02.
validity review proceedings to require clear and convincing evidence, just as the courts require.

2. Making Interbranch Estoppel Bilateral

To put these proposed reforms into perspective, it is helpful to consider the example of trademark law. Just as patents are subject to \textit{ex post} review of their validity in the PTAB, so also are trademark registrations subject to review in the Trademark Trial and Appeal Board (TTAB). The TTAB is also an administrative alternative to judicial resolution, and once the registration is canceled, the revocation of the trademark registration is also subject to nonmutual defensive collateral estoppel—i.e., the registration is thereafter invalid as against the world.

These analytical similarities between the PTAB and TTAB extend to the paradigm of court-agency substitution as well. Section 14 of the Lanham Act authorizes the TTAB to hear a cancellation proceeding,\textsuperscript{441} and Section 37 authorizes the federal courts to do the same.\textsuperscript{442} Though a bare claim for judicial cancellation is disfavored, this is only in the jurisdictional sense that Section 37 does not create an independent right of action. Instead, such a claim must either be presented in response to a suit for infringement or must accompany a preemptive suit for declaratory judgment by the challenger.\textsuperscript{443} Ultimately, then, this fully mirrors the two possible postures for judicial invalidation of patents as well.

However, unlike the \textit{ex post} reevaluation of patent rights, court-agency substitution in reevaluating the validity of trademark registrations is, in important respects, considerably simpler. For example, there is no meaningful distinction between standard and nonstandard petitioners in the TTAB. In the PTAB, \textit{inter partes} review is primarily a forum for standard petitioners: about 70\% seek review in defensive response to an infringement lawsuit already filed against them in the federal courts.\textsuperscript{444} What the remaining 30\%, who are nonstandard petitioners, seek instead is preemptive review akin to declaratory relief.\textsuperscript{445} Covered business method review, meanwhile, is available only as a defensive response to infringement litigation, making nonstandard petitioning impossible.

There is also no statutory constraint between parallel proceedings in the TTAB and the federal courts as there is for the PTAB. The one-year time bar for bringing an \textit{inter partes} review in the PTAB does not exist for bringing a cancellation under Section 14 in the TTAB. The discretionary power of courts to grant stays (or not) pending resolution of a parallel agency proceeding is the same in both contexts.

\textsuperscript{441} 15 U.S.C. § 1064.
\textsuperscript{444} See Strategic Decision Making, supra note 2, at 73; see supra text accompanying note 61.
\textsuperscript{445} See Strategic Decision Making, supra note 2, at 73; see supra text accompanying note 61.
Indeed, the only substitutionary influence in trademark cancellation is the preclusive power of whichever tribunal decides first, the TTAB or the court.

Finally, and perhaps most importantly, Section 14 proceedings provide a two-stage structure for seeking administrative cancellation in the TTAB, akin to what post-grant review provides prospectively for patents issued under the first-inventor-to-file provisions of the AIA. For trademarks that have been registered for fewer than five years, cancellation is available on the basis of any of the substantive statutory requirements for registrability under Section 2 of the Lanham Act. These include showing that the mark does not contain matter that

(1) is immoral, deceptive, scandalous, or disparaging; (2) consists of a national, state, or municipal flag or insignia; (3) consists of the name, portrait, or signature of certain individuals; (4) is likely to cause confusion with existing marks; (5) is merely descriptive, misdescriptive, functional, or primarily merely a surname; or (6) would likely cause dilution by blurring or tarnishment.

Cancellation is also available on the further grounds that the trademark “has become generic, functional, or abandoned; was obtained fraudulently or contrary to the provisions of sections 2(a), (b), or (c); or is being used to misrepresent the source of the goods or services.” For trademarks that have been registered for more than five years, only these further grounds are available.

This mirrors the two-stage process for challenging post-AIA patents, which are subject to broad challenges in PGR for the first nine months after their issuance and are subject thereafter to narrower challenges in IPR and CBM. However, the myriad additional structural differences among PGR, IPR, and CBM are not present in Section 14 cancellation proceedings, leaving a much simpler agency substitute in trademark law.

This pattern of simplicity extends beyond the contours of adjudicatory power itself to post-adjudication estoppel as well. A final judgment in a cancellation proceeding in the TTAB under Section 14 precludes relitigation of the same issues in future TTAB proceedings as well. Similarly, as the Supreme Court recently held in B&B Hardware v. Hargis Industries, preclusion from a final judgment in a TTAB cancellation proceeding extends to the federal courts as well. The facts of B&B Hardware are analogous to administrative patent validity review not only in terms of court-agency substitution but in legal substance as well.

B&B had registered the trademark SEALTIGHT for metal fasteners in the aerospace industry, and Hargis, three years later, tried to register the trademark SEALTITE for metal screws in metal and post-frame buildings. Concerned with potential consumer confusion, B&B asserted its priority by opposing Hargis’s

447. Vacca, supra note 443, at 303–04 (summarizing 15 U.S.C. § 1052(a)–(f)).
448. Id. at 304 (summarizing 15 U.S.C. § 1064(3)).
449. See supra notes 52–58 and accompanying text.
450. ALTMAN & POLLACK, supra note 443, § 23:37.
452. Id. at 144–46.
registration for SEALTITE in a pre-issuance proceeding before the TTAB. The TTAB agreed that Hargis’s proposed SEALTITE trademark would have been confusingly similar to B&B’s SEALTIGHT trademark and ruled in favor of B&B. Meanwhile, B&B had also sued Hargis for infringing B&B’s trademark registration on SEALTIGHT, and the TTAB reached its conclusion while that court case was still pending. B&B then argued to the district court that the TTAB’s conclusion foreclosed Hargis from disputing that the trademarks were confusingly similar. However, the district court rejected B&B’s estoppel argument on the grounds that the TTAB was not an Article III court and so could not bind the district court by preclusion. The Supreme Court reversed, explaining that when duplicative litigation arises in multiple tribunals, “the decision of the first tribunal usually must be followed by the second, at least if the issue is really the same”—and that to do otherwise “wastes litigants’ resources and adjudicators’ time, and it encourages parties who lose before one tribunal to shop around for another.”

But just as importantly, and since long before B&B Hardware, the reverse has also been true. A final judgment in a judicial cancellation proceeding under Section 37 precludes relitigation of the same issues in future TTAB proceedings. This is all subject to the ordinary requirements of preclusion that the issue was “actually litigated and determined by a valid and final judgment” and that “the determination is essential to the judgment.” Thus, what is particularly telling about the example of bilateral court-agency estoppel in trademark law is not B&B Hardware, though that opinion was certainly a significant surprise to many in both the administrative law and intellectual property fields.

For illuminating the institutional design of the PTAB, the question that B&B Hardware answers is merely about how final administrative judgments influence subsequent judicial proceedings. On that question, the AIA provides an express statutory answer in § 315(e)(2) for inter partes review and in § 325(e)(2) for post-grant review and covered business method review. Instead, the problematic duplication in ex post patent validity review has arisen from the Patent Office ignoring the prior judicial conclusions of the federal courts. And in this regard, the lesson of trademark law is clear and longstanding: even if the first adjudicator is a generalist court and the second adjudicator an expert agency, the prior decision controls.

453. Id.
454. Id. at 146–47.
455. Id.
456. Id.
457. Id.
458. Id. at 139–42.
459. Id.
460. RESTATEMENT (SECOND) OF JUDGMENTS § 27 (AM. L. INST. 1982).
3. Codifying Court-Aligned Claim Construction

Detailed analysis of the two key aforementioned design parameters—timing\(^\text{461}\) and institutional competence\(^\text{462}\)—suggests that if all else is equal, then the analysis and decision of the first tribunal to adjudicate patent validity should control. However, if all else is not equal because, say, the first tribunal was merely a generalist court and the second tribunal was an expert agency with more robust access for properly incentivized litigants, even then, the relative virtues of the agency still do not justify the use of divergent substantive standards for evaluating patentability. The method for construing the patent claims should be the same in the agency as it previously was in the court, the *Phillips v. AWH Corp.* standard.

As of late 2018, the Patent Office has expressed agreement with this view and has put into effect a rule formally ending the PTAB’s approach of giving claims their broadest reasonable interpretation and instead using the *Phillips* approach of the district courts.\(^\text{463}\) This was a sound and welcome change, one that was met with strong support among law and economics scholars of the patent system.\(^\text{464}\) The twenty-one-professor response especially highlighted that divergence in the standards of claim construction—and of the standards of proof for establishing invalidity—was the key hindrance to generating preclusive effect for the first tribunal that decides on patent validity.\(^\text{465}\) Removing these divergences, beginning with a unified court-agency standard for claim construction, offered “an opportunity to reduce needless duplicative litigation, [and w]hen taken, that opportunity would eliminate the associated costs for the USPTO, the courts, and the parties alike.”\(^\text{466}\)

The Patent Office did take this opportunity, but what one agency administration gives, the next might take away. The 2018 rule change itself was just that sort of reversal, for the agency’s initial choice of the broadest reasonable interpretation had been quite purposeful. One of the earliest challenges against the rules package that embodied PTAB practice had targeted the use of broadest reasonable interpretation in *inter partes* review proceedings, and the Patent Office had defended its choice—successfully—all the way to the Supreme Court in *Cuozzo v. Lee*.\(^\text{467}\) Therefore, it would be prudent during the current season of reform to entrench the *Phillips* approach into PTAB review by statute in order to ensure its continued vitality.

CONCLUSION

The institutional design problem of seeking efficient case resolutions in patent law by substituting agency expertise for the generalism, cost, delay, and error of

---

466. *Id.* at 7.
judicial process is a perennially difficult one. The AIA and its administrative trial proceedings reflect the most ambitious implementation of the substitution thesis in the history of U.S. patent law. As that system enters into its second decade, it is necessary and proper to consider what problems still remain and to identify where they arise.

Likewise, in pressing for reform, it is important that the corrections reflect the lessons of past judicial victories and defeats as well as the results of prior experiments in regulatory policy. Most importantly, structural change to the design of court-agency substitution should be resilient and sustainable beyond the political leadership of the day.

To that discourse, this Article has made three principal contributions. Descriptively, it offers the most extensive empirical account to date of duplication in PTAB litigation. Analytically, it offers a systematic explanation of the statutory and institutional sources of that duplication. Normatively, it concludes with a framework for rethinking the substitution thesis and with targeted legislative proposals to modernize the PTAB for the new decade.

**Tables and Figures**

*Figure 1.* Grant Rates of Contested Motions to Stay U.S. District Litigation Pending *Inter Partes Review*, 3-year Average, 2012–2020
Figure 2. Distribution of Computers & Communications-Related Patents Across Five High-Volume Patent Districts

Figure 3. Distribution of Electrical & Electronic-Related Patents Across Five High-Volume Patent Districts
Figure 4. Distribution of Mechanical-Related Patents Across Five High-Volume Patent Districts

Figure 5. Distribution of Chemical-Related Patents Across Five High-Volume Patent Districts
**Figure 6.** Distribution of Drugs & Medical-Related Patents Across Five High-Volume Patent Districts

**Figure 7.** Motions to Amend in *Inter Partes Review*, Retrospective 12-month Average

---

*En banc* rehearing granted in *Aqua Products* (Aug. 2016)
Figure 8. Outcomes of Motions to Amend in Inter Partes Review, Retrospective 12-month Average

Figure 9. The Immediate Impact of SAS Institute on Petitioner Filing Behavior (4-month windows immediately before and after April 24, 2018)

<table>
<thead>
<tr>
<th>Disputes (defined as a set of petitions between the same petitioner and patent owner)</th>
<th>Petitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre- vs. Post-SAS Institute (aggregate of IPR, CBM, and PGR filings)</td>
<td>$-6.3%$</td>
</tr>
<tr>
<td>Pre- vs. Post-SAS Institute (IPR filings only)</td>
<td>$-2.2%$</td>
</tr>
</tbody>
</table>

PTAB guidance issued after *Aqua Products* (Nov. 2017)
Figure 10. Distribution of Lag Between First IPR Petition and First Federal Court Litigation, in Years
Figure 11. Distribution of Lag Between First IPR Petition and Last Federal Court Filing Prior to Petition, in Years